



Exploitation de la couleur pour la segmentation et l'analyse d'images

Ludovic Macaire

► To cite this version:

Ludovic Macaire. Exploitation de la couleur pour la segmentation et l'analyse d'images. Interface homme-machine [cs.HC]. Université des Sciences et Technologie de Lille - Lille I, 2004. tel-00386862

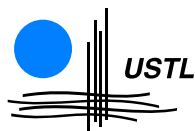
HAL Id: tel-00386862

<https://theses.hal.science/tel-00386862>

Submitted on 22 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université des Sciences et Technologies de Lille

Laboratoire LAGIS

UMR-CNRS-8146



EXPLOITATION DE LA COULEUR POUR LA SEGMENTATION ET L'ANALYSE D'IMAGES

Mémoire présenté
en vue de l'obtention de

l'Habilitation à Diriger des Recherches

par

Ludovic Macaire

Maître de Conférences à l'UFR d'I.E.E.A.

Soutenu publiquement le 19 novembre 2004 devant le jury composé de :

Jean-Marc CHASSERY	Directeur de Recherche au LIS de Grenoble	Rapporteur
Jean-Michel JOLION	Professeur à l'I.N.S.A de Lyon	Rapporteur
Sylvie PHILIPP-FOLIGUET	Professeur à l'E.N.S.E.A.	Rapporteur
Olivier COLOT	Professeur à l'U.S.T.L.	Examineur
Christine FERNANDEZ-MALOIGNE	Professeur à l'Université de Poitiers	Examineur
Jack-Gérard POSTAIRE	Professeur à l'U.S.T.L.	Examineur
Alain TRÉMEAU	Professeur à l'Université de Saint-Étienne	Examineur

Remerciements

Pendant ces quinze années d'activités professionnelles à l'Université des Sciences et Technologies de Lille, j'ai été constamment soutenu par l'amitié et la confiance de mon directeur de thèse, Jack-Gérard Postaire. Après m'avoir ouvert la porte de l'enseignement supérieur et de la recherche, il m'a tout appris. Je lui en suis extrêmement reconnaissant.

Messieurs Pierre Vidal, Christian Vasseur et Philippe Vanheeghe, directeurs des Laboratoires dans lesquels j'ai travaillé, m'ont accueilli aux premiers jours de ma thèse et ont été témoins de tout ce parcours. Ils m'ont donné les moyens de travailler dans d'excellentes conditions, je les en remercie.

Je remercie Madame Sylvie Philipp-Foliguet et Messieurs Jean-Marc Chassery et Jean-Michel Jolion d'avoir accepté la tâche ingrate de rédiger un rapport sur ce manuscrit.

Madame Christine Fernandez-Maloigne et Monsieur Alain Trémeau ont parcouru des centaines de kilomètres pour participer à ce jury. J'en suis très touché et les en remercie chaleureusement. Je remercie Monsieur Olivier Colot d'avoir accepté d'assumer le rôle de président de jury.

Durant toutes ces années, je me suis partagé entre l'enseignement et la recherche, aussi je tiens à remercier Pierre Bonnet pour m'avoir entraîné vers quelques aventures pédagogiques dans la joie et la bonne humeur.

Tous les travaux présentés dans ce mémoire n'auraient pas vu le jour sans le travail des doctorants que j'ai eu le plaisir d'encadrer. Je remercie Vincent Ultré, Vénérée Rakotomalala, Nicolas Vandenbroucke, Michael Fontaine et Damien Muselet pour la confiance qu'ils m'ont témoignée.

Je ne suis jamais dans mon bureau. Je préfère travailler dans la salle commune où je retrouve Christophe, parfois Claudine, François et Luc. Ce ne sont pas des collègues, mais des amis.

Nombreuses sont les personnes qui m'ont donné d'indispensables coups de main au cours de ces années. J'espère ne pas avoir oublié chaque fois de les remercier. Si ce n'est pas le cas, qu'elles me pardonnent aujourd'hui.

Table des matières

Remerciements	3
Table des matières	7
Avant-propos	9
Notice individuelle	13
Activités d'enseignement	17
Laboratoires de rattachement	23
Encadrement et jurys de thèse de Doctorat	25
Rayonnement scientifique	29
Publications	33
1 Introduction	37
1.1 Contexte des travaux	37
1.1.1 Formation d'une image couleur	38
1.1.2 Chaîne de traitements des images couleur	45
1.1.3 Codage des systèmes de représentation de la couleur	45
1.1.4 Exploitation de la couleur pour la segmentation des images	47
1.1.5 Exploitation de la couleur pour la reconnaissance d'objets	55
1.2 Travaux présentés	57
1.2.1 Systèmes de représentation de la couleur et segmentation	58
1.2.2 Régions de l'image et classes de pixels	64
1.2.3 Reconnaissance d'objets sous éclairage non contrôlé	67
1.3 Organisation du mémoire	74
1.4 Bibliographie	74
2 Espaces couleur adaptés à la classification de pixels	79

2.1	Introduction	79
2.2	Construction de l'espace couleur hybride adapté	80
2.2.1	Échantillon d'apprentissage	80
2.2.2	Espace couleur hybride candidat	81
2.2.3	Construction de l'espace couleur hybride adapté	82
2.2.4	Critère d'arrêt	85
2.3	Application à la segmentation d'images de foot-ball	86
2.3.1	Apprentissage supervisé	87
2.3.2	Classification des pixels-joueurs	89
2.3.3	Résultats	90
2.4	Conclusion	102
2.5	Bibliographie	102
3	Segmentation par classification spatio-colorimétrique des pixels	105
3.1	Introduction	105
3.1.1	Propriétés de connexité	106
3.1.2	Plan du chapitre	107
3.2	Degré de connexité et d'homogénéité d'un sous-ensemble-couleur	108
3.2.1	Sous-ensemble couleur	108
3.2.2	Degré de connexité	109
3.2.3	Degré d'homogénéité	113
3.2.4	Degré de connexité et d'homogénéité	115
3.3	Construction des classes de pixels	116
3.3.1	Classes définies par des domaines-couleur	117
3.3.2	Description de la procédure	118
3.3.3	Sélection des domaines-couleur candidats	120
3.3.4	Construction des classes de pixels	123
3.4	Résultats	128
3.4.1	Les images	129
3.4.2	Évaluation qualitative de la qualité de segmentation	135
3.4.3	Évaluation quantitative de la qualité de segmentation	135
3.5	Conclusion	137
3.6	Bibliographie	140
4	Reconnaissance d'objets sous éclairage non contrôlé	141
4.1	Introduction	141
4.1.1	Reconnaissance d'objets éclairés par différents illuminants	141

4.1.2	Plan du chapitre	142
4.2	Hypothèses basées sur les mesures de rang	143
4.2.1	Hypothèse de conservation des mesures de rang	144
4.2.2	Hypothèse des plus proches mesures de rang	146
4.3	Histogrammes couleur égalisés	147
4.4	Histogrammes couleur spécifiques aux couples d'images	150
4.4.1	Couple d'histogrammes couleur spécifiques	150
4.4.2	Couple de tables mono-dimensionnelles de correspondance	151
4.4.3	Sélection des couples de niveaux	152
4.4.4	Exemple	154
4.5	Résultats expérimentaux	156
4.5.1	Reconnaissance d'objets de la base de l'Université Simon Fraser	156
4.5.2	Reconnaissance d'objets de la base de l'Université East Anglia	158
4.5.3	Discussion	161
4.6	Conclusion	161
4.7	Bibliographie	162
Conclusion et perspectives		165
Publications choisies		167
Color image segmentation by analysis of 3D histogram with fuzzy morphological filters		169
Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis		195
Color person image indexing and retrieval		223
Identification of retinal vessels by color image analysis		245
Determination of compatibility coefficients for color edge detection by relaxation		269

Avant-propos

Ce mémoire retrace mes travaux effectués depuis mon arrivée en 1989 au Centre d'Automatique de Lille de l'Université des Sciences et Technologies de Lille. L'origine de ce rapport date probablement du jour où l'on m'a posé la question :

"Quand passes-tu ton HDR ?"

Au début de ma carrière, peu d'éléments m'incitaient à y répondre favorablement.

Jugez plutôt.

Une thèse industrielle en convention CIFRE en collaboration avec une PME sidérurgique n'est pas la voie la plus appropriée pour obtenir un poste de Maître de Conférences. Mais la découverte du monde de la recherche a été pour moi, qui me destinait à une carrière d'ingénieur, une révélation. Par un heureux concours de circonstances, les effets conjugués d'un manque de candidats et d'un nombre important de postes à pourvoir m'ont permis d'être nommé en 1993 sur un poste de maître de conférences à l'UFR d'IEEA de l'Université des Sciences et Technologies de Lille.

Suivent alors deux années plus consacrées à l'animation d'une crèche parentale (deux bébés en 14 mois...) qu'à la rédaction de publications. Il est vrai également qu'en 1995 mon sujet de thèse, le contrôle-qualité de produits plats en défilement continu par l'analyse temps réel d'images ligne en niveaux de gris, commence à passer du domaine de la recherche à celui du transfert technologique. Si au niveau de ma formation à l'innovation et à la recherche et développement en entreprise, mon expérience CIFRE a été très bénéfique, le cadre de travail sur les lignes de production était éloigné des préoccupations académiques des chercheurs universitaires.

C'est par hasard que Vincent Ultré, doctorant au Centre d'Automatique de Lille de 1994 à 1996, me parle de ses travaux consacrés à la segmentation d'images couleur. Quel changement ! Je passe de l'analyse d'images lignes en niveaux de gris à celle d'images matricielles couleur...

Ensemble, nous découvrons la spécificité des problèmes posés par l'exploitation de la couleur pour analyser des images numériques. Nous tentons de comprendre pourquoi des méthodes classiques de classification de données multi-dimensionnelles ne permettent pas d'obtenir une segmentation pertinente d'images que pourtant nous jugeons visuellement comme étant "simples à segmenter". Le mémoire de thèse de Vincent Ultré pose à ce sujet plus de questions qu'il n'apporte de réponses.

Des éléments de réponse fort pertinents nous sont fournis par l'un des rapporteurs de la thèse de Vincent Ultré, Jean Bajon, Professeur à l'ENSEEIH de Toulouse et spécialiste de l'acquisition d'images couleur. Les images considérées étaient tout simplement acquises sans ajustement préalable des paramètres d'acquisition, rendant difficile la discrimination entre les classes de pixels en présence dans les images.

Jean Bajon juge, à juste titre, que notre travail n'est pas du niveau d'une thèse de doctorat et refuse dans un premier temps de rédiger son rapport sur le manuscrit de la thèse de Vincent Ultré. Je ne peux pas dire que ma carrière de co-encadrant de Doctorants commence à ce moment-là par un franc succès...

Jean Bajon a la gentillesse d'accepter de nous rencontrer et de répondre à nos questions de candides sur la formation des images numériques couleur. Il nous incite à retravailler le manuscrit et surtout à dialoguer avec la "communauté des spécialistes de l'analyse des images numériques couleur". La seconde version du manuscrit est finalement acceptée par les deux rapporteurs et la thèse est soutenue en janvier 1996.

Cette première expérience, peu glorieuse, m'a fait prendre conscience des tous les risques que comporte le co-encadrement de jeunes Doctorants. Avec les suivants, toutes les précautions ont été prises et les thèses ont été soutenues sans accroc.

"Qu'attends-tu pour passer ton HDR ?"

C'est ainsi, qu'après avoir co-encadré quelques Doctorants, j'ai eu droit régulièrement à cette question qui avait le don de m'agacer quelque peu. Surtout que le terme attendre, traduisant un état passif, me semble incompatible avec le sens que je donne à l'HDR. On attend le train, sa femme, certains l'heure de la retraite, mais on n'attend pas pour passer une HDR. Comme si l'âge ou le nombre d'années passées en tant qu'enseignant-chercheur suffit pour être digne d'encadrer des chercheurs...

Je ne pense pas avoir attendu, j'ai préféré prendre le temps de me former au contact de Jack-Gérard Postaire à l'exigeante fonction d'encadrement de chercheurs, généralement assumée par les Professeurs d'Université. Ma perception du rôle des Professeurs d'Université au sein d'un laboratoire explique ma démarche.

"Être ou ne pas être (Professeur d'Université)"

Au sein du laboratoire universitaire dans lequel je travaille, la recherche est organisée selon une structure pyramidale à trois niveaux.

Les Doctorants sont les chevilles ouvrières de la recherche. Ils lisent les publications, développent les programmes, testent leurs idées ainsi que celles de leur(s) directeur(s), rédigent les différents rapports de recherche. Leurs fonctions sont principalement scientifiques, rarement administratives. Cette période est un aboutissement de leurs études. Elle est l'occasion d'approfondir leurs connaissances dans des domaines pointus, mais surtout elle doit être la période où ils échangent, confrontent leurs idées avec les autres Doctorants et enseignants-chercheurs du laboratoire. C'est pour cette raison que les Doctorants de notre équipe travaillent dans la même salle, afin de favoriser au mieux ces échanges.

Les Maîtres de Conférences forment la population la plus importante des permanents du laboratoire. Cette population se divise en deux catégories. Les Maîtres de Conférences de la première catégorie consacrent la majorité de leur temps aux lourdes activités administratives et assument des responsabilités d'enseignement, ce qui permet à ceux de la seconde catégorie d'entreprendre une activité soutenue de recherche. Cette activité de recherche s'exerce, soit sous la forme de travaux individuels, soit sous la forme de co-encadrements de Doctorants.

Étant dans une situation hiérarchique intermédiaire et ayant peu de pouvoir de décision, les relations

du Maître de Conférences avec les Doctorants et Professeurs sont généralement assez sereines. Ils sont respectés par les Doctorants en qualité de permanents. Ils collaborent avec les Professeurs sur des projets de recherche ou d'enseignement. N'étant pas au sommet de la pyramide, ils sont rarement jugés par les autres membres du laboratoire. Le statut du Maître de Conférences lui permet de mener ses activités comme il le désire, sans répercussions profondes sur ses revenus. Cette liberté d'actions est très rare dans notre société. Chaque jour, je tente de l'apprécier à sa juste valeur et d'en être digne.

Les Professeurs constituent le corps décisionnel du laboratoire. Ils sont recrutés et jugés principalement sur leurs activités de recherche. Ainsi, certains sont tentés de réduire au minimum leur activité pédagogique.

Les relations des Professeurs avec les autres membres du laboratoire ne sont pas toujours très harmonieuses. Ils sont en concurrence avec les autres Professeurs du laboratoire pour partager les moyens matériels mais surtout humains (bourses, postes...). De ce fait, certains n'hésitent pas à utiliser toutes les armes psychologiques qui sont à leur disposition (intimidation, séduction, dénigrement, manipulation, clientélisme) pour s'accaparer un maximum de ces moyens.

Il est vrai que la célèbre "commission de spécialistes", fondement de notre système de recrutement, constitue un champ privilégié d'expérimentations de toutes ces armes. Il est rare que tous les membres de cette commission, chargée de sélectionner un candidat pour pourvoir à un poste de Maître de Conférences ou de Professeur, n'aient aucun intérêt avec la prise de décision du recrutement. Ces membres sont donc juges (ils évaluent l'adéquation de chaque candidat au poste à pourvoir) et parties (plusieurs candidats peuvent être membres du laboratoire où travaillera le recruté). Le fonctionnement de cette commission devient tragique (unité de lieu, unité d'action, intérêts divergents) quand plusieurs membres du laboratoire postulent à un poste dont le profil est suffisamment évasif pour laisser le champ libre à tous les excès. Qui sait si TF1 ou M6 ne seraient pas friands de tels scénarios pour leurs émissions de télé-réalité basées sur l'élimination des candidats ? Toujours est-il que la "commission de spécialistes" est un formidable générateur de rancoeurs et de conflits en tout genre.

Le remplacement de cette commission par un dispositif de recrutement indépendant des intérêts personnels des permanents du laboratoire serait une solution satisfaisante pour éviter des conflits de personnes qui entachent la vie du laboratoire et freinent les coopérations entre membres de différentes équipes.

La concurrence entre Professeurs du laboratoire était auparavant aiguës par la relation bijective entre une équipe et un Professeur. Heureusement, cette structure a évalué incitant les Professeurs de notre laboratoire à coopérer. Actuellement, le LAGIS est composé de grosses équipes incluant plusieurs Professeurs chargés de coopérer. Je suis persuadé que la structure en réseau de l'internet a favorisé cette notion de coopération.

Le rôle du Professeur est ingrat. En effet, mener des travaux de recherche en collaboration avec un Professeur ne peut être que le fruit d'un désir et non d'une contrainte. A la différence de nombreux postes de direction, le Professeur doit convaincre les Doctorants et Maîtres de Conférences de travailler avec lui. Cette démarche nécessite alors d'établir une relation de confiance avec les Doctorants et Maître de Conférences. Confiance sur les qualités morales du Professeur (Qui a envie de travailler avec quelqu'un qui ne reconnaît pas son travail ? Qui désire collaborer avec un mandarin ?). Confiance également sur les compétences scientifiques et techniques du Professeur (un encadrant qui ne reconnaît pas son incompétence dans un domaine le conduit à ne pas justifier et expliquer ses décisions, mais à les imposer).

Un Professeur est donc jugé par ses pairs pour partager les moyens et par les Maîtres de Conférences et Doctorants pour collaborer sur ses travaux de recherche. Comme il est exposé aux jugements des

autres membres du laboratoire dans lequel il travaille, la médiocrité de son activité le place dans une situation difficile à assumer.

"C'est décidé, je passe mon HDR"

Je ne me suis senti mûr pour assumer la co-direction de recherche qu'après plusieurs années. Je tiens au terme co-encadrement d'un Doctorant. J'ai encore et toujours besoin de confronter mes idées, propos et surtout mes écrits avec un autre directeur de recherche. Jack-Gérard Postaire a assumé (avec brio, mais je m'arrête là car ce n'est pas la page des remerciements) ce rôle depuis le début de ma carrière.

Partant du principe qu'il y a toujours plus d'idées dans deux têtes que dans une, je n'envisage pas d'assumer seul des encadrements de Doctorants. Les Doctorants ont tout à gagner à être co-encadrés par deux directeurs de recherche qui coopèrent.

L'habilitation à co-encadrer des recherches étant absente de la liste de nos diplômes, j'ai choisi de rédiger ce mémoire en vue d'obtenir l'Habilitation à Diriger des Recherches.

Vous n'y trouverez pas une synthèse de l'ensemble des travaux effectués depuis mon arrivée au laboratoire. Plusieurs d'entre eux ont été arrêtés, faute de résultats probants ou de perspectives attrayantes. Voulant conjuguer rédaction avec plaisir, j'ai préféré dégager trois thèmes de recherche sur lesquels je suis toujours en train de travailler et, qui sait ?, d'avancer...

Je vous souhaite autant de plaisir à lire ce manuscrit que j'en ai eu à le rédiger.

Ludovic Macaire, notice individuelle

Né le 6 octobre 1965 à Haubourdin (59)

Marié, 3 enfants

Maître de Conférences à l'UFR d'IEEA (Informatique, Électronique, Électro-technique, Automatique) de l'Université des Sciences et Technologies de Lille (USTL)

Adresse professionnelle

Laboratoire LAGIS UMR - CNRS - 8146

Université des Sciences et Technologies de Lille

Cité Scientifique - Bâtiment P2

59655 Villeneuve d'Ascq CEDEX

Tél : 03 20 43 41 69 (direct), 03 20 43 48 76 (secrétariat)

Fax : 03 20 43 65 67

ludovic.macaire@univ-lille1.fr

Cursus universitaire

Juin 1988 Diplôme d'Ingénieur de l'Université de Technologies de Compiègne (UTC)

Juin 1989 DEA "Contrôle des Systèmes", UTC, Mention AB

Sept 1989 Inscription en thèse au Centre d'Automatique de Lille (CAL),

sous la direction du Professeur Jack-Gérard POSTAIRE

Oct 1992 Attaché Temporaire d'Enseignement et de Recherche, EUDIL

Janv 1993 Doctorat de "Productique : Automatique et Informatique Industrielle"

USTL, mention très honorable avec les félicitations du jury

Oct 1993 Maître de Conférences, USTL, UFR d'IEEA

Doctorat

de l'Université des Sciences et Technologies de Lille, spécialité "Productique : Automatique et Informatique Industrielle", soutenu le 20 janvier 1993, mention très honorable avec les félicitations du jury.

Inspection d'états de surface par vision linéaire. Application au contrôle qualité continu en sidérurgie.

Membres du jury :

- Pierre VIDAL, président
- Bernard DUBUISSON, rapporteur
- Claude LAURGEAU, rapporteur
- Luc DUVIEUBOURG, examinateur
- Jack-Gérard POSTAIRE, directeur de recherche

Thématiques de recherche

- **Analyse d'images couleur**
- **Sélection automatique d'espaces couleur adaptés à des problèmes de segmentation d'images**
- **Segmentation d'images par classification spatio-colorimétrique de pixels**
- **Reconnaissance d'objets sous éclairage non contrôlé**

Publications

- **4** chapitres de livres : 2 dans des livres internationaux et 2 dans un ouvrage national collectif
- **4** articles dans des revues internationales avec référé : 1 dans Computer Vision and Image Understanding, 3 dans Machine Graphics and Vision
- **1** article dans une revue nationale avec référé : Revue d'Automatique et de Productique Appliquée
- **44** communications dans des congrès : 32 conférences internationales avec comité de lecture, 12 conférences nationales ou internationales mineures
- **1** article et **3** communications pour la diffusion de la connaissance

Rayonnement scientifique

- Collaborations
 - **3** avec des laboratoires internationaux (2 Maroc et 1 Belgique)
 - **4** avec des laboratoires nationaux (INRETS, IRCOM-SIC, PSI, LASL)
- **1** communication invitée dans un congrès international avec actes
- Relecture de publications :
 - **3** revues internationales
 - **1** revue nationale
 - **2** congrès internationaux

Responsabilités collectives :

- Membre élu de la commission de spécialistes 61^{ème} section de l'USTL de 1998 à 2004
- Membre élu au conseil du laboratoire I3D de 1998 à 2003
- Membre élu au conseil du laboratoire LAGIS depuis 2004

Encadrement

- Contrat d'encadrement doctoral et de recherche depuis le 1^{er} octobre 1998, renouvelé en 2002
- **5** thèses (4 soutenues, 1 en cours)
- **8** stages de recherche de DEA

Enseignement

1992 - 2004, 12 ans d'enseignement à temps plein (192h TD/an), UFR d'IEEA, USTL

- Temps réel
- Supervision
- Programmation structurée
- Logique combinatoire et séquentielle
- Traitement d'images

Activités d'enseignement

J'enseigne à l'Université des Sciences et Technologies de Lille (USTL) depuis 1992 : Attaché Temporaire d'Enseignement et de Recherche pendant 1 an avant d'être nommé en 1993 Maître de Conférences à l'UFR d'IEEA (Informatique, Électronique, Électro-technique, Automatique) de l'USTL. J'ai acquis une expérience de 12 années d'enseignement à temps plein (192 heures TD par an). Ce chapitre détaille les heures académiques de cours, travaux dirigés et travaux pratiques, ainsi que l'activité d'encadrement de projets déconnectée de mon activité de recherche. Dans tout ce qui suit, les nombres d'heures indiqués sont en équivalent TD.

Bilan général

Principales matières enseignées

Les activités effectuées par matière sur l'ensemble des années sont détaillées plus loin.

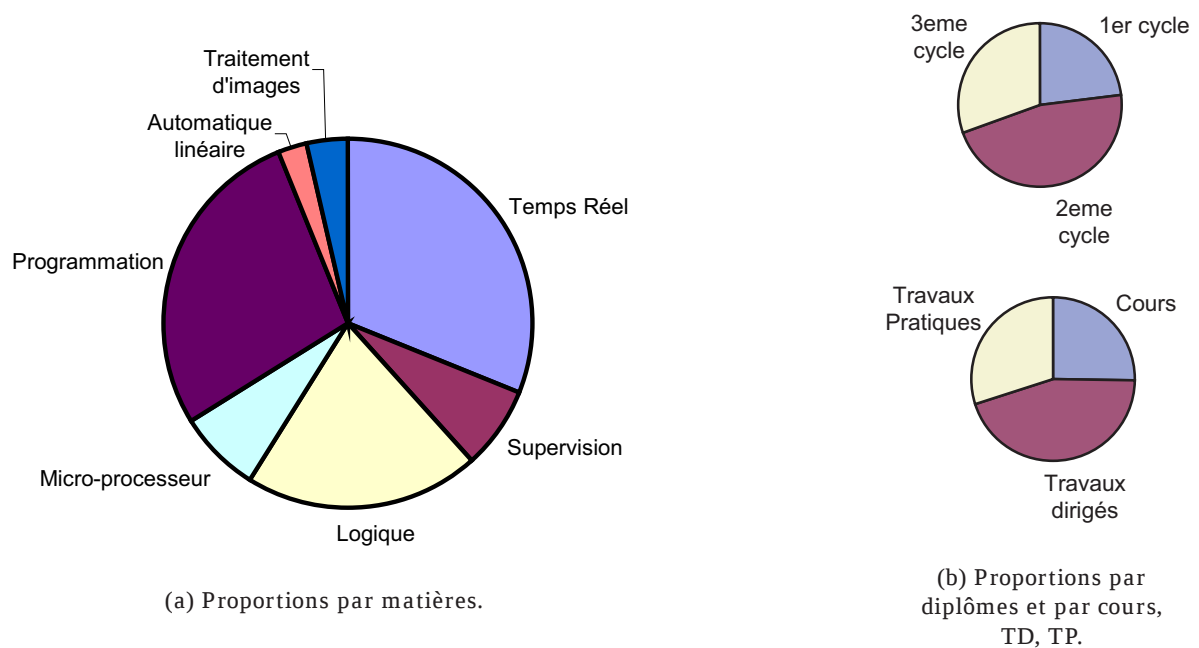


FIG. E.1 – Proportions par matières, par diplômes et par cours, TD, TP.

Diplômes

- 1^{er} cycle :
 - DEUG STPI (Sciences et Techniques pour l'Ingénieur)
- 2^{ème} cycle :
 - Licence EEA (Electronique, Electro-technique, Automatique)
 - Licence SPI (Sciences pour la Production Industrielle)
 - Maîtrise EEA
 - Ecole d'Ingénieur EUDIL option IMA (Informatique-Mesures-Automatique)
- 3^{ème} cycle :
 - DESS SP (Systèmes de Production)
 - DESS IIDC (Double Compétence en Informatique Industrielle)
 - DESS GI (Génie Industriel)

Activités d'intérêt collectif

- Responsable pédagogique
 - module Automatique et Logique - DEUG STPI 1^{ère} année- depuis 1993
 - module Temps Réel - Licence SPI - depuis 1995
- Responsable de l'organisation des TP
 - module Programmation Structurée - Licence EEA - 1993-1999
 - module Programmation Avancée - Maîtrise EEA - depuis 1993
- Rédaction de supports de cours pour
 - module Temps Réel - Licence SPI
 - module Temps Réel - DESS SP et DC
 - module Supervision - DESS GI et IIDC
- Création de TP
 - 4 séances de TP de logique combinatoire sur maquettes dédiées
 - 5 séances de TP de systèmes temps réel - Développement d'applications sous un noyau temps réel
 - 3 séances de TP de programmation en Langage C
 - 4 séances de TP de supervision avec le progiciel Intouch

Bilan par matière

Les volumes horaires exprimés en heures TD indiquent une moyenne sur l'ensemble des années.

Temps-réel

- Caractéristiques des systèmes temps-réel
- Spécification des systèmes temps-réel
- Spécificités des systèmes d'exploitation temps-réel
- Outils de coopération entre tâches

DESS SP - depuis 1993 - 21h C - 6 h TP -

DESS IIDC - depuis 1993 - 18 h C - 8 h TP -

Licence SPI - depuis 1995 - 21h TD -

Supervision

- Les fonctions d'un superviseur de processus industriels : synoptique, alarme, archivage temporel sélectif
- Structure des logiciels de supervision
- Méthodologie de développement d'une application de supervision
- Interfaces HMI pour la surveillance de processus industriels
- Communication entre logiciels de supervision et calculateurs, automates industriels

DESS IIDC - depuis 1999 - 6 h TP -

DESS GI - depuis 1999 - 12 h C - 11 h TD - 8 h TP -

Logique

- Introduction à la logique booléenne
- Circuits intégrés de logique combinatoire
- Circuits arithmétiques simples : addition et soustraction de nombres entiers positifs
- Eléments de logique séquentielle
- Circuits synchrones

DEUG STPI - depuis 1993 - 6h C - 28h TD -

Programmation

- Outils d'analyse d'un problème informatique
- Structures de contrôle pour la programmation structurée
- Structures de données évoluées pour la programmation avancée
- Application à la programmation en langage C

Maîtrise EEA - depuis 1993 - 20hTD - 20 h TP -

Licence EEA - depuis 1993 - 10hTD -

Traitement d'images

- Formation d'une image numérique
- Modélisation du bruit d'acquisition
- Filtrage d'une image
- Restauration d'une image

Maitrise EEA - de 1993 à 2003 - 12 h TD -

Automatique linéaire

- Initiation à l'automatique
- Analyse temporelle et fréquentielle de systèmes du premier ordre

DEUG STPI 1ere année - de 1993 à 1998 - 12 h TD -

Microprocesseur

- Organisation interne : UAL, séquenceur, registres
- Organisation externe : bus, chronogrammes
- Langage machine et assembleur
- Les interruptions

EUDIL option IMA - 1992 - 40hTD - 152 h TP -

Encadrements de Projets

Ayant encadré 2 à 4 projets d'étudiants par an, il est difficile d'établir une liste exhaustive de l'ensemble des sujets abordés. Voici un extrait des sujets des derniers projets que j'ai suivis :

- Création d'un système automatique de traduction de documents braille par analyse d'images - Les supports des documents examinés simple face et double face sont l'aluminium, la résine et le papier. Projet d'un binôme d'étudiants de DESS SP (3 mois à temps plein) - depuis 1998.
- Création d'un système de recherche d'images couleur par le contenu - Implantation sur l'internet sous forme de bases de données dynamiques - Projet d'un binôme d'étudiants de DESS SP (3 mois à temps plein) - depuis 2002.
- Étude de faisabilité d'un système de contrôle qualité de pièces d'alternateur. Implantation sur un système de vision industrielle Insight - Projet d'un binôme d'étudiants de DESS IIDC (1,5 mois à temps plein) - 2001.
- Mise en place d'un réseau Fipway d'automates TSX37 et connexion avec le logiciel de supervision Intouch - Projet d'un binôme d'étudiants de DESS SP (3 mois à temps plein) - 2000.
- Régulation d'un processus thermique avec le logiciel de supervision Intouch - Projet d'un binôme d'étudiants de DESS IIDC (1,5 mois à temps plein) - 1999.

Une expérience pédagogique : le DEUG STPI

A ma nomination en 1993 sur un poste de Maître de Conférences à l'UFR d'IEEA de l'USTL, j'ai été chargé de mettre en place le module d'automatique et de logique figurant au DEUG STPI nouvellement créé.

Ce DEUG s'adresse à des étudiants qui ne sont pas très à l'aise avec les matières fondamentales telles que les mathématiques ou les sciences physiques. Ces étudiants doivent atteindre en deux ans un niveau dans les matières fondamentales et technologiques suffisant pour suivre notamment la licence EEA dans de bonnes conditions.

L'enjeu pédagogique est excitant et nécessite la mise en place d'une structure pédagogique innovante, qui s'appuie sur un ordre chronologique original du déroulement du module. L'ordre classique est le suivant : les cours fondamentaux précèdent les travaux dirigés pendant lesquels les étudiants reprennent les notions vues en cours à partir de cas précis. Les travaux pratiques succèdent aux travaux dirigés permettant aux étudiants de tester sur des maquettes les solutions élaborées lors des travaux dirigés.

Dans le module d'automatique et logique, les étudiants découvrent les fondements de la matière lors des travaux pratiques. Les énoncés sont rédigés de telle sorte qu'ils contiennent tous les éléments dont les étudiants ont besoin. Lors des séances de travaux pratiques, l'enseignant doit faire preuve de beaucoup de disponibilité pour répondre aux nombreuses questions des étudiants. Les travaux

pratiques des étudiants sont évalués à partir des manipulations effectuées et du rapport décrivant ces manipulations.

Nous faisons face à deux types d'étudiants, ceux qui découvrent effectivement la matière et ceux qui ont déjà acquis des notions lors de leur scolarité. Pour répondre à cette situation, les exercices sont divisés en deux niveaux. Au premier niveau, les exercices de découverte de la matière doivent être effectués par tous les étudiants. Les exercices du second niveau s'adressent aux étudiants ayant déjà acquis des notions sur cette matière. La résolution de ces exercices leur permet d'améliorer leur note de travaux pratiques.

Suit alors le cours magistral, où les notions abordées en travaux pratiques sont décrites de manière fondamentale. Des exercices sont ensuite résolus par les étudiants lors des travaux dirigés.

Les énoncés des travaux dirigés sont également adaptés aux étudiants de premier cycle. Les problèmes sont décomposés en une série d'exercices de base. Un des étudiants est désigné pour résoudre l'exercice au tableau pendant que les autres le font sur leurs cahiers et vérifient la solution proposée au tableau. En suivant cette démarche, les étudiants abordent un nombre plus restreint d'exercices que lorsque l'enseignant fournit lui-même les réponses au tableau. Par contre, l'étudiant se sent obligé de suivre le déroulement des travaux dirigés, car il sait qu'il peut être interrogé à tout moment. Comme l'étudiant résout le problème au tableau, l'enseignant peut juger du niveau d'acquisition des notions. Enfin, cette démarche règle quelques détails de discipline en respectant les préceptes suivants : "Dernier arrivé à la séance de travaux dirigés, premier interrogé au tableau", "Étudiant sonné (par un portable), étudiant interrogé"...

Créé en 1993 avec 28 étudiants, le DEUG STPI est actuellement suivi par 225 étudiants. Cette démarche pédagogique est l'un des éléments d'explication de cette nette augmentation des effectifs.

Laboratoires de rattachement

Malgré les changements apparents de laboratoire, je n'ai jamais cessé de travailler avec la même équipe, et en particulier avec mon directeur de thèse Jack-Gérard Postaire.

En 1989, j'ai commencé mes travaux de thèse au sein du Centre d'Automatique de Lille.

En 1998, le laboratoire I^3D (Interaction, Image et Ingénierie de la Décision) est le résultat de la fusion du Centre d'Automatique de Lille et du Laboratoire de Mesures Automatiques.

En janvier 2002, le laboratoire I^3D a été associé au Département STIC du CNRS en tant que FRE 2497 et le laboratoire LAIL (Automatique et Informatique Industrielle de Lille) reconduit dans son association en tant qu'UMR 8021.

En janvier 2004, ces deux laboratoires ont fusionné pour donner naissance au Laboratoire LAGIS, (Laboratoire d'Automatique, Génie Informatique et Signal). Ce laboratoire commun à l'Ecole Centrale de Lille et à l'Université des Sciences et Technologies de Lille (USTL) est une Unité Mixte de Recherche (UMR 8146) du CNRS.

1989 - 1997 Centre d'Automatique de Lille

UPRES EA 1022 - Directeur : Pierre Vidal

Equipe Image et Décision - Responsable : Jack-Gérard Postaire

1998 - 2003 Laboratoire I^3D (Interaction, Image et Ingénierie de la Décision)

FRE CNRS 2497 - Directeur : Christian Vasseur

Groupe de Recherche "Scènes Dynamiques" - Responsable : Jack-Gérard Postaire

2004 Laboratoire LAGIS (Laboratoire d'Automatique, Génie Informatique et Signal)

UMR CNRS 8146 - Directeur : Philippe Vanheeghe

Equipe "Vision et Image" - Responsable : Jack-Gérard Postaire

Encadrements et jurys de thèse de Doctorat

Thèses co-encadrées

1. **Vincent Ultré**

Contribution à la segmentation d'images de mosaïque couleur

Thèse de Doctorat, Université des Sciences et Technologies de Lille

soutenue en janvier 1996

co-encadrement à 50% avec Jack-Gérard Postaire.

2. **Vénérée Rakotomalala**

Reconstruction bi-dimensionnelle de vaisseaux réiniens par analyse d'images couleur du fond de l'oeil

Thèse de Doctorat, Université des Sciences et Technologies de Lille

soutenue en décembre 1999

co-encadrement à 50% avec Jack-Gérard Postaire.

3. **Nicolas Vandenbroucke**

Segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimétriques adaptés. Application à l'analyse d'images de football

Thèse de Doctorat, Université des Sciences et Technologies de Lille

soutenue en décembre 2000

co-encadrement à 50% avec Jack-Gérard Postaire.

4. **Michael Fontaine**

Segmentation non supervisée d'images couleur par analyse de la connexité des pixels

Thèse de Doctorat, Université des Sciences et Technologies de Lille

soutenue en décembre 2001

co-encadrement à 50% avec Jack-Gérard Postaire.

5. **Damien Muselet**

Reconnaissance automatique de personnes sous éclairage non contrôlé par analyse d'images couleur

Thèse de Doctorat, Université des Sciences et Technologies de Lille

débutée en octobre 2001

co-encadrement à 50% avec Jack-Gérard Postaire.

Encadrement de stages de DEA

1. **Maurice Delplanque**

Application de la morphologie au contrôle qualité de galets de roulement

DEA de Productique : Automatique et Informatique Industrielle, Université des Sciences et Technologies de Lille

juillet 1993

co-encadrement avec Jack-Gérard Postaire.

2. **Jean Caron**

Algorithmes de segmentation d'images lignes adaptés à la détection de défauts d'aspect apparaissant sur des bandes métalliques laminées à froid

DEA de Productique : Automatique et Informatique Industrielle, Université des Sciences et Technologies de Lille

septembre 1993

co-encadrement avec Jack-Gérard Postaire.

3. **Nicolas Vandenbroucke**

Reconnaissance de personnes par analyse d'images couleur. Application au suivi de joueurs au cours d'un match de football

DEA de Productique : Automatique et Informatique Industrielle, Université des Sciences et Technologies de Lille

juillet 1996

co-encadrement avec Jack-Gérard Postaire.

4. **Michael Fontaine**

Classification non supervisée de pixels pour la segmentation d'images en niveaux de gris par analyse de la connexité. Application à la détection de lésions rétinienne dans des images du fond de l'oeil

DEA de Productique : Automatique et Informatique Industrielle, Université des Sciences et Technologies de Lille

juillet 1998

co-encadrement avec Jack-Gérard Postaire.

5. **Matthieu Skrzypniak**

Indexation d'images de personnes par analyse de matrices de co-occurrences couleur

DEA de Productique : Automatique et Informatique Industrielle, Université des Sciences et Technologies de Lille

juillet 2000

co-encadrement avec Jack-Gérard Postaire.

6. **Damien Muselet**

Invariants couleur pour l'indexation d'images de personnes

DEA de Productique : Automatique et Informatique Industrielle, Université des Sciences et Technologies de Lille

juillet 2001

co-encadrement avec Jack-Gérard Postaire.

7. Nicolas Nédonsel

Acquisition et classification d'images multi-spectrales

DEA de Productique : Automatique et Informatique Industrielle, Université des Sciences et Technologies de Lille

juillet 2002

co-encadrement avec Jack-Gérard Postaire.

8. Norredine Hemche

Segmentation d'images couleur par analyse floue d'un graphe d'adjacence de régions

DEA de Productique : Automatique et Informatique Industrielle, Université des Sciences et Technologies de Lille

septembre 2002

co-encadrement avec Jack-Gérard Postaire.

Participations à des jurys de thèses

1. Abdheramane Elkharraz

Classification automatique floue par réseaux de neurones. Application à l'analyse des images couleur

Thèse de Doctorat de Faculté des Sciences de Rabat

soutenue en mars 1999

(examineur).

2. Vincent Lebrun

Caractérisation des pierres ornementales par analyse d'images

Thèse de Doctorat de l'Université de Liège - Faculté des Sciences Appliquées

soutenue en mars 2002

(examineur).

3. Julien Dombre

Systèmes de représentation multi-échelles pour l'indexation et la restauration d'archives médiévales couleur

Thèse de Doctorat de l'Université de Poitiers - Faculté des Sciences Fondamentales et Appliquées

soutenue en décembre 2003

(examineur).

Rayonnement scientifique

Collaborations internationales

1. Maroc : Université de Fès, Ecole Supérieure de Technologie (EST) de Fès, Séjour à Fès en 1997
Constitution et analyse d'une base d'images pour la restauration de monuments historiques à Fès
Deux publications en commun.
2. Maroc : CNCPRST de la Faculté des Sciences de Rabat, séjour à Rabat en 1999.
Classification automatique floue par réseaux de neurones. Application à l'analyse des images couleur
Participation au jury de Thèse de Doctorat d'Abdheramane Elkharraz.
3. Belgique : Laboratoire MICA, Université de Liège, séjours de 2000 à 2002.
Contrôle-qualité de l'aspect de dalles de marbre par analyse d'images couleur
Participation au jury de Thèse de Doctorat de Vincent Lebrun
Une publication en commun.

Collaborations nationales

Ces collaborations se sont déroulées dans le cadre de projets institutionnels, mais également dans le cadre de collaborations ponctuelles.

Collaboration dans le cadre du GDR ISIS

De 1998 à 2001, j'ai appartenu au groupe qui animait **l'opération Couleur** (OT.3.2) dans le cadre du GT3 (systèmes de segmentation) du GDR ISIS (un livre en commun).

Collaborations sur projets institutionnels

1. 1996 - 1998 : **Analyse Couleur d'Images Numériques du Fond de l'Oeil**
Soutenu par le Ministère de l'Education Nationale, de la Recherche et de la Technologie dans le cadre des actions en 1996 sur la télé-médecine et les technologies pour la santé
En collaboration avec le Service des Maladies Infectieuses et du Voyageur de l'hôpital Dron de Tourcoing
Quatre publications en commun.

2. 2000-2001 : **Indexation d'images couleurs dermatologiques**

Soutenu par le GDR ISIS dans le cadre des projets Jeunes Chercheurs 2000-2001

En collaboration avec le Laboratoire PSI de l'Université de Rouen et le Laboratoire IRCOM-SIC de l'Université de Poitiers

Une publication en commun.

Collaborations hors projet institutionnel

Comme il s'agit de collaborations hors projet institutionnel, je cite dans cette partie les sujets et les chercheurs avec qui j'ai eu le plaisir de collaborer.

1. **Reconnaissance de personnes par analyse d'images couleur**

Avec Louahdi Khoudour du Laboratoire LEOST de l'INRETS (Institut National de la Recherche sur les Transports et la Sécurité) Villeneuve d'Ascq

Deux publications en commun.

2. **Identification de véhicules par analyse d'images couleur sous éclairage non contrôlé**

Avec Cina Motamed du Laboratoire LASL de l'Université du Littoral

Deux publications en commun.

Communication invitée dans un congrès international avec actes

P. Lambert et **L. Macaire**, "Filtering and segmentation : the specificity of color images", In the Proc. of the Int. Conf. on Color in Graphics and Image Processing, volume 1, Saint-Etienne, pages 57-64, octobre 2000.

Référé d'articles

Pour des revues internationales

1. IEEE Transactions on Fuzzy Systems
2. Eurasip Journal on Applied Signal Processing
3. Pattern Recognition Letters

Pour une revue nationale

1. Traitement du Signal

Pour des congrès internationaux

1. CGIP'2000 : International Conference on Color in Graphics and Image Processing
2. CGIV 2002 : First European Conference on Colour in Graphics, Imaging and Vision

Contrats Industriels

Ces deux contrats se sont déroulés dans le cadre de ma thèse au Centre d'Automatique de Lille de l'Université des Sciences et Technologies de Lille.

1. 1989-1993 :

Inspection automatique d'aspect de surface de produits métallurgiques plats

Fabrique de Fer de Maubeuge - Convention CIFRE.

2. 1995-1997 : **Développement d'un système d'inspection automatique temps réel d'aspect de surface de produits métallurgiques plats**

Direction de Recherche et Développement de Systèmes Industriels de Matra Cap Systemes et Fabrique de Fer de Maubeuge.

Publications

Chapitres de livres

1. A. Gillet, **L. Macaire**, C. Botte-Lecocq et J. G. Postaire. Chapitre "Color image segmentation by analysis of 3D histogram with fuzzy morphological filters", pages 154–177, Dans *Fuzzy filters for image processing- Studies in fuzziness and soft computing*, Authors of the book : M. Nachtegael, D. Van der Weken, D. Van De Ville and E.E. Kerre. Springer-Verlag Editor, 2002.
2. A. Gillet, C. Botte-Lecocq, **L. Macaire** et J. G. Postaire. Chapitre "Application of fuzzy mathematical morphology for unsupervised color pixels classification", pages 69–75, Dans *Data Analysis Classification and Related Methods*, Authors of the book : H. Kiers, J. Rasson, P. Groenen and M. Schader. Springer-Verlag Editor, 2000.

Articles dans des revues scientifiques avec comité de lecture

1. N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis.". *Computer Vision and Image Understanding*, 90(2) :190–216, 2003.
2. D. Muselet, **L. Macaire** et J. G. Postaire. "Color person image indexing and retrieval". *Machine Graphics and Vision*, 11(2/3) :257–283, 2002.
3. V. Rakotomalala, **L. Macaire**, J. G. Postaire et M. Valette. "Identification of retinal vessels by color image analysis". *Machine Graphics and Vision*, 7(4) :725–743, 1998.
4. V. Ultré, **L. Macaire** et J. G. Postaire. "Determining compatibility coefficients for color contour detection by relaxation". *Machine Graphics and Vision*, 5(1/2) :265–275, 1996.
5. V. Ultré, **L. Macaire** et H. Deleplanque. "Détection de contours d'images couleurs par étiquetage probabiliste". *Revue Automatique et Productique Appliquée*, 8(2) :315–320, 1995.

Communications avec actes dans des congrès internationaux

1. L. Busin, N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color space selection for unsupervised color image segmentation by analysis of connectedness properties". In *Proc. of the Int. Conf. on Complex Systems, Intelligence and Modern Technology Applications*, pages 311–316, Cherbourg, 2004.
2. L. Busin, N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color space selection for unsupervised color image segmentation by histogram multithresholding". In *Proc. of the IEEE Int. Conf. on Image Processing*, page à paraître, Singapour, 2004.
3. T. Leclercq, L. Khoudour, **L. Macaire** et J. G. Postaire. "Compact color video signature by principal component analysis". In *Proc. of the Int. Conf. on Computer Vision and Graphics*, page à paraître, Varsovie (Pologne), 2004.
4. D. Muselet, **L. Macaire** et J. G. Postaire. "Co-occurrences of adapted features for object recognition across illumination changes". In *Proc. of the Int. Conf. on Computer Vision and Graphics*, page à paraître, Varsovie (Pologne), 2004.
5. D. Muselet, C. Motamed, **L. Macaire** et J. G. Postaire. "Co-occurrence matrices of color feature vectors for multi-camera vehicle identification". In *Proc. of the Int. Conf. on Advanced Concepts for Intelligent Vision Systems*, pages 22–29, Ghent (Belgique), 2003.

6. D. Muselet, C. Motamed, **L. Macaire** et J. G. Postaire. "Color invariant features for multi-camera vehicle identification". In *Proc. of the IMACS Multiconference Computational Engineering in Systems Applications*, volume 4, pages 232–235, Villeneuve d'Ascq, 2003.
7. D. Muselet, **L. Macaire**, J. G. Postaire et L. Khoudour. "Color invariant for person images indexing". In *Proc. of the first European Conf. on Colour in Graphics, Image and Vision*, pages 236–240, Poitiers, 2002.
8. M. Sofka, R. Benslimane, **L. Macaire**, M. Rudko et J. G. Postaire. "Archeological mosaic image indexing by color-based segmentation and skeleton extraction". In *Proc. of the second IEEE Int. Symp. on Signal Processing and Information Technology*, pages 327–331, Marrakesh (Maroc), 2002.
9. A. Gillet, **L. Macaire**, C. Botte-Lecocq et J. G. Postaire. "Color image segmentation by fuzzy morphological transformation of the 3D color histogram". In *Proc. of the tenth IEEE Int. Conf. on Fuzzy Systems*, volume 2, pages 824–824, Melbourne (Australie), 2001.
10. V. Lebrun et **L. Macaire**. "Aspect inspection of marble tiles by colour line scan cameras". In *Proc. of the Int. Conf. on Quality Control by Artificial Vision*, volume 2, pages 403–408, Le Creusot, 2001.
11. M. Fontaine, **L. Macaire** et J. G. Postaire. "Color image segmentation based on a multiscale analysis of pixel connectivity properties". In *Proc. of the Int. Conf. on Color in Graphics and Image Processing*, volume 1, pages 78–83, Saint-Etienne, 2000.
12. M. Fontaine, **L. Macaire** et J. G. Postaire. "Image segmentation based on an original multiscale analysis of the pixel connectivity properties". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 1, pages 804–807, Vancouver (Canada), 2000.
13. M. Fontaine, **L. Macaire** et J. G. Postaire. "Pixel classification scheme for color image segmentation". In *Proc. of the Int. Symp. on Image/Video Communications over Fixed and Mobile Networks*, pages 1–8, Rabat (Maroc), 2000.
14. M. Fontaine, **L. Macaire** et J. G. Postaire. "Unsupervised segmentation based on connectivity analysis". In *Proc. of the IAPR Int. Conf. on Pattern Recognition*, volume 1, pages 660–663, Barcelone (Espagne), 2000.
15. A. Gillet, **L. Macaire**, C. Botte-Lecocq et J. G. Postaire. "Color image segmentation based on fuzzy mathematical morphology". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 3, pages 340–343, Vancouver, 2000.
16. A. Gillet, **L. Macaire**, C. Botte-Lecocq et J. G. Postaire. "Fuzzy unsupervised color image segmentation". In *Proc. of the Int. Conf. on Color in Graphics and Image Processing*, volume 1, pages 141–146, Saint-Etienne, 2000.
17. P. Lambert et **L. Macaire**. "Filtering and segmentation : the specificity of color images (keynote)". In *Proc. of the Int. Conf. on Color in Graphics and Image Processing*, volume 1, pages 57–64, Saint-Etienne, 2000.
18. N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color image segmentation by supervised pixels classification in a color texture features space". In *Proc. of the IAPR Int. Conf. on Pattern Recognition*, volume 3, pages 625–628, Barcelone (Espagne), 2000.
19. N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color systems coding for color image processing". In *Proc. of the Int. Conf. on Color in Graphics and Image Processing*, volume 1, pages 180–185, Saint-Etienne, 2000.
20. N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Unsupervised color texture features extraction and selection for soccer images segmentation". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 2, pages 800–803, Vancouver (Canada), 2000.
21. M. Fontaine, **L. Macaire** et J. G. Postaire. "Fundus images segmentation by unsupervised classification". In *Proc. of the 12th Conf. on Vision Interface*, pages 296–302, Trois-Rivières (Canada), 1999.
22. V. Rakotomalala, **L. Macaire**, J. G. Postaire et M. Valette. "Bidimensional retinal blood vessel reconstruction by a new color edge tracking procedure". In *Proc of the IEEE Symp. on Image Analysis and Interpretation*, pages 232–237, Tucson (USA), 1998.
23. N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color pixels classification in an hybrid color space". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 1, pages 176–180, Chicago (USA), 1998.

24. J. Caron, **L. Macaire** et J. G. Postaire. "Real-time aspect inspection by texture analysis". In *Proc. of the Int. Conf. on Quality Control by Artificial Vision*, pages 242–250, Le Creusot, 1997.
25. N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Soccer player recognition by pixels classification in an hybrid color space". In *Proc. of the SPIE Int. Conf. on Multispectral and Hyperspectral Imagery III*, volume 3071, pages 23–33, Orlando (USA), 1997.
26. N. Vandenbroucke, **L. Macaire**, C. Vieren et J. G. Postaire. "Contribution of a color classification to soccer players tracking with snakes". In *Proc. of the IEEE Int. Conf. on System, Man, and Cybernetics*, volume 4, pages 3660–3665, Orlando (USA), 1997.
27. V. Ultré et **L. Macaire**. "Multispectral edge detection by relaxation algorithm". In *Proc. of the SPIE Int. Symp. on Electronic Imaging : Science & Technology*, volume 2666, pages 252–258, San Jose (USA), 1996.
28. V. Ultré, **L. Macaire** et J. G. Postaire. "Determination of compatibility coefficients for color edge detection by relaxation". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 3, pages 1045–1048, Lausanne (Suisse), 1996.
29. **L. Macaire** et J. G. Postaire. "Automated visual inspection of galvanized and painted metallic strips". In *Proc. of the 7th IEEE European Comp. Conf. COMPEURO 93*, pages 8–15, Paris, 1993.
30. **L. Macaire** et J. G. Postaire. "Flaw detection on galvanized metallic strips in real-time by adaptive thresholding". In *Proc. of the SPIE Computer Vision for Industry Symp. EUROPTO*, pages 200–205, Munich, 1993.
31. **L. Macaire** et J. G. Postaire. "Real-time control of galvanized coating aspect by a texture inspection system". In *Proc. of the IEEE Int. Conf. on Systems, Man and Cybernetics*, pages 493–498, Le Touquet, 1993.
32. **L. Macaire** et J. G. Postaire. "Real-time detection and identification of flaws on galvanized and painted strip". In *Proc. of the 5th Int. Conf. on Condition Monitoring and Diagnostic Engineering Management COMADEM 93*, pages 152–157, Bristol (UK), 1993.

Communications avec actes dans des congrès nationaux

1. D. Muselet, **L. Macaire**, P. Bonnet et J. G. Postaire. "Reconnaissance d'objets sous éclairage non contrôlé par analyse d'histogrammes invariants et adaptés". Dans *Actes de CORESA'2004 Journées d'études et d'échanges COMpression et REprésentation des Signaux Audiovisuels*, pages 203–206, Villeneuve d'Ascq, 2004.
2. D. Muselet, **L. Macaire**, P. Bonnet et J. G. Postaire. "Reconnaissance d'objets grâce à l'analyse des composantes". Dans *Actes du GRETSI'03 Colloque sur le Traitement du Signal et des Images.*, volume 3, pages 30–33, Paris, 2003.
3. A. Gillet, **L. Macaire**, C. Botte-Lecocq et J. G. Postaire. "Segmentation d'images couleur par application d'une transformation morphologique floue à l'histogramme couleur 3D". Dans *Actes de LFA'2001 Rencontres francophones sur la Logique Floue et ses Applications*, pages 153–160, La Rochelle, 2001.
4. M. C. Larabi, N. Richard, C. Fernandez et **L. Macaire**. "Aide au diagnostic pour les cancers de peau basée sur une indexation par la couleur, la texture et la peau". Dans *Actes de ICISP'2001 Int. Conf. on Image and Signal Processing.*, pages 1055–1062, Agadir (Maroc), 2001.
5. A. Gillet, **L. Macaire**, C. Botte-Lecocq et J. G. Postaire. "Segmentation d'images couleur par classification non supervisée de pixels grâce aux opérateurs morphologiques flous". Dans *Actes de CORESA'2000 Journées d'études et d'échanges COMpression et REprésentation des Signaux Audiovisuels*, pages 131–138, Poitiers, 2000.
6. M. Skrzypniak, **L. Macaire** et J. G. Postaire. "Indexation d'images de personnes par analyse de matrices de co-occurrences couleur". Dans *Actes de CORESA'2000 Journées d'études et d'échanges COMpression et REprésentation des Signaux Audiovisuels*, pages 411–418, Poitiers, 2000.
7. M. Fontaine, **L. Macaire** et J. G. Postaire. "Segmentation non supervisée d'images par analyse de la connexité, Application à la détection de lésions virales sur des images de fond d'oeil". Dans *Actes du GRETSI'99 Colloque sur le Traitement du Signal et des Images*, pages 35–38, Vannes, 1999.

8. V. Rakotomalala, **L. Macaire**, J. G. Postaire et M. Valette. "Analyse d'images pour la reconstruction 2D du réseau vasculaire rétinien". Dans *Actes du Colloque Int. sur le Traitement d'Images et les Systèmes de Vision Artificielle*, volume 1, pages 121–134, Oujda (Maroc), 1998.
9. V. Rakotomalala, **L. Macaire**, J. G. Postaire et M. Valette. "Reconstruction 2D du réseau vasculaire rétinien par un suivi récursif en couleur". Dans *Actes du 9ème Forum des Jeunes Chercheurs en Génie Biologique et Médical*, volume 1, pages 72–73, Brest, 1998.
10. N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Un algorithme de classification couleur. Application à la reconnaissance de joueurs de football". Dans *Actes du Colloque Int. sur le Traitement d'Images et les Systèmes de Vision Artificielle*, volume 1, pages 121–133, Oujda (Maroc), 1998.
11. N. Vandenbroucke, **L. Macaire** et C. Vieren. "Suivi automatique de personnes en mouvement par analyse d'images couleurs successives. Application au suivi de joueurs de football". Dans *Actes du GRETSI'97 Colloque sur le Traitement du Signal et des Images*, volume 2, pages 917–920, Grenoble, 1997.
12. Y. Ruichek, J. G. Postaire, **L. Macaire** et J. C. Burie. "Implantation neuronale pour la mise en correspondance de primitives en stéréovision linéaire". Dans *Actes de la Conf. Int. AMSE on Communications, Signal and Systems*, volume 1.2, pages 574–582, Rabat, 1995.

Communications sans actes et séminaires

1. **L. Macaire**. "Reconstruction 2D du réseau vasculaire rétinien par un suivi récursif en couleur". Dans *Réunion Texture et couleur organisée par l'Operation Imagerie Couleur du GDR ISIS*, juin 1999.
2. **L. Macaire**. "Exploitation de l'information couleur pour la segmentation d'images". Dans *Réunion Analyse d'images couleur organisée par l'Operation Imagerie Couleur du GDR ISIS*, juin 1998.

Diffusion de la connaissance

1. N. Vandenbroucke et **L. Macaire**. "Invariants couleur - méthodes et applications". *Techniques de l'Ingénieur*, 2004 à paraître.
2. **L. Macaire** et D. Muselet. "Invariants couleur - méthodes et applications". Dans *Actes de l'Ecole de printemps du GDR ISIS Operation Imagerie Couleur*, Dijon, avril 2003.
3. **L. Macaire**. "La segmentation d'images couleur par classification". Dans *Actes de l'Ecole de printemps du GDR ISIS Operation Imagerie Couleur*, Pau, avril 2001.
4. **L. Macaire**. "La segmentation d'images couleur". Dans *Actes de l'Ecole d'été du GDR ISIS Operation Imagerie Couleur*, Saint-Etienne, septembre 1999.

Chapitres dans un ouvrage collectif

1. S. Philipp, P. Lambert, O. Colot et **L. Macaire**. Chapitre "Segmentation", pages 200–255, Dans *Image couleur : de l'acquisition au traitement*, Auteurs du livre : A. Trémeau, C. Fernandez-Maloigne et P. Bonton. Editions Dunod, Paris, 2004.
2. N. Vandenbroucke, **L. Macaire**, J. G. Postaire et C. Vieren. Chapitre "Applications - Analyse de séquences d'images couleur de matchs de football pour le suivi des joueurs", pages 342–351, Dans *Image couleur : de l'acquisition au traitement*, Auteurs du livre : A. Trémeau, C. Fernandez-Maloigne et P. Bonton. Editions Dunod, Paris, 2004.

Article dans une revue hors du domaine avec comité de lecture

1. A. Zarghili, M. Fontaine, R. Benslimane, **L. Macaire** et J. G. Postaire. "Indexation et recherche des images par similarité de formes. application à la recherche documentaire pour les arts décoratifs arabo-andalous". *Archéologie Islamique*, 11 :1–12, 2001.

Chapitre 1

Introduction

1.1 Contexte des travaux

Nos travaux portent principalement sur **l'exploitation de la couleur pour la segmentation et l'analyse d'images numériques**. Ce domaine a été abondamment étudié depuis plusieurs décennies afin de résoudre des problèmes posés notamment par le contrôle qualité industriel, le diagnostic médical et la communication via des outils multi-médias.

L'analyse des images numériques couleur couvre un champ d'investigation plus vaste que celui couvert par l'analyse des images en niveaux de gris car il importe généralement de respecter les propriétés physiques et psycho-physiologiques de notre propre perception des couleurs.

La reconnaissance d'objets sous différents éclairages constitue un exemple significatif de cette problématique spécifique à l'analyse des images couleur. Le système de perception et d'interprétation humain des couleurs est capable d'identifier la couleur d'un objet, même lorsqu'il est éclairé par des dispositifs différents.

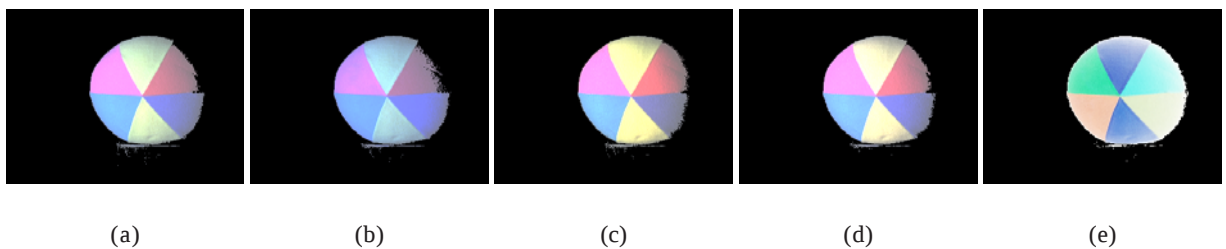


FIG. 1.1 – Les images (a),..., (d) représentent un objet identique éclairé par un dispositif propre à chaque image et observé selon les mêmes conditions de prise de vue. L'image (e) contient un objet différent de celui représenté par les images (a),..., (d). Ces images extraites de la base de l'Université Simon Fraser sont disponibles sur l'internet à l'adresse : http://www.cs.sfu.ca/colour/image_db.

Les images des figures 1.1(a),..., 1.1(d) représentent le même ballon, observé dans les mêmes conditions de prise de vue, mais éclairé par des dispositifs différents pour chacune des quatre acquisitions. L'image de la figure 1.1(e) représente un second ballon, différent de celui contenu dans les images des figures 1.1(a),..., 1.1(d). Bien que les couleurs du premier ballon varient d'une image à l'autre, nous sommes capables en examinant visuellement ces images de constater que les images des figures 1.1(a),..., 1.1(d) contiennent le même ballon, et que celui-ci est différent de celui représenté par l'image de la figure 1.1(e). Cet exemple montre que nous sommes capables de nous affranchir des

conditions d'éclairage afin d'identifier les objets en les caractérisant par leurs couleurs.

Par contre, les variations d'éclairage posent un problème pour la reconnaissance automatique d'objets par analyse d'images numériques couleur. Les méthodes d'analyse tentent alors de s'affranchir des conditions d'éclairage afin d'identifier les couleurs des objets représentés dans les images. Pour cela, elles se basent sur différents modèles qui caractérisent les variations des couleurs des pixels provoquées par des changements d'éclairage.

C'est pourquoi il nous semble essentiel de décrire brièvement le processus de formation d'une image numérique couleur avant de situer et présenter nos travaux sur l'exploitation de la couleur pour l'analyse et la segmentation d'images numériques.

1.1.1 Formation d'une image couleur

De la lumière à la couleur

La couleur d'un objet, telle que nous le percevons, est une notion très complexe qui obéit à des lois physiques, physiologiques et psychologiques particulières. Elle peut être considérée comme le résultat d'une combinaison des quatre éléments suivants (voir figure 1.2) :

- la source lumineuse,

Elle est caractérisée par sa répartition spectrale relative d'énergie, notée $S(\lambda)$. Cette fonction $S(\lambda)$ ne possède pas d'unité et correspond à la normalisation à 1, ou à 100, de la répartition spectrale d'énergie pour une longueur d'onde particulière, en général $\lambda = 560nm$. Certaines sources correspondant à des conditions courantes d'observation ont été normalisées par la CIE (Commission Internationale de l'Éclairage) sous le nom d'illuminants.

- le matériau,

Il réfléchit et absorbe une partie des rayons lumineux pour constituer le stimulus de couleur. Les caractéristiques de réflexion des matériaux sont estimées par le facteur de luminance lumineuse, appelé également réflectance spectrale et noté $\beta(\lambda)$. Il correspond au rapport entre l'énergie de la lumière réfléchie par le matériau et l'énergie de la lumière réfléchie par un diffuseur parfait observé dans les mêmes conditions d'éclairage et d'observation. Le stimulus de couleur $C(\lambda)$ s'exprime alors par la relation $C(\lambda) = \beta(\lambda) \times S(\lambda)$.

- le récepteur,

L'oeil focalise le stimulus de couleur issu de l'interaction lumière-matériau et le projette sur sa partie photo-sensible, la rétine.

- le système d'interprétation.

Il est constitué par le cerveau qui permet d'identifier une couleur grâce à différents attributs. Comme notre perception de la couleur dépend de l'interprétation établie par notre cerveau, cette perception est subjective.

Mesure de la couleur

Cependant, de nombreuses applications industrielles nécessitent de mesurer de manière objective la couleur d'un matériau. C'est dans cette problématique que s'inscrit la colorimétrie, science de la mesure de la couleur, qui obéit à des lois établies pour la plupart grâce des expériences d'appariement des couleurs.

Une expérience d'appariement consiste à demander à un observateur d'égaliser visuellement un stimulus de couleur projeté sur une moitié d'un écran blanc avec un mélange approprié de couleurs

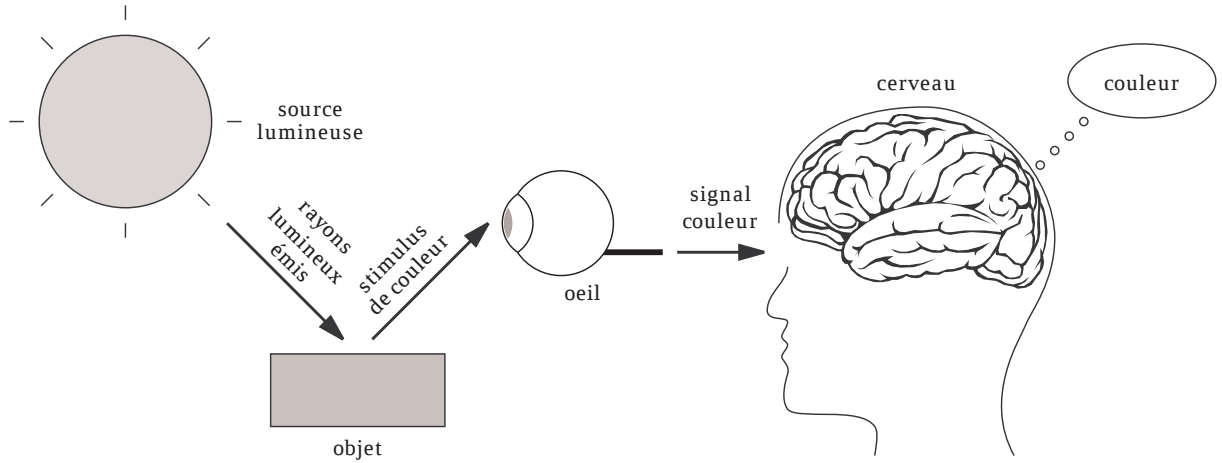
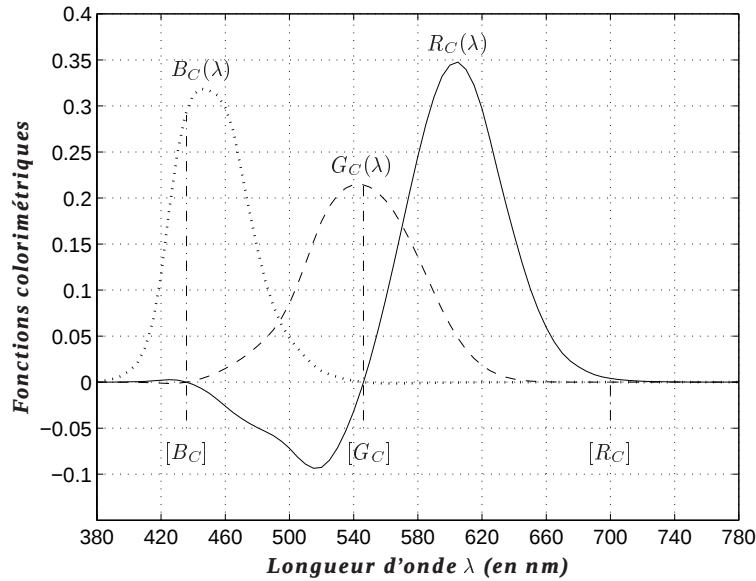


FIG. 1.2 – Perception humaine de la couleur.

primaires, superposées sur la seconde moitié de l'écran. Cette expérience permet de démontrer le principe de la trivariance visuelle : trois couleurs primaires sont nécessaires et suffisantes pour reproduire toutes les couleurs par synthèse additive, à savoir par le mélange de lumières colorées correspondant chacune à une des trois primaires. En 1931, la CIE a défini un ensemble de trois couleurs primaires, le rouge noté $[R_C]$, le vert noté $[G_C]$ et le bleu noté $[B_C]$, caractérisées par les fonctions colorimétriques que nous notons $R_C(\lambda)$, $G_C(\lambda)$ et $B_C(\lambda)$ (voir figure 1.3). Ces trois primaires ont été choisies de telle sorte que le mélange unitaire reproduise l'impression du spectre équi-énergétique (illuminant E).

FIG. 1.3 – Fonctions colorimétriques de la CIE 1931 $R_C(\lambda)$, $G_C(\lambda)$ et $B_C(\lambda)$.

Les composantes trichromatiques (R, G, B) d'un stimulus de couleur $C(\lambda)$ s'expriment par l'équation suivante :

$$\begin{cases} R = k \times \int_{\lambda=380}^{\lambda=780} R_C(\lambda) C(\lambda) d\lambda, \\ G = k \times \int_{\lambda=380}^{\lambda=780} G_C(\lambda) C(\lambda) d\lambda, \\ B = k \times \int_{\lambda=380}^{\lambda=780} B_C(\lambda) C(\lambda) d\lambda. \end{cases} \quad (1.1)$$

Le facteur k est choisi afin de normaliser les composantes trichromatiques par rapport à la fonction colorimétrique $G_C(\lambda)$ et par rapport à la source lumineuse qui est caractérisée par sa répartition spec-

trale relative d'énergie $S(\lambda)$. Lorsque $S(\lambda)$ est normalisée à 100, k s'exprime par l'équation suivante :

$$k = \frac{100}{\int_{\lambda=380}^{\lambda=780} S(\lambda) G_C(\lambda) d\lambda}. \quad (1.2)$$

Cette normalisation permet d'obtenir des composantes trichromatiques sans unité.

La figure 1.4 illustre le principe de mesure des composantes trichromatiques (R, G, B) d'un stimulus de couleur $C(\lambda)$.

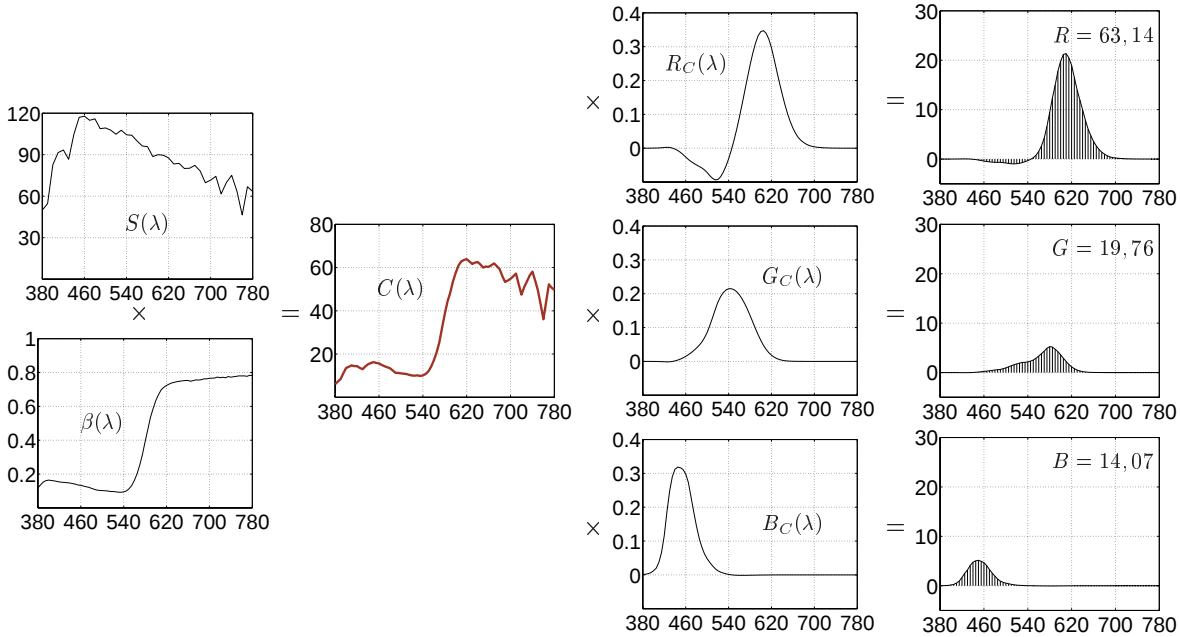


FIG. 1.4 – Exemple de mesure des composantes trichromatiques R , G et B d'un stimulus de couleur $C(\lambda)$, issu de la réflexion d'un matériau éclairé par une source lumineuse $S(\lambda)$ dont le facteur de luminance lumineuse est $\beta(\lambda)$. La fonction $S(\lambda)$ est celle de l'illuminant D_{65} ($k = 0,0517$). Toutes ces fonctions sont sans unité.

Afin d'exprimer une mesure d'un stimulus de couleur grâce aux couleurs primaires $[R_C]$, $[G_C]$ et $[B_C]$ définies par la CIE, il est nécessaire de définir un système de représentation de la couleur. A ces trois primaires, nous faisons correspondre respectivement trois vecteurs directeurs normés $\mathbf{R}_C$, $\mathbf{G}_C$ et $\mathbf{B}_C$ qui forment le repère d'un espace tri-dimensionnel d'origine O . Dans cet espace, chaque stimulus de couleur est représenté par un point-couleur $\mathbf{C}$ dont les coordonnées sont les composantes trichromatiques R , G et B . Certains de ces points ont des coordonnées négatives puisqu'ils correspondent à des stimuli de couleur qui ne peuvent pas être reproduits par synthèse additive. Les points correspondant à des stimuli de couleur dont les composantes trichromatiques sont positives sont contenus dans un cube, connu sous le nom de cube des couleurs (voir figure 1.5). L'origine O correspond au noir ($R = G = B = 0$) tandis que le blanc de référence est défini par le mélange unitaire des trois primaires ($R = G = B = 1$). La droite passant par les points *Noir* et *Blanc* est appelée axe des gris, axe des couleurs neutres ou encore axe achromatique. En effet, les points de cette droite représentent des nuances de gris allant du noir au blanc.

Les composantes trichromatiques d'un stimulus de couleur sont liées à sa luminance, qui est l'attribut d'une sensation visuelle selon laquelle une surface paraît émettre plus ou moins de lumière. Deux stimuli de couleur peuvent ainsi posséder le même caractère chromatique, que nous appelons chrominance, mais avoir des composantes trichromatiques R , G et B différentes à cause de leur luminance. Afin d'obtenir des composantes qui ne tiennent compte que de la chrominance, il convient donc de nor-

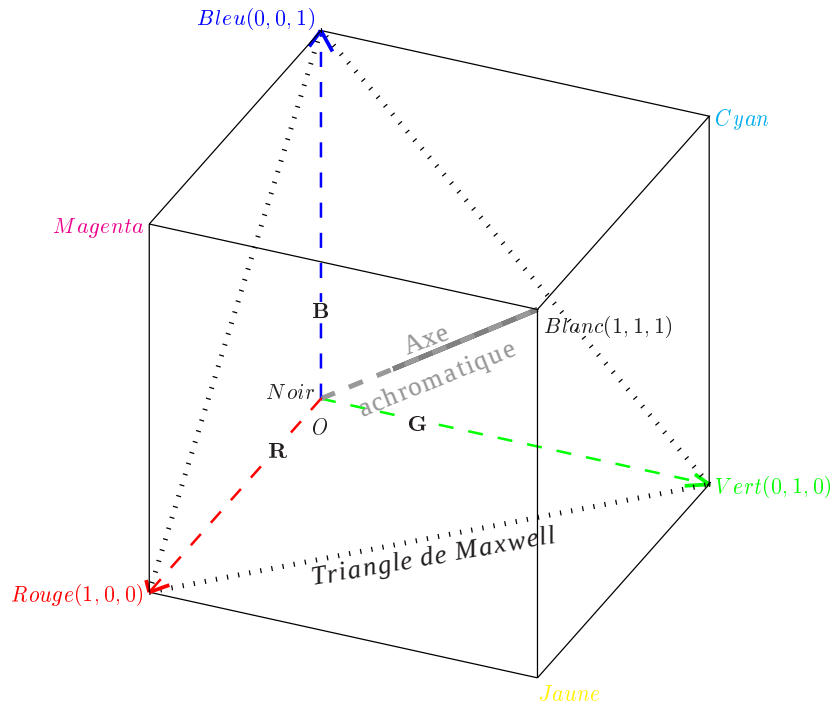


FIG. 1.5 – Cube des couleurs.

maliser les valeurs des composantes trichromatiques par rapport à la luminance. Ceci est réalisé en divisant chaque composante trichromatique par la somme des trois. Les composantes ainsi obtenues sont appelées coordonnées trichromatiques, coordonnées réduites ou encore composantes normalisées. Elles sont notées r , g , b et sont définies par :

$$\begin{cases} r = \frac{R}{R+G+B}, \\ g = \frac{G}{R+G+B}, \\ b = \frac{B}{R+G+B}. \end{cases} \quad (1.3)$$

Systèmes de représentation de la couleur

Le choix des primaires n'est pas unique. En théorie, il existe autant de systèmes de représentation de la couleur qu'il y a de systèmes de primaires. Un système de représentation se définit par le choix des primaires utilisées et du blanc de référence qui fixe leurs valeurs unitaires. Par ailleurs, d'autres systèmes de représentation de la couleur ont été conçus sans définir de nouvelles primaires car ils sont fondés sur le respect de propriétés physiques, physiologiques et psychologiques de notre perception des couleurs.

Avec Vandenbroucke [Van00], nous avons classé les systèmes de représentation existants en quatre familles (voir figure 1.6) :

- **Les systèmes de primaires,**
tels que le système (X, Y, Z) de la CIE et les systèmes (R_*, G_*, B_*) proposés par différents standards de définition de primaires (CIE, EBU European Broadcasting Union, FCC Federal Communications Commission. Le symbole $*$ représente l'un de ces standards.)
- **Les systèmes luminance-chrominance,**
dans lesquels une composante représente la luminance et deux composantes la chrominance d'un stimulus de couleur. Les composantes d'un système luminance-chrominance sont évaluées à partir

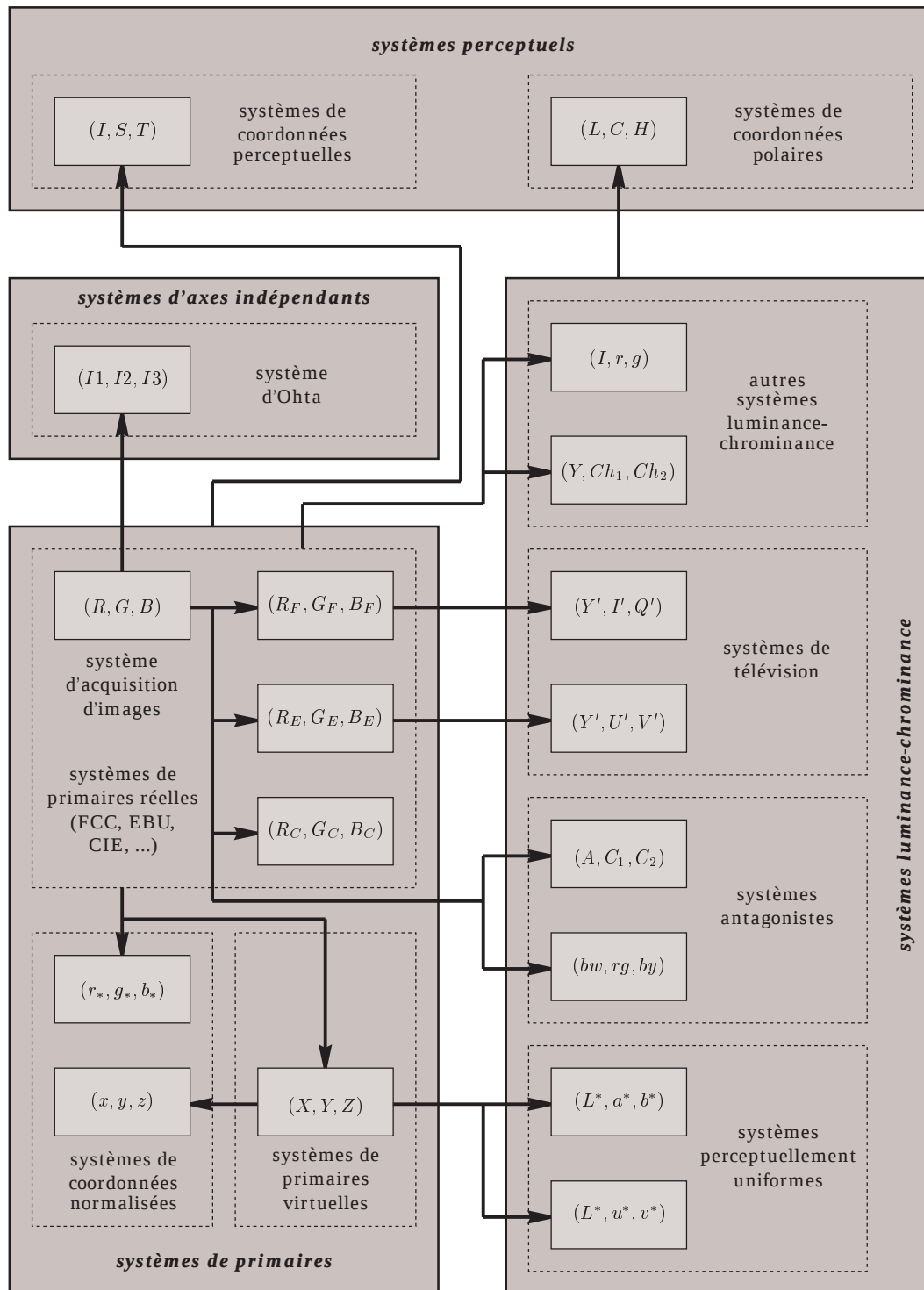


FIG. 1.6 – Les familles de systèmes de représentation de la couleur [Van00].

des composantes trichromatiques (R_*, G_*, B_*) , soit par une transformation linéaire définie en utilisant une matrice de passage, soit par une transformation non linéaire. Le type de transformation est lié à la nature même du système. Ainsi, nous distinguons différents types de systèmes luminance-chrominance :

- Les systèmes perceptuellement uniformes qui possèdent une métrique permettant d'établir une correspondance entre une différence entre deux couleurs telle qu'elle est perçue par l'homme et une distance dont la mesure reflète cette différence.
- Les systèmes de télévision qui permettent de séparer l'information de chrominance de l'information de luminance pour la transmission des signaux de télévision.
- Les systèmes antagonistes qui tentent de reproduire le modèle de la théorie des couleurs opposées de Hering.
- Les autres systèmes notés (I, r, g) ou (Y, Ch_1, Ch_2) , où Y désigne la luminance et Ch_1 et Ch_2 la chrominance d'un stimulus de couleur. Carron a proposé un système (Y, Ch_1, Ch_2) utilisé par de nombreuses méthodes d'analyse d'images numériques couleur [Car95].

– **Les systèmes perceptuels,**

qui représentent la couleur selon des entités telles que la luminance, la teinte et la saturation. La teinte correspond aux dénominations des couleurs telles que rouge, vert, bleu, jaune, ... Elle correspond à la longueur d'onde dominante d'un stimulus de couleur, c'est à dire la longueur d'onde pour laquelle l'énergie correspondante est la plus élevée. Le blanc, le noir ou les gris sont des couleurs qui n'ont pas de teinte. La saturation est une grandeur permettant d'estimer le niveau de coloration d'une teinte indépendamment de la luminosité. La saturation représente la pureté de la couleur perçue comme vive, pâle, terne, ...

Nous distinguons deux familles de systèmes perceptuels :

- Les systèmes de coordonnées polaires ou cylindriques. Ils correspondent à l'expression en coordonnées polaires des composantes de systèmes luminance-chrominance et se différencient donc par le système duquel ils se déduisent.
 - Les systèmes humains de perception de la couleur. Ils sont évalués directement à partir des composantes trichromatiques d'un système de primaires et se différencient par les relations exprimant la luminosité, la teinte ou la saturation.
- **Les systèmes d'axes indépendants,**

qui résultent de différentes méthodes statistiques fournissant les composantes les moins corrélées possible. Dans le domaine du traitement d'images couleur, une expérience menée par Ohta en 1980 sur une série de huit images différentes a permis de déterminer le système de représentation de la couleur noté $(I1, I2, I3)$ et basé sur la transformation de Karhunen-Loeve [OKS80].

Dans la figure 1.6 qui représente ces systèmes de représentation, les rectangles grisés correspondent à ces quatre familles et les flèches indiquent les transformations pour passer d'un système à un autre. Ces transformations sont détaillées dans [Sèv96, Van00]. Chacun des systèmes présentés est relié par ces flèches directement ou indirectement au système d'acquisition de l'image (R, G, B) . Nous distinguons, au sein de chaque famille, des sous-familles qui sont représentées par des rectangles aux contours discontinus.

A chaque système de représentation de la couleur correspond un espace tri-dimensionnel couleur, dans lequel un stimulus de couleur est représenté par un point-couleur. C'est pour cette raison, que dans la suite du mémoire, nous faisons référence aussi bien à des espaces couleur qu'à des systèmes de représentation de la couleur.

Ces différents systèmes sont utilisés pour représenter les couleurs des éléments de surface des objets projetés sur les pixels d'une image numérique couleur.

Image numérique couleur

Dans le domaine de la vision artificielle, les images couleur sont généralement acquises par une caméra vidéo couleur, puis numérisées par un ordinateur via une carte d'acquisition (voir figure 1.7).

L'acquisition des images couleur passe par les deux points essentiels suivants :

- un calibrage colorimétrique de la chaîne d'acquisition des images couleur est nécessaire pour que la couleur des objets observés par une caméra soit correctement représentée dans les images acquises.
- la connaissance des conditions d'acquisition des images couleur nécessite de spécifier le blanc de référence utilisé, l'éclairage employé et les fonctions colorimétriques des capteurs qui permettent de choisir les transformations adaptées vers d'autres systèmes de représentation de la couleur.

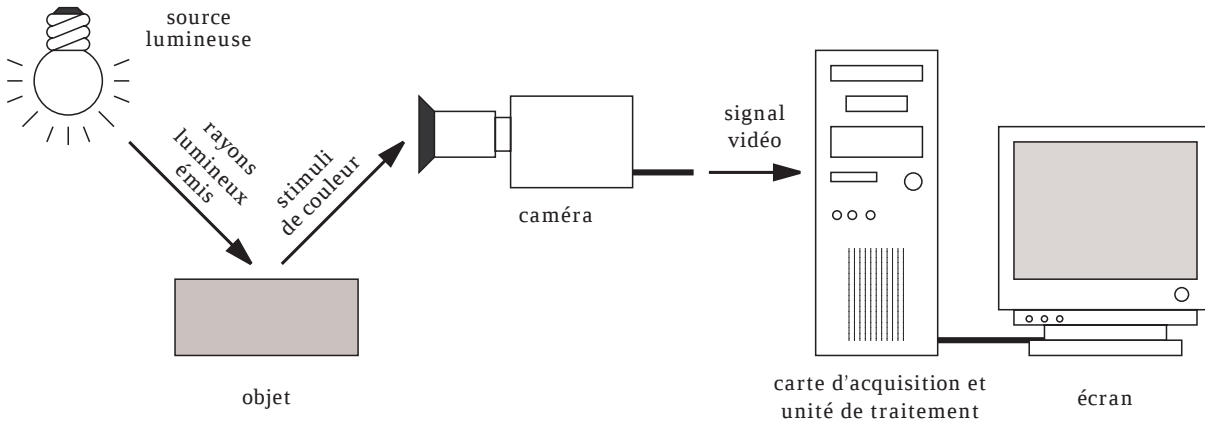


FIG. 1.7 – Vision artificielle.

La majorité des travaux d'analyse des images couleur ont été développés en s'appuyant sur deux hypothèses simplificatrices. Les images couleur sont spatialement bi-dimensionnelles, même si elles résultent de la projection de scènes observées tri-dimensionnelles. Elles sont multi-composantes, c'est à dire que chaque pixel est caractérisé par un *point-couleur* dans un *espace tri-dimensionnel des couleurs* ξ .

Ceci conduit à représenter une image numérique couleur I par trois signaux échantillonnés bi-dimensionnels à support et à valeurs bornés que nous notons $I(x, y)$ avec $[x, y] \in \mathbb{N}^2$ et $0 \leq x \leq X - 1$; $0 \leq y \leq Y - 1$; où x et y représentent les coordonnées du pixel P dans le plan image de taille $X \times Y$ et où $I(x, y)$ représente la couleur de ce pixel.

A chaque pixel P , nous associons un point-couleur, noté $I(P)$, repéré dans un espace tri-dimensionnel couleur ξ par trois coordonnées $I^i(P)$, $i = 1, 2, 3$, qui représentent les valeurs des trois composantes couleur quantifiées sur L niveaux.

L'image couleur I peut être également décomposée en trois *images de composante* I^i , $i = 1, 2, 3$, où le pixel P est caractérisé par le niveau de la i^{eme} composante couleur.

L'image numérique couleur est une représentation de la scène observée par une caméra couleur, où le point-couleur $I(P)$ associé à chaque pixel P correspond à une mesure de la couleur de l'élément de surface projeté sur ce pixel P . Les coordonnées du point-couleur dépendent également de la caméra et de l'illuminant. Ils peuvent s'exprimer par un modèle simple de formation de la couleur qui suppose que l'objet est composé d'éléments de surface plans, Lambertiens, perpendiculaires à l'axe optique de la caméra et sur lesquels les rayons lumineux issus du dispositif d'éclairage subissent une réflexion diffuse uniforme.

Lorsque la couleur des pixels est codée dans le système d'acquisition de l'image (R, G, B) , les niveaux des composantes rouge, verte et bleue sont alors notés $(I^R(P), I^G(P), I^B(P))^T$.

Les niveaux $I^i(P)$, $i = R, G, B$, du pixel P sont fonctions de :

- $E(\lambda)$ qui caractérise la distribution spectrale de l'illuminant,
 - $\beta(\lambda)$ qui définit la réflectance spectrale de l'élément de surface projeté sur le pixel P ,
 - $S^i(\lambda)$, $i = R, G, B$, qui caractérisent les sensibilités spectrales des trois capteurs de la caméra [ST97].
- Ils s'expriment par :

$$I^i(P) = \int_{\lambda=380}^{\lambda=780} S^i(\lambda) \beta(\lambda) E(\lambda) d\lambda, \quad i = R, G, B. \quad (1.4)$$

Ce modèle simple, basé sur une équation analogue à l'équation (1.1), montre que la formation de la couleur de chaque pixel obéit à des lois complexes.

Après avoir décrit brièvement la formation d'une image couleur, nous nous proposons d'examiner les procédés constituant la chaîne de traitements des images numériques couleur.

1.1.2 Chaîne de traitements des images couleur

Notre contribution s'inscrit dans le traitement automatique des images numériques couleur, dont l'objectif est de parvenir à interpréter le contenu d'une image analysée. La chaîne de traitements d'une image couleur se décompose en plusieurs étapes successives qui partent de l'analyse de la couleur de chaque pixel pour obtenir une interprétation du contenu de l'image (voir figure 1.8).

Préalablement à tout traitement, la couleur de chaque pixel représentée dans l'espace d'acquisition (R, G, B) peut être convertie dans un des systèmes de représentation de la couleur ξ décrits dans la figure 1.6. Cette phase effectue alors le codage du système de représentation de la couleur sélectionné.

La seconde étape consiste à filtrer l'image afin de mettre en évidence l'information "utile" qu'elle contient et/ou d'atténuer l'information "inutile". La frontière entre information "utile" et information "inutile" n'est pas simple à fixer car elle dépend de la nature de l'image et de l'application envisagée. Le filtrage des images couleur, qui a fait l'objet d'études approfondies dans [Lam02], n'est pas abordé dans ce mémoire.

La segmentation est un traitement bas-niveau qui, à partir de l'image couleur originale ou de l'image filtrée, s'attache à reconstruire les régions qui correspondent aux objets présents dans l'image. La segmentation transforme l'image couleur en une image où chaque pixel est caractérisé par une étiquette correspondant à la région à laquelle il appartient.

Une fois les régions reconstruites, l'étape d'analyse tente en général de les décrire à l'aide d'attributs qui permettent de reconnaître les objets contenus dans l'image.

Dans ce mémoire, nous nous arrêtons à cette étape d'analyse qui précède l'étape d'interprétation sémantique du contenu de l'image grâce à la manipulation d'informations symboliques.

1.1.3 Codage des systèmes de représentation de la couleur

La transformation du système (R, G, B) en un système de représentation de la couleur quelconque ξ , constitué de trois composantes que nous notons (T_1, T_2, T_3) , modifie souvent la dynamique des composantes couleur de chaque pixel. Ainsi, l'intervalle de valeurs entières non signées $[0, 255]$ auquel appartiennent les valeurs de chaque composante R , G et B se transforme en un intervalle de valeurs réelles signées ayant d'autres bornes. Ceci peut poser plusieurs types de problèmes en traitement d'images tel que par exemple, le calcul de distances euclidiennes entre points-couleur. La distance euclidienne permet d'évaluer les différences de couleur et a servi comme métrique pour la conception de plusieurs

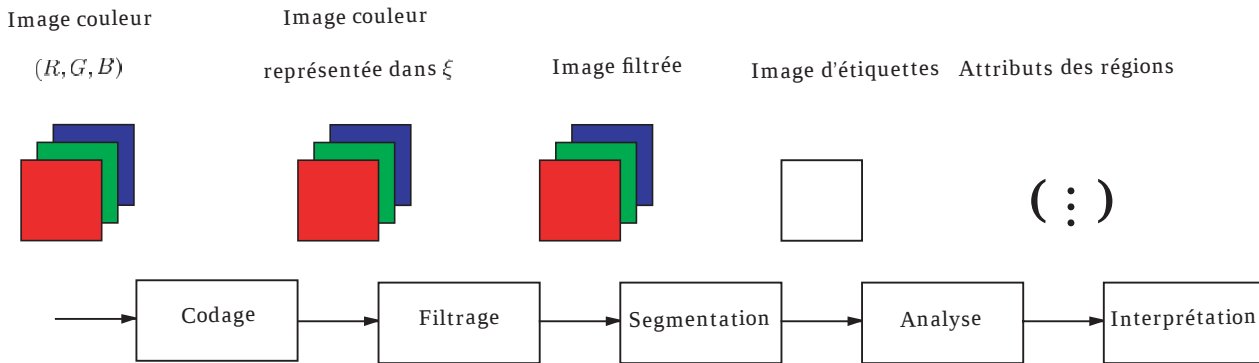


FIG. 1.8 – Chaîne de traitements d'une image couleur.

systèmes de la CIE. La comparaison de distances euclidiennes entre deux couleurs codées selon différents systèmes de représentation ne peut s'effectuer que si les axes de l'espace de représentation ont une dynamique comparable.

Ceci explique la nécessité de réaliser un codage des systèmes de représentation de la couleur de telle sorte que la dynamique des composantes codées, notées T'_1 , T'_2 et T'_3 , soit la même que celle des composantes R , G et B . Ce codage, dont le principe est illustré figure 1.9, consiste à décaler, normaliser et arrondir les valeurs des composantes transformées T_1 , T_2 et T_3 tout en tenant compte des propriétés intrinsèques du système correspondant.

Vandenbroucke propose une méthode de codage dont la principale caractéristique est qu'elle conserve la forme du gamut propre à chaque système de représentation de la couleur [VtP00]. Ce codage s'effectue de telle sorte que la distance euclidienne entre deux points-couleur soit cohérente avec les propriétés du système considéré.

Les couleurs des pixels étant codées dans un système de représentation, nous pouvons aborder la première étape du traitement qui consiste à segmenter l'image.

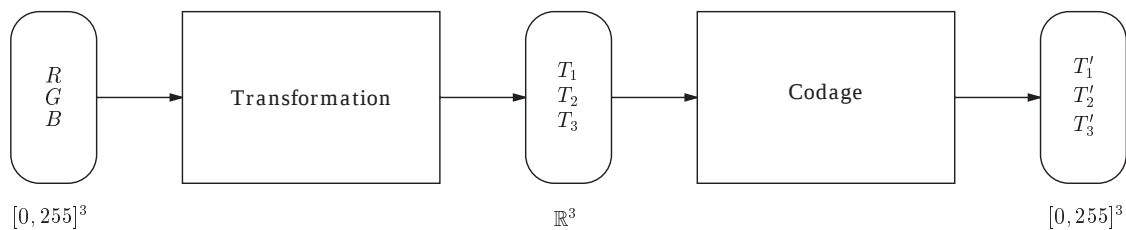


FIG. 1.9 – Codage des systèmes de représentation de la couleur.

1.1.4 Exploitation de la couleur pour la segmentation des images

Introduction

La segmentation d'une image couleur consiste à partitionner l'image en régions disjointes. En général, on suppose que les différentes couleurs présentes dans l'image correspondent principalement aux différentes orientations et propriétés de réflexion des surfaces des objets observés ainsi qu'à la présence d'ombres. Les méthodes de segmentation analysent ces couleurs afin de distinguer les différents objets qui composent une scène observée par une caméra couleur. Les différences entre les couleurs présentes dans l'image sont plus ou moins bien mises en évidence selon le système de représentation de la couleur utilisé.

Nous divisons les méthodes de segmentation d'images couleur en deux familles principales, selon l'hypothèse admise sur la formation des régions.

La première suppose que les régions adjacentes représentant des objets différents présentent des discontinuités locales de leurs couleurs à leurs frontières. Les méthodes liées à cette approche, dite approche "contours", tentent de rechercher les pixels contours couleur correspondant aux variations locales significatives des couleurs des pixels.

La seconde suppose qu'une région est constituée de pixels connexes partageant des propriétés colorimétriques similaires. Les méthodes relatives à cette approche, dite approche "régions", recherchent dans l'image des sous-ensembles de pixels connexes dont les couleurs sont homogènes.

Nous n'allons pas présenter ici une liste exhaustive des méthodes de segmentation d'images, mais plutôt tenter de mettre en évidence les principes adoptés par les méthodes de chacune de ces deux familles.

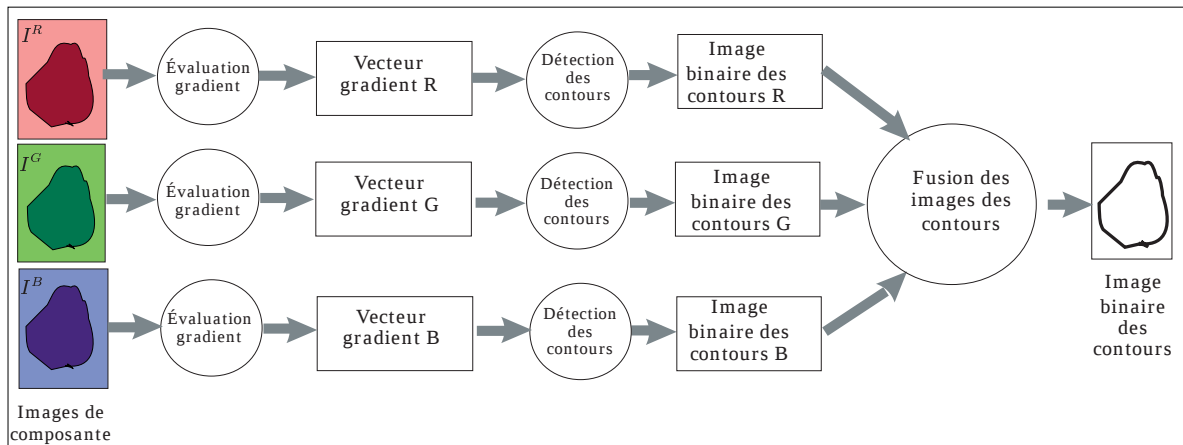
Approche "contours"

La détection des pixels contours couleur se définit comme la recherche de discontinuités locales de la couleur. Cette recherche s'effectue soit par une analyse d'une seule image de composante, soit par des analyses conjointes des images de composante, soit par une analyse "vectorielle" de l'image couleur. Nous nous limitons à la présentation des méthodes d'extraction des pixels contours couleur fondées sur la recherche des maxima d'une ou plusieurs normes de gradients. Le résultat de la détection des pixels contours couleur est une image binaire constituée des pixels contours et des pixels non contours.

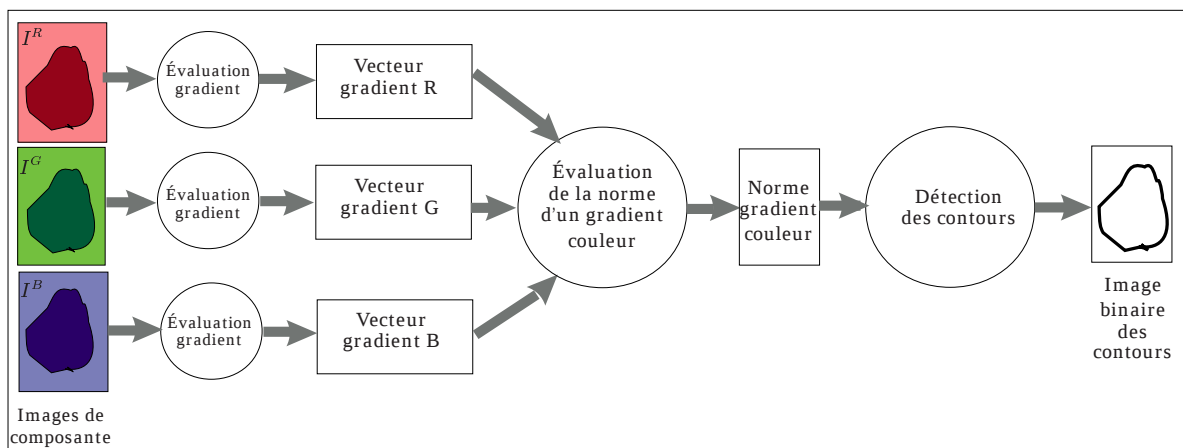
Les méthodes de détection des pixels contours couleur peuvent se diviser en quatre groupes :

- les méthodes qui analysent un vecteur gradient calculé à partir d'une seule image de composante. Par exemple, l'exploitation de l'image de luminance permet d'appliquer les méthodes de détection des pixels contours dans les images en niveaux de gris. Cette démarche s'avère satisfaisante quand les niveaux de luminance des pixels représentant les objets suffisent à les distinguer.
- les méthodes qui fusionnent les images binaires de contours obtenues par des analyses des images de composante [CP90] (voir figure 1.10(a)).
- les méthodes qui exploitent la norme d'un gradient couleur déterminée à partir des vecteurs gradients calculés dans les images de composante [LC99] (voir figure 1.10(b)).
- les méthodes qui expriment un vecteur gradient couleur à partir de vecteurs gradients calculés dans les images de composante. Les pixels contours couleur sont détectés par une analyse du vecteur gradient couleur [Zen86, LC91] (voir figure 1.10(c)).

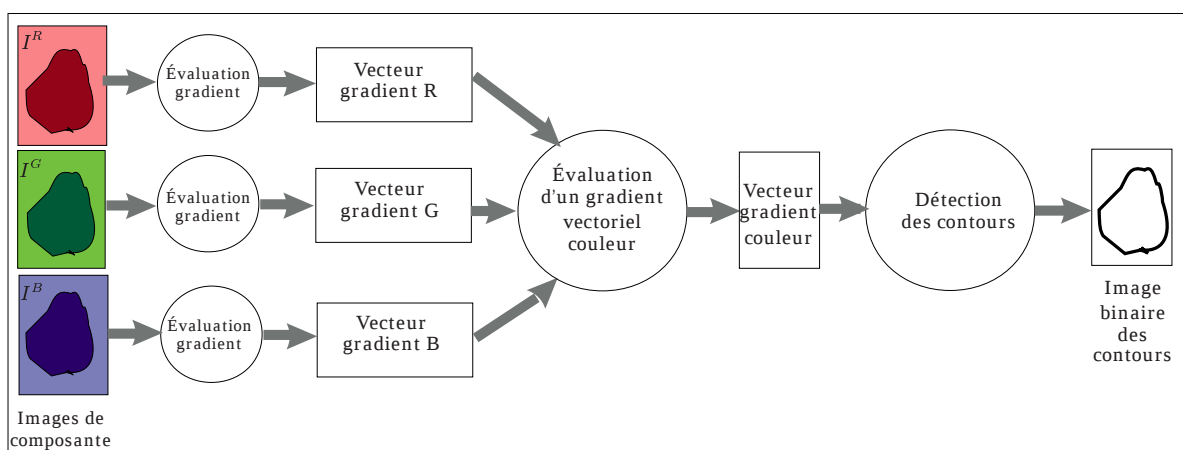
Pour obtenir une partition de l'image en régions, il est nécessaire de chaîner les pixels contours afin qu'ils constituent des lignes fermées correspondant aux frontières des régions [CP90].



(a) Analyse des images binaires de contours pour la détection des pixels contours couleur.



(b) Analyse de la norme d'un gradient couleur pour la détection des pixels contours couleur.



(c) Analyse d'un gradient vectoriel couleur pour la détection des pixels contours couleur.

FIG. 1.10 – Différentes stratégies pour la détection des pixels contours couleur.

Les méthodes de détection des pixels contours couleur sont adaptées à la segmentation d'images où les couleurs des régions sont contrastées. Elles ne fournissent pas toujours des résultats satisfaisants quand les contours des régions sont difficilement perceptibles dans l'image. Par ailleurs, reconstruire les lignes fermées délimitant les régions à partir des pixels contours couleur détectés reste un problème délicat. Par conséquent, de nombreux auteurs préfèrent l'approche "régions" pour segmenter des images numériques couleur.

Approche "régions"

La segmentation d'une image I en régions $\mathcal{R}_i$ a pour but de regrouper des pixels connexes ayant des couleurs similaires, afin de constituer des régions de couleurs homogènes, ces régions étant dans la mesure du possible liées à un objet ou à une partie d'un objet représenté dans l'image [Zuc76].

Les pixels de chaque région doivent respecter des critères d'homogénéité et de connexité. L'homogénéité d'une région $\mathcal{R}_i$ est définie par un prédicat d'uniformité, noté $Pred(\mathcal{R}_i)$. Ce prédicat est vrai si $\mathcal{R}_i$ est homogène, faux dans le cas contraire.

Les régions doivent respecter les quatre conditions suivantes :

- $I = \bigcup_i \mathcal{R}_i$,
- $\mathcal{R}_i$ est constituée de pixels connexes pour tout i ,
- $Pred(\mathcal{R}_i) = \text{vrai}$ pour tout i ,
- $Pred(\mathcal{R}_i \cup \mathcal{R}_j) = \text{faux}$ pour tout $i \neq j$, $\mathcal{R}_i$ et $\mathcal{R}_j$ étant adjacents dans I .

La première condition implique que chaque pixel de l'image doit appartenir à une région $\mathcal{R}_i$ et que l'union de toutes les régions correspond à l'image entière. La deuxième condition est relative à la structure des régions. Elle définit une région comme un sous-ensemble de pixels connexes. La troisième condition exprime que chaque région doit respecter un prédicat d'uniformité. La dernière condition implique la non-réalisation de ce même prédicat pour toute réunion de deux régions adjacentes.

Le résultat de la segmentation est une image dans laquelle est attribuée à chaque pixel une étiquette correspondant à la région à laquelle il appartient.

Les méthodes de reconstruction de régions peuvent se diviser en deux grandes familles, selon que l'analyse s'effectue dans le plan image ou dans un espace couleur [Lt00, CJSW01].

Analyse dans le plan image Les méthodes de segmentation par analyse spatiale du plan image regroupent des pixels connexes présentant des couleurs similaires pour reconstruire les régions de l'image. Elles représentent l'image sous la forme de structures de données de type tétra-arbre (quad-tree) ou de type graphe d'adjacence de régions.

Les méthodes de division-fusion des régions procèdent à l'analyse d'un tétra-arbre dont chaque noeud possède exactement quatre noeuds fils, exceptés les noeuds terminaux [Cou91]. La racine de l'arbre correspond à l'image entière dont les nombres de lignes et de colonnes doivent être des puissances de deux (voir figure 1.11). Chaque noeud représente un bloc, c'est à dire une zone de l'image de forme rectangulaire, dont les nombres de lignes et de colonnes sont également des puissances de deux.

Chaque bloc associé à un noeud du tétra-arbre de la partition initiale est analysé de façon itérative afin de décider s'il doit être divisé en quatre sous-blocs. L'analyse itérative s'arrête lorsque chaque sous-bloc respecte un critère d'homogénéité. A ce stade de l'analyse, certains blocs adjacents dans l'image présentent des caractéristiques colorimétriques identiques. C'est pourquoi ces couples de blocs sont fusionnés. L'analyse s'arrête lorsqu'il n'existe plus de couple qui respecte le critère de fusion. L'inconvénient majeur du tétra-arbre réside dans la rigidité du découpage qu'il impose. Il conduit à

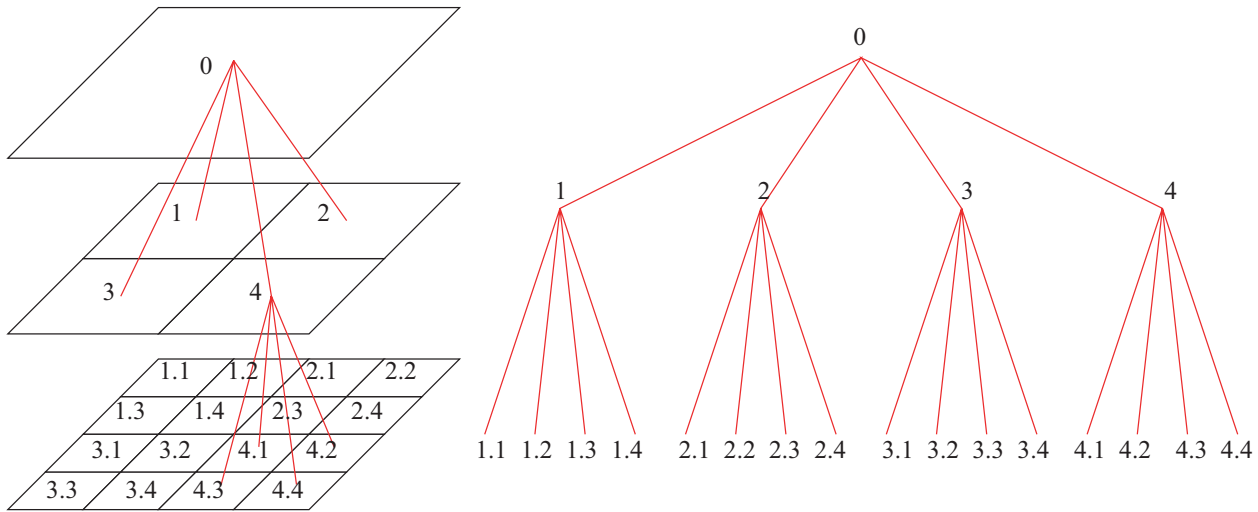


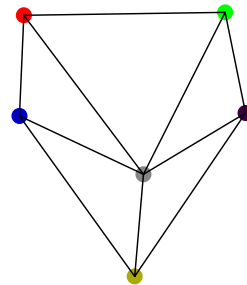
FIG. 1.11 – Structure pyramidale du tétra-arbre.

une partition globale de l'image qui ne respecte pas toujours la forme des régions présentes dans l'image.

Les méthodes de croissance de régions [TB97, SP97] et de fusion de régions [TC00] s'appuient sur l'analyse d'un graphe d'adjacence de régions. Il s'agit d'un graphe non-orienté dont chaque noeud représente une région et chaque arête représente une adjacence entre deux régions. Chaque région peut être constituée initialement d'un pixel dans le cadre de la croissance de régions, ou d'un sous-ensemble de pixels connexes quand l'image a été préalablement sur-segmentée. La figure 1.12(b) représente, à titre d'exemple, le graphe d'adjacence de régions obtenu à partir de l'image pré-segmentée de la figure 1.12(a). L'analyse itérative du graphe consiste à fusionner, à chaque itération, deux noeuds reliés par une arête à condition qu'ils respectent un critère de fusion.



(a) Image originale.



(b) Graphe d'adjacence de régions de l'image de la figure 1.12(a).

FIG. 1.12 – Graphe d'adjacence de régions.

Ces méthodes privilégient l'interaction spatiale entre les pixels, et ne font intervenir les propriétés colorimétriques des pixels qu'au stade de la décision de fusion ou division des régions.

Elles se distinguent par leur ordre de parcours des différents noeuds et par les critères de fusion ou de division utilisés. Certaines méthodes nécessitent le réglage de nombreux seuils et paramètres, tels que ceux requis par les critères de fusion ou division, et ceux employés par les critères d'arrêt de l'analyse

itérative. L'ajustement de ces paramètres est délicat, de telle sorte qu'il est difficile de prévoir si on aboutira à une sous-segmentation ou à une sur-segmentation de l'image [Lt00].

Ces limites ont conduit de nombreux auteurs à s'intéresser à l'analyse de la distribution des points-couleur représentant les pixels dans un espace couleur.

Analyse dans un espace couleur Les méthodes qui analysent la distribution des points-couleur dans un espace couleur considèrent que chaque pixel P est représenté par un point-couleur $I(P)$ dans l'espace couleur. Nous avons vu que les points-couleur peuvent être représentés dans des espaces couleur autres que l'espace (R, G, B) (voir figure 1.6). La performance d'un algorithme de segmentation dépend alors du choix de l'espace couleur [VtP03]. De nombreux auteurs ont tenté de déterminer les espaces couleur qui sont les mieux adaptés à leurs problèmes spécifiques de segmentation d'images couleur [LCK94]. Malheureusement, il n'existe pas un espace couleur qui permet d'obtenir des résultats satisfaisants pour la segmentation de tous les types d'images [VtP98]. **La première partie de nos travaux est consacrée à la construction d'espaces couleur adaptés à des problèmes spécifiques de segmentation d'images.**

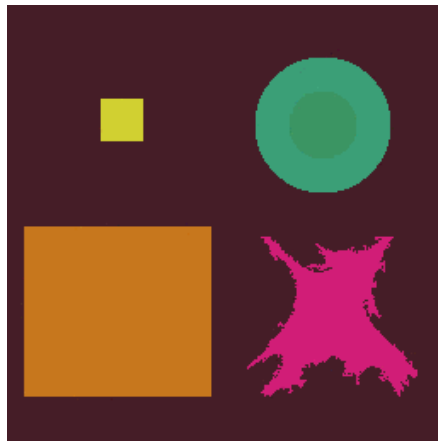
Les méthodes qui analysent la distribution des points-couleur représentant les pixels supposent que les régions homogènes de l'image donnent naissance à des nuages de points-couleur identifiables dans l'espace couleur considéré, chaque nuage définissant une classe de pixels qui partagent des propriétés colorimétriques similaires. Ces méthodes recherchent des groupements de points-couleur dans l'espace couleur considéré afin de construire des classes de pixels. Il s'agit donc de méthodes de classification automatique.

La figure 1.13 montre que les points-couleur associés aux régions dont les couleurs sont distinctes forment des nuages identifiables dans l'espace (R, G, B) . Elle montre également le cas, qui n'a rien d'exceptionnel, des deux disques verts concentriques où les points-couleur associés à ces deux régions adjacentes donnent naissance à deux nuages qui se chevauchent dans l'espace (R, G, B) , ce qui rend difficile leur discrimination.

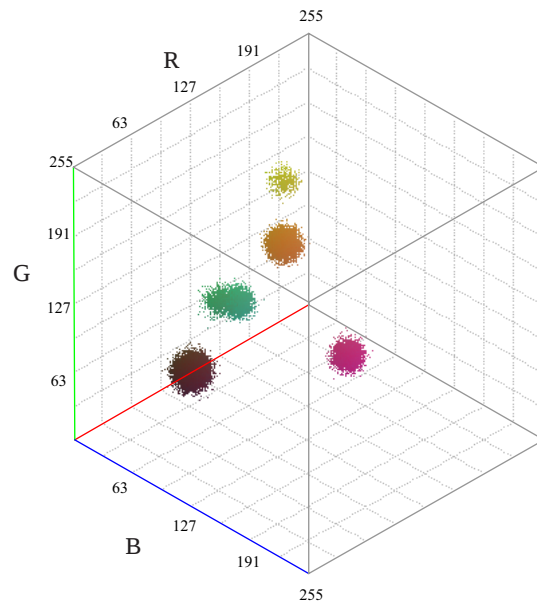
Les classes de pixels sont construites par des méthodes liées à une approche métrique [IK89, TNA95, UA94, Sch97a, PY99, CL02] ou par des méthodes relatives à une approche statistique [OKS80, Tom92, SPK98, PYL98, GtBLP02]. Quand les classes sont construites, chaque pixel est assigné à l'une d'entre elles par une règle de décision. Les étiquettes des régions sont affectées aux pixels connexes assignés aux mêmes classes afin de construire l'image segmentée.

Les méthodes liées à une approche métrique déterminent les centres de gravité des classes, notamment en minimisant des critères permettant d'évaluer la dispersion et la compacité des classes. Ces méthodes donnent généralement de bons résultats quand les nuages formés par les points-couleur sont globulaires et bien séparés dans l'espace couleur. Les pixels sont assignés aux classes en comparant des distances qui séparent leurs points-couleur des centres des classes dans l'espace couleur. La plupart de ces méthodes analysent globalement l'image et nécessitent soit de connaître au préalable le nombre de classes, soit si celui-ci est inconnu, de régler finement des paramètres afin de déterminer le nombre de classes en présence dans la population des pixels.

Les méthodes relatives à une approche statistique construisent les classes par une analyse de l'histogramme couleur $H[I]$ d'une image I . L'image est associée à une modélisation probabiliste dont l'histogramme couleur représente l'estimation de la loi de densité de probabilité d'apparition des couleurs des pixels. L'histogramme couleur est implanté sous la forme d'une structure tri-dimensionnelle de cellules repérées par les trois coordonnées des points-couleur considérés. Chaque cellule $H[I](c)$ associée au point-couleur c indique le nombre de pixels dans l'image caractérisés par ce point-couleur divisé



(a) Image couleur constituée de six régions.



(b) Nuages de points dans l'espace couleur (R, G, B) correspondant aux pixels de l'image de la figure 1.13(a).

FIG. 1.13 – Régions homogènes de l'image et nuages de points-couleur dans l'espace couleur.

par la taille de l'image, notée $X \times Y$.

Son analyse permet de mettre en évidence des domaines de l'espace dits "modaux", à savoir des zones de fortes concentrations de points-couleur. Ces domaines sont séparés par des "vallées", caractérisées par de faibles concentrations de points-couleur. Ces méthodes supposent que chaque domaine modal définit une classe de pixels. Ainsi, la construction des classes revient à détecter les domaines modaux de l'espace couleur. Ces méthodes sont adaptées aux cas où les classes de pixels sont relativement équiprobables. Par conséquent, comme les régions de petite taille sont peu représentées par les cellules de l'histogramme couleur, les modes auxquels elles donnent naissance risquent de ne pas être détectés par une analyse de l'histogramme couleur (voir figure 1.14).

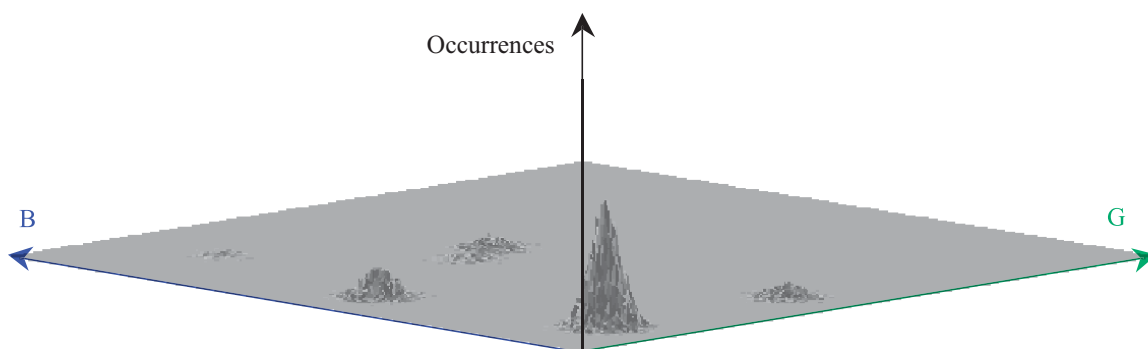
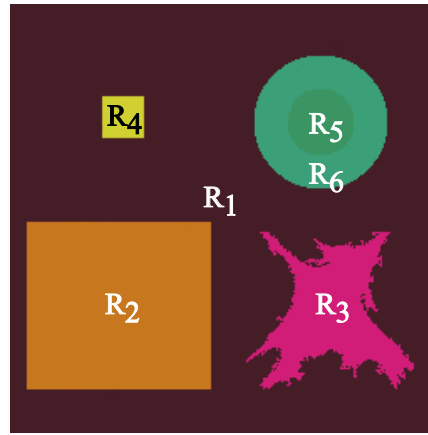


FIG. 1.14 – Histogramme couleur de l'image de la figure 1.13(a) qui ne tient compte que des niveaux des composantes G et B des pixels. Les petites régions de l'image sont sous-représentées par comparaison aux grandes régions de l'image.

Comme l'histogramme couleur est une structure très encombrante en termes de place mémoire, une solution consiste à projeter l'histogramme couleur sur les trois composantes afin de constituer les trois histogrammes mono-dimensionnels $H^i[\mathbf{I}]$ des images de composante I^i , $i = R, G, B$. Dans ce cas, les domaines modaux de l'espace sont supposés être des parallélépipèdes rectangles, dont les projections sur les composantes couleur constituent des intervalles modaux. Les intervalles modaux sur chaque composante sont détectés par une recherche des pics de l'histogramme mono-dimensionnel correspondant.

Un pic d'un histogramme mono-dimensionnel peut contenir les niveaux des pixels provenant de plusieurs régions différentes de l'image (voir figure 1.15). Les pixels de chaque région doivent être regroupés en une classe spécifique de pixels, définie par un domaine modal de l'espace couleur. Par conséquent, les méthodes de détection des domaines modaux traitent le cas où un intervalle représentant un pic d'un histogramme mono-dimensionnel contient les intervalles modaux de plusieurs domaines modaux.

Lorsque sont retrouvés les intervalles représentant les pics de chaque histogramme mono-dimensionnel, l'espace couleur est alors partitionné en parallélépipèdes rectangles grâce au produit cartésien de ces intervalles. Une analyse de la population des pixels dont les points-couleur se trouvent dans ces parallélépipèdes permet d'identifier les domaines modaux de l'espace couleur (voir figure 1.16).



(a) Image originale.

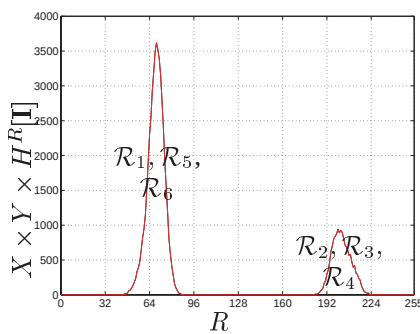
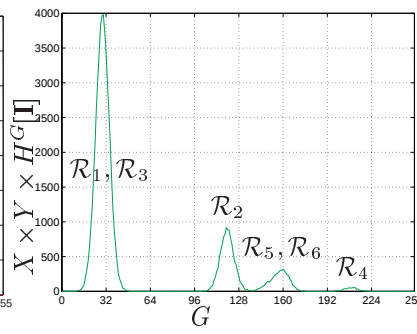
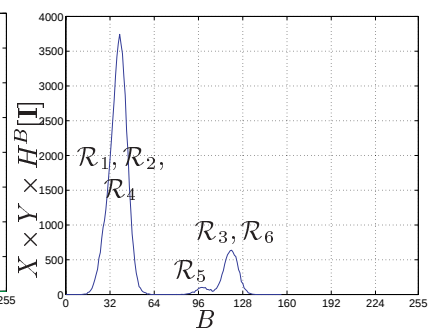
(b) Histogramme mono-dimensionnel de l'image de composante I^R .(c) Histogramme mono-dimensionnel de l'image de composante I^G .(d) Histogramme mono-dimensionnel de l'image de composante I^B .

FIG. 1.15 – Histogrammes mono-dimensionnels des images de composante issues de l'image couleur de la figure 1.15(a).

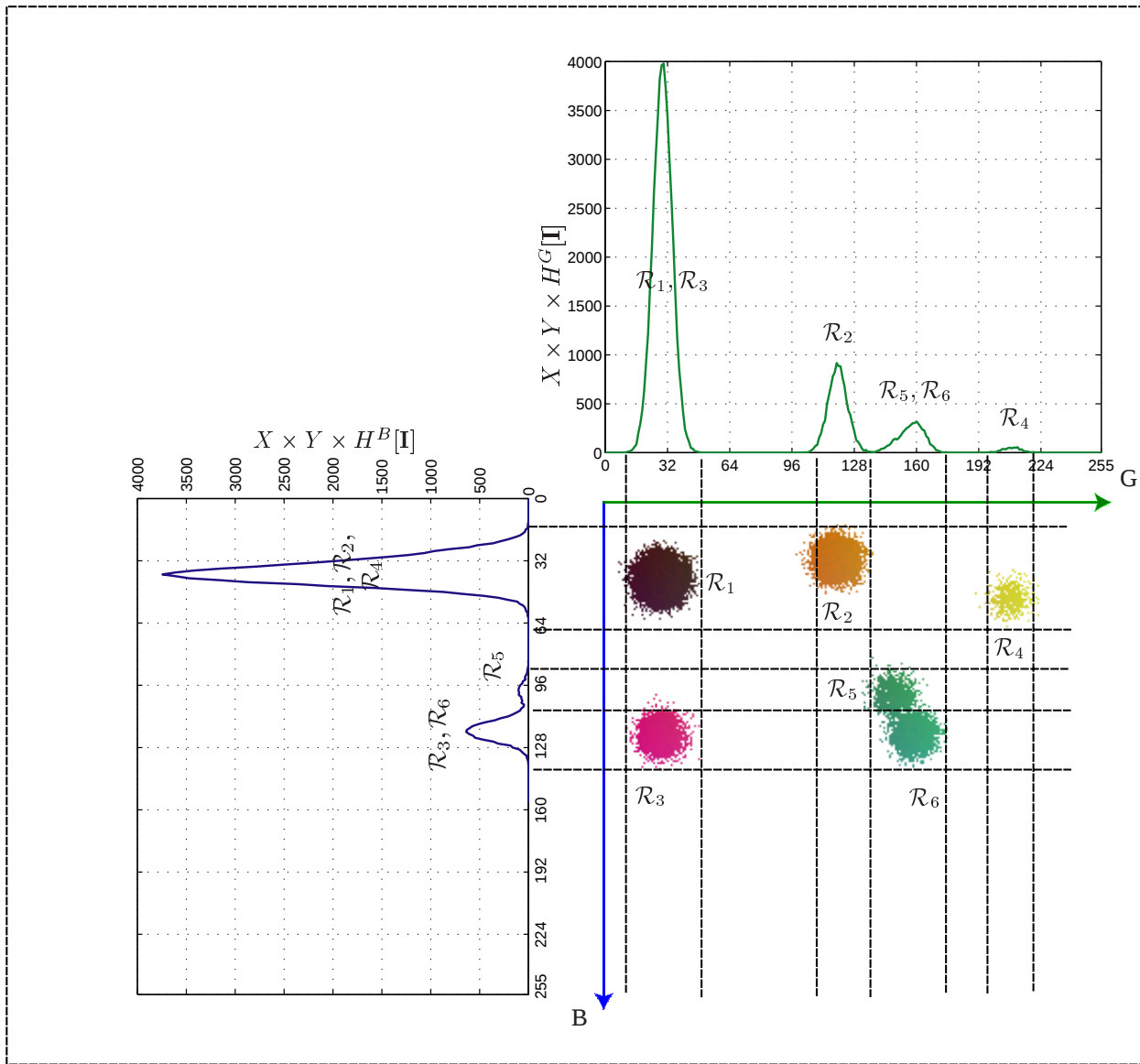


FIG. 1.16 – Détection des domaines modaux du plan chromatique (G, B) par analyse des histogrammes mono-dimensionnels des images de composante I^G et I^B de l'image de la figure 1.15(a).

L'approche par classification de pixels est efficace lorsque les nuages de points-couleur représentatifs des régions sont bien séparés dans l'espace couleur. C'est généralement le cas lorsque les images sont constituées d'un faible nombre de régions ayant des couleurs bien différenciées. Cependant, pour certaines images, le nombre de régions est élevé et leurs couleurs sont proches les unes des autres. Les nuages de points se chevauchent alors dans l'espace couleur, ce qui limite les performances de ces méthodes de classification automatique.

Par ailleurs, nous pouvons reprocher aux méthodes de classification de privilégier les propriétés colorimétriques des pixels par rapport à leurs propriétés spatiales. En effet, elles ne sont pas prises en compte assez tôt dans le processus de segmentation et n'interviennent que dans un deuxième temps afin de reconstruire les régions en recherchant des sous-ensembles de pixels connexes appartenant aux mêmes classes. Ainsi, lors de la classification, le contexte spatial est négligé et des nuages de points-couleur dans l'espace couleur peuvent ne pas correspondre à des régions significatives dans l'image analysée. **L'introduction de l'information spatiale dans le processus de construction des classes de pixels constitue la seconde partie de nos travaux.**

Nous avons vu que la phase de segmentation reconstruit les régions de l'image de telle sorte que chacune d'elles corresponde, dans la mesure du possible, à un objet dans la scène observée par la caméra. Les attributs extraits des régions sont alors examinés afin de reconnaître les objets représentés dans les images.

1.1.5 Exploitation de la couleur pour la reconnaissance d'objets

Lorsque les techniques de segmentation sont maîtrisées, nous pouvons passer à un niveau supérieur de l'analyse des images couleur, celui de la reconnaissance des objets. Nous abordons ce problème par la recherche de toutes les images d'une base qui contiennent le même objet que celui représenté dans une image requête. Nous supposons que toutes les images candidates ainsi que l'image requête contiennent un seul objet placé sur un fond uniforme. Ce problème est un cas particulier du domaine de la recherche d'images par l'exemple [TFMB04] (voir figure 1.17).

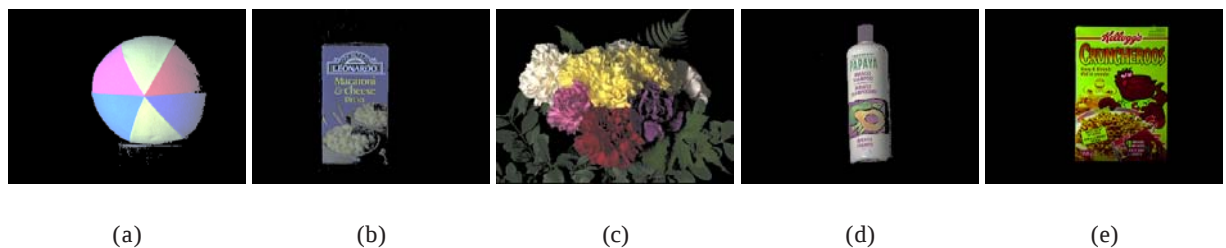


FIG. 1.17 – Les images (a),..., (e) contiennent un seul objet placé sur un fond uniforme. Ces images sont extraites de la base publique de l'Université Simon Fraser.

Une étape préalable de segmentation est nécessaire afin de ne considérer que les pixels qui représentent l'objet contenu dans chaque image. Ces pixels forment une ou plusieurs régions qui sont caractérisées par leurs attributs.

Les attributs des régions reflètent leurs propriétés géométriques [TW03], colorimétriques [MM96], ou une combinaison des deux [SG00, LRFt01]. Les attributs des régions qui représentent l'objet contenu dans l'image sont regroupés au sein d'une "signature" de l'image.

La reconnaissance d'objets est obtenue par la comparaison de la signature de l'image requête avec celles des images candidates. Cette comparaison est fondée sur une mesure de similarité entre ces signatures [SCZ01]. Les images candidates sont triées en fonction de mesures de similarité entre leurs signatures et celle de l'image requête. Sont retrouvées ainsi les images candidates qui contiennent des objets semblables à celui représenté dans l'image requête (voir figure 1.18).

L'histogramme couleur constitue l'une des signatures des images les plus utilisées [SB91]. La principale limitation de l'histogramme est qu'il ne fournit aucune information sur la répartition spatiale des pixels ayant des couleurs similaires dans l'image. C'est pourquoi des signatures spatio-colorimétriques ont été proposées, dont les plus significatives sont :

- le corrélogramme couleur, tableau indexé par des couples de couleurs, dont la $d^{\text{ième}}$ entrée pour le couple de couleurs (c_α, c_β) spécifie la probabilité de trouver un pixel de couleur c_α à une distance d dans l'image d'un pixel de couleur c_β [HKM⁺97]. Sensible à la résolution spatiale de l'image, le corrélogramme est encore plus encombrant en termes de place mémoire que l'histogramme couleur.
- le "joint-histogram" dont chaque cellule indique le nombre de pixels ayant des caractéristiques spatio-colorimétriques identiques, basées notamment sur la norme d'un gradient couleur [PZ99].
- le "spatial-chromatic-histogram" dont chaque cellule associée à un point-couleur contient le nombre

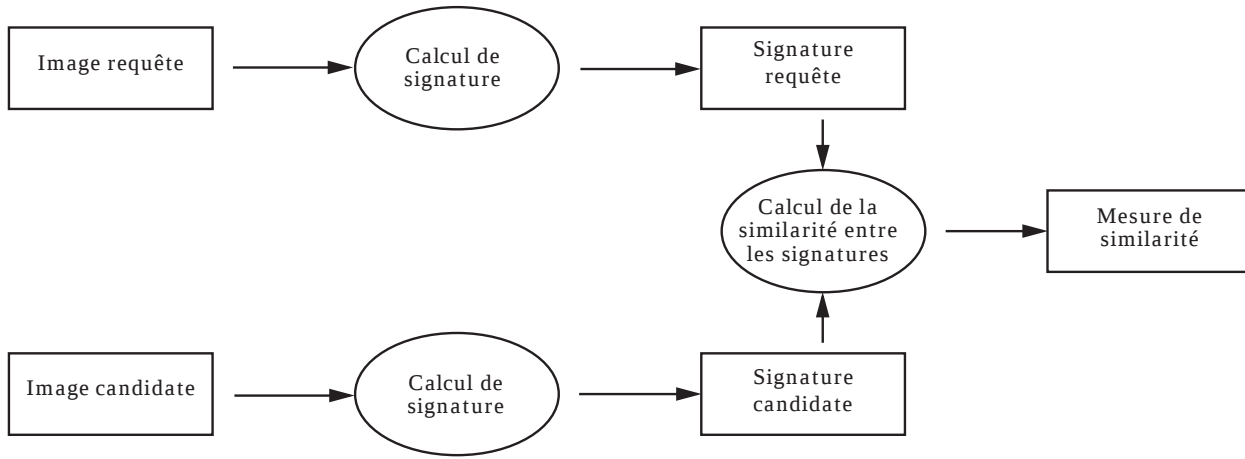


FIG. 1.18 – Comparaison de signatures.

de pixels caractérisés par ce point-couleur, les coordonnées du barycentre de ces pixels dans le plan image et la dispersion spatiale de ces pixels autour du barycentre [CCL⁺01]. L'information spatiale étant basée sur les coordonnées des pixels, cette signature est notamment sensible à toute translation spatiale des objets contenus dans les images.

Pour décrire comment la couleur peut être exploitée pour reconnaître des objets, nous nous proposons d'examiner l'une des signatures basées exclusivement sur les couleurs des pixels, l'histogramme couleur.

Plusieurs mesures de similarité entre histogrammes ont été proposées [Sch97b]. Nous retenons la mesure la plus courante et considérons que la similarité entre l'histogramme couleur $\mathbf{H}[\mathbf{I}_{req}]$ de l'image requête $\mathbf{I}_{req}$ et l'histogramme couleur $\mathbf{H}[\mathbf{I}_{cand}]$ de l'une des images candidates $\mathbf{I}_{cand}$ est mesurée grâce à leur intersection [SB91] :

$$Inter(\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}]) = \sum_{c \in \xi} \min(h[\mathbf{I}_{req}](c), h[\mathbf{I}_{cand}](c)), \quad (1.5)$$

avec :

$$h[\mathbf{I}](c) = \frac{X \times Y \times \mathbf{H}[\mathbf{I}](c)}{O[\mathbf{I}]}, \quad (1.6)$$

où $O[\mathbf{I}]$ indique le nombre de pixels qui représentent l'objet dans l'image $\mathbf{I}$ de taille $X \times Y$. La valeur contenue dans la cellule $h[\mathbf{I}](c)$ indique le nombre de pixels de l'image $\mathbf{I}$ associés au point-couleur c normalisé par le nombre de pixels qui représentent l'objet dans l'image. L'intersection entre les histogrammes varie entre 0, indiquant que les objets contenus dans les deux images sont différents et 1, permettant de conclure que les objets sont semblables.

Cette mesure de similarité est basée sur la somme des minima entre les cellules des deux histogrammes associées aux mêmes points-couleur. Quand les images à comparer contiennent le même objet, la mesure n'est proche de 1 que si les points-couleur associés aux pixels qui représentent les mêmes éléments de surface dans les deux images sont identiques.

Or, nous avons vu que les coordonnées des points-couleur représentatifs des pixels dépendent entre autre des conditions d'acquisition. Ceci explique que les performances en termes de reconnaissance d'objets atteintes par l'intersection entre les histogrammes couleur ne sont satisfaisantes que lorsque les conditions d'acquisition des images sont parfaitement maîtrisées et identiques [FF95]. **La troisième partie de nos travaux aborde les problèmes soulevés par la reconnaissance d'objets**

par analyse de leurs couleurs lorsque les conditions d'acquisition diffèrent d'une image à l'autre, en particulier lorsque différents illuminants de caractéristiques inconnues sont utilisés lors des prises de vue.

1.2 Travaux présentés

Nous ne tenons pas à fournir une liste exhaustive des travaux auxquels nous avons apporté notre contribution. Ces travaux se sont déroulés dans le cadre de six thèses. Nous en avons co-encadrées cinq avec Jack-Gérard Postaire dont quatre ont été soutenues. De ces six thèses, nous avons retenu **trois thèmes** qui constituent les étapes d'une chaîne de traitements des images couleur.

Une fois l'image couleur acquise, se pose le problème de la sélection de l'espace de représentation de la couleur qui est le mieux adapté aux traitements à effectuer sur l'image. Des éléments de réponse sont apportés par une analyse bas-niveau de l'image qui étudie **le lien entre les systèmes de représentation de la couleur et les performances d'un algorithme de segmentation**. Cette analyse tente de déterminer les systèmes de représentation de la couleur adaptés à des problèmes spécifiques de segmentation d'images couleur. Elle est le fruit du travail effectué avec Nicolas Vandenbroucke durant sa thèse [Van00].

L'espace de représentation de la couleur étant sélectionné, nous abordons l'étape de niveau intermédiaire qui consiste à segmenter l'image par classification de pixels. Le point clé de la classification est **la construction de classes de pixels qui correspondent aux régions de l'image**. Cette démarche, initialisée par Michael Fontaine lors de sa thèse [Fon01], débouche sur le développement de procédures de segmentation par classification de pixels qui analysent des propriétés spatio-colorimétriques de sous-ensembles de pixels.

Lorsque les régions sont reconstruites par la segmentation, nous atteignons l'étape de haut niveau qui tente de reconnaître les objets contenus dans l'image. **Les problèmes soulevés par les modifications des images couleur consécutives à des changements d'illuminants pour la reconnaissance d'objets** sont au coeur de la thèse en cours de Damien Muselet. Nous tentons de les résoudre en proposant des méthodes spécifiques de reconnaissance d'objets par analyse d'histogrammes couleur.

Les travaux qui ne rentrent pas dans cette chaîne de traitements ne sont pas présentés dans ce mémoire. Ils ont principalement abordé la segmentation des images couleur et portent sur :

- la détection des pixels contours couleur [UtP96].

L'objectif de cette approche est d'éviter l'étape de seuillage par hystérésis des maxima locaux de la norme d'un gradient couleur effectuée par les procédures classiques de détection de pixels contours. L'ajustement des seuils qui influence la qualité de la détection des pixels contours couleur, se révèle délicat pour obtenir une image des pixels contours non bruitée et représentant l'ensemble des contours.

L'approche proposée s'appuie sur une méthode de relaxation basée sur un étiquetage probabiliste. Elle regroupe les pixels en deux classes, celles des pixels contours et celle des pixels non-contours. Pour ce faire, elle évalue les probabilités initiales d'appartenance des pixels aux deux classes en fonction des normes d'un gradient couleur. Les probabilités de chaque pixel d'appartenir aux deux classes sont alors mises à jour par une procédure itérative de relaxation qui tient compte des probabilités d'appartenance aux classes de ses pixels voisins.

Quoique très originale, la méthode proposée ne permet malheureusement pas d'atteindre les performances obtenues par des méthodes existantes de détection de pixels contours couleur en termes de localisation et de détection. Ce travail effectué avec Vincent Ultré lors de sa thèse [Ult96], a été le

point de départ de notre démarche portant sur l'exploitation de la couleur pour l'analyse des images. Il a été pour nous l'occasion de découvrir la spécificité des problèmes posés par l'analyse des images couleur. Toutes ces raisons expliquent pourquoi nous avons choisi de ne pas relater ce travail dans ce mémoire.

- l'analyse d'images couleur du fond de l'oeil [RtPV98].

Cette étude s'inscrit dans le développement d'un système d'analyse automatique d'images couleur du fond de l'oeil afin de suivre l'évolution de lésions rétinienne au cours de traitements thérapeutiques. La problématique consiste à retrouver et localiser les régions représentant les lésions dans les images du fond de l'oeil d'un même patient acquises à différents stades de sa maladie. Comme les vaisseaux rétiens sont les repères les mieux adaptés à la localisation des lésions détectées, nous avons développé un algorithme de reconstruction des vaisseaux rétiens grâce à une méthode itérative de suivi de leurs contours. Les informations spatiales et colorimétriques sont fusionnées grâce à une coopération de techniques de segmentation couleur orientées contours et régions.

Ce travail, effectué avec Vénérée Rakotomalala durant sa thèse [Rak99], a répondu à un problème très spécifique posé par l'analyse d'images médicales. De ce fait, nous ne jugeons pas opportun d'en détailler son contenu au sein de ce rapport.

- la construction de classes de pixels par une détection floue des domaines modaux d'un espace couleur [GtBLP02].

Pour ce faire, nous considérons que les points-couleur représentant les pixels de l'image constituent un sous-ensemble flou caractérisé par sa fonction d'appartenance à un domaine modal. Cette fonction est évaluée grâce à une analyse de la concavité de l'histogramme couleur. Des transformations morphologiques floues sont ensuite appliquées de manière itérative à cette fonction d'appartenance afin de mettre en évidence les domaines modaux de l'espace couleur.

Ce travail est le fruit d'une collaboration avec Aymeric Gillet doctorant, et Claudine Lecocq sa co-directrice de thèse, spécialiste de l'analyse de données multi-dimensionnelles par transformations morphologiques floues. Il constitue une application des recherches fondamentales sur la détection des domaines modaux développées par Claudine Lecocq à la segmentation des images couleur. Il appartient donc à Claudine Lecocq de le faire figurer dans ses travaux de recherche.

Une partie des articles joints à ce document présentent les contenus de ces travaux qui ne sont pas évoqués dans ce mémoire, et complètent ainsi notre présentation. Nous nous proposons maintenant d'évoquer nos contributions répondant aux problématiques s'inscrivant dans les trois thèmes retenus et précédemment cités.

1.2.1 Systèmes de représentation de la couleur et segmentation

La problématique

Différents systèmes pouvant être retenus pour représenter les couleurs des pixels d'une image, l'un des premiers phénomènes à étudier est l'influence du système utilisé sur les performances des algorithmes de segmentation.

Pour illustrer notre approche de ce problème, examinons la figure 1.19 extraite de [Van00] où chacune des images contient un joueur de football. Ces images ont subi un pré-traitement afin de ne retenir que les "pixels joueurs" qui représentent les joueurs. Ces joueurs sont divisés en quatre classes qui correspondent aux joueurs de deux équipes qui portent des tenues de couleurs différentes et aux deux gardiens de but.

Du fait que le short, le maillot et les chaussettes de chaque joueur sont de couleurs différentes, les composantes couleur des pixels joueurs d'une même classe peuvent être hétérogènes et très dispersées,

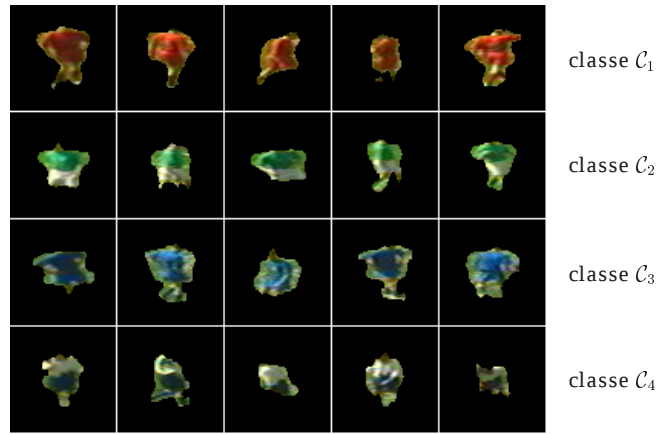


FIG. 1.19 – Ensemble des images contenant chacune un des joueurs sélectionnés pour cet exemple.

ce qui rend difficile la caractérisation d'une classe et sa discrimination vis-à-vis des autres classes.

Il est alors plus judicieux de caractériser un pixel joueur, non pas directement par ses composantes couleur mais par les moyennes des composantes couleur des pixels joueurs présents dans une certaine fenêtre de voisinage centrée sur ce pixel joueur. Nous proposons que chaque pixel joueur constitue une observation de la classe à laquelle il appartient. Chaque observation est caractérisée par un vecteur d'attributs dont les coordonnées sont les moyennes des composantes couleur des pixels joueurs présents dans la fenêtre centrée sur le pixel joueur correspondant à cette observation.

Dans chacune des images, nous prélevons aléatoirement un nombre de pixels joueurs. Durant cette phase, nous prenons soin à ce qu'un pixel joueur ne soit pas prélevé plus d'une fois et à ce qu'il y ait le même nombre de pixels joueurs prélevés dans chaque image. La figure 1.20 représente dans l'espace (R, G, B) les moyennes des composantes des pixels joueurs calculées dans des fenêtres centrées sur les pixels joueurs prélevés.

Nous avons représenté sur la figure 1.21 les moyennes des composantes des pixels joueurs prélevés dans l'espace couleur (Y, Ch_1, Ch_2) de Carron [Car95]. Les figures 1.20 et 1.21 montrent que les observations forment des nuages de points plus compacts dans l'espace couleur (Y, Ch_1, Ch_2) de Carron que dans l'espace (R, G, B) . Cet exemple illustre l'influence de l'espace couleur utilisé pour représenter les nuages d'observations qui définissent les classes de pixels à distinguer.

Il semble judicieux de déterminer parmi les nombreux systèmes de représentation de la couleur présentés dans la figure 1.6, ceux qui permettent d'effectuer la segmentation des images la plus satisfaisante suivant des critères à définir. Vandenbroucke tente de recenser les auteurs qui comparent différents systèmes de représentation de la couleur afin de fournir des éléments de réponse à cette question [Van00].

Ces études peuvent être classées en deux catégories selon que les systèmes de représentation de la couleur les mieux adaptés sont déterminés selon un critère visuel ou selon un critère numérique. Le critère visuel correspond à une appréciation subjective de la qualité des images segmentées [OKS80, Dai89, LR97] tandis que le critère numérique est fondé sur une mesure quantitative sur ces images [OPR78, BCS98, Lez00].

Les critères utilisés se limitent au cas des images analysées et aux algorithmes utilisés. En effet, la démarche la plus courante consiste à segmenter une image par un algorithme puis de juger des résultats soit visuellement soit par la mesure d'un taux d'erreur. L'algorithme utilisé et le critère d'évaluation de la qualité d'un système influencent donc aussi le choix de celui-ci.

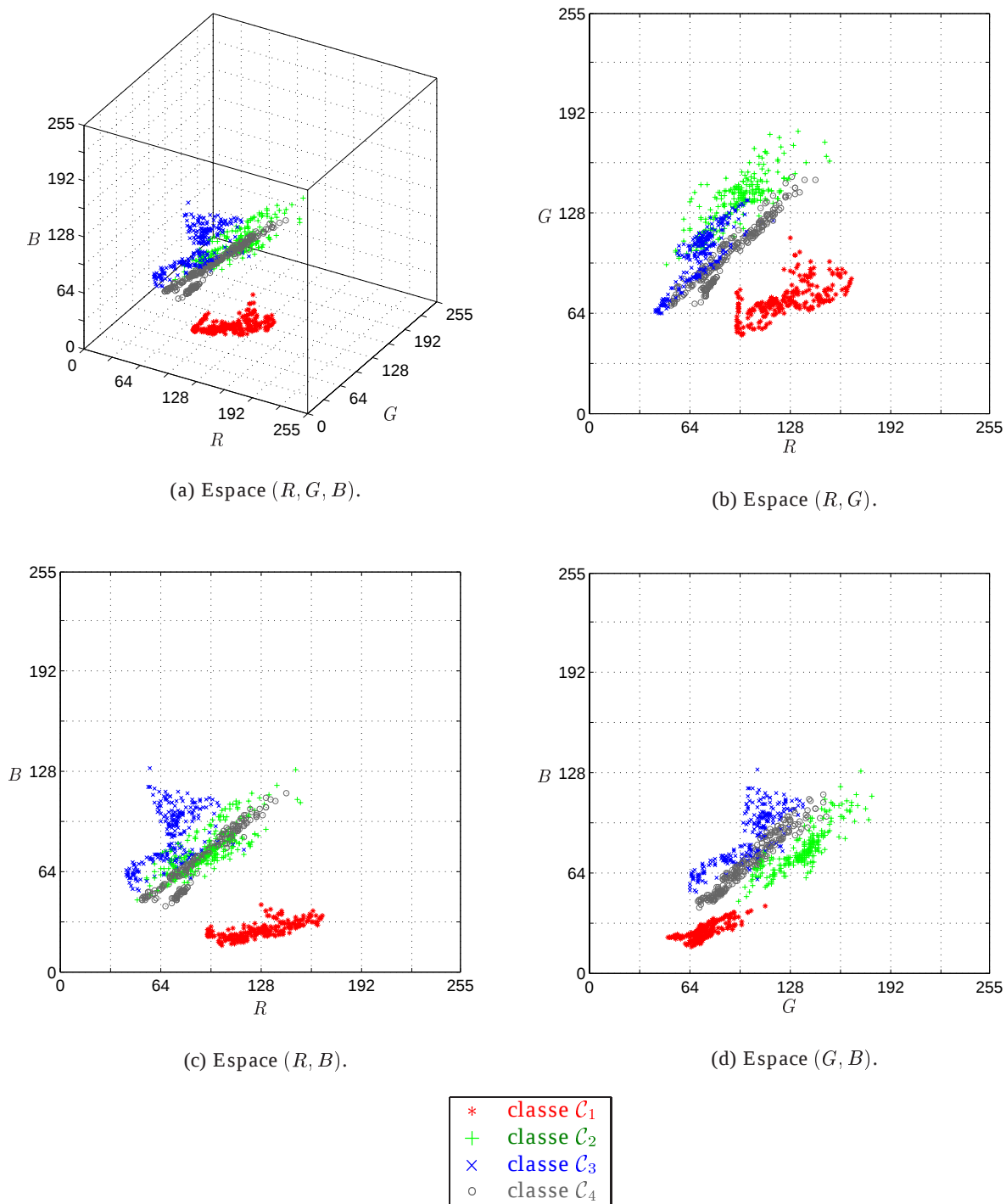


FIG. 1.20 – Représentation dans l'espace (R, G, B) des moyennes des composantes couleur calculées dans des fenêtres centrées sur les pixels joueurs qui ont été prélevés dans les images de la figure 1.19.

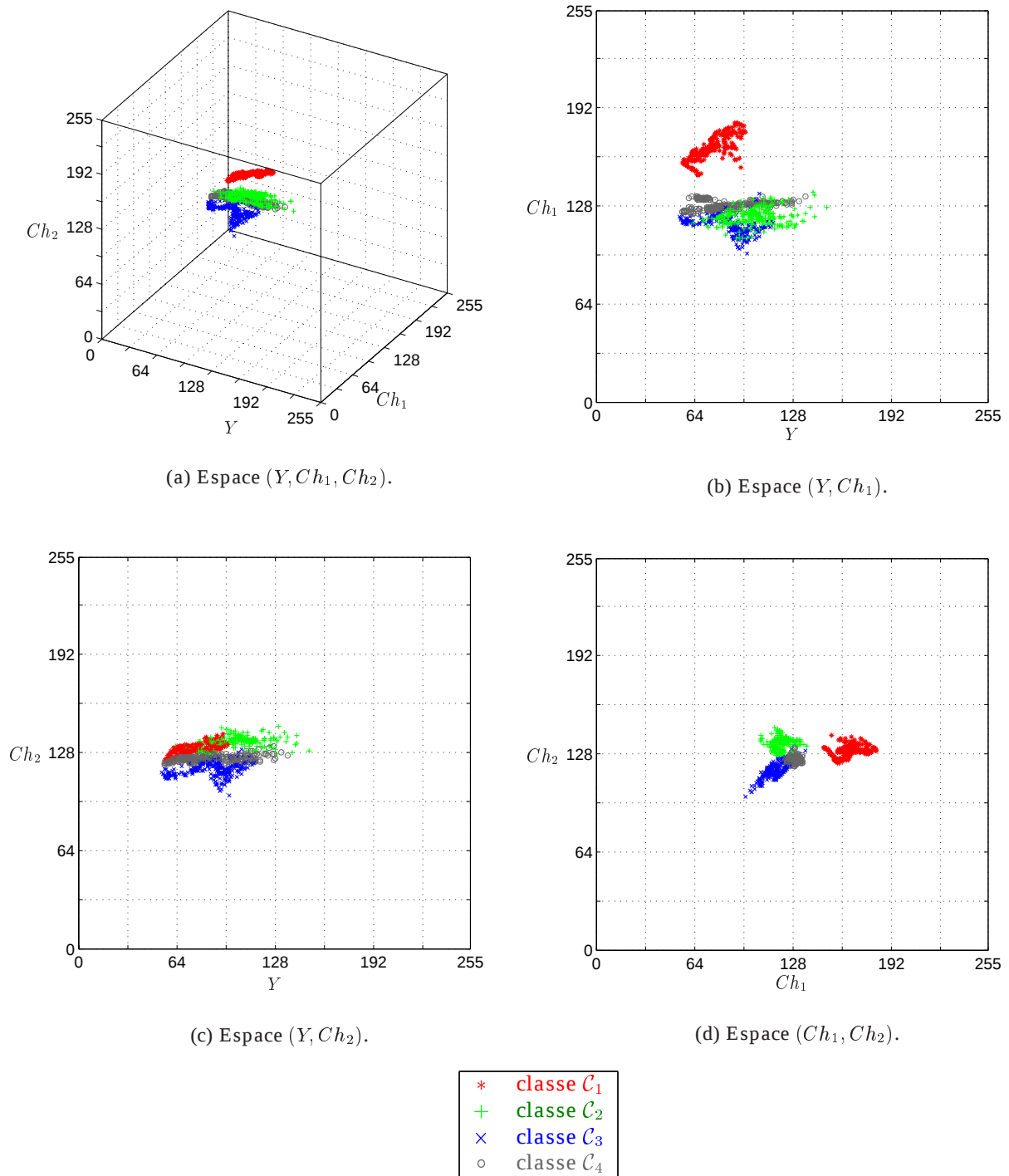


FIG. 1.21 – Représentation dans l'espace (Y, Ch_1, Ch_2) de Carron des moyennes des composantes couleur calculées dans des fenêtres centrées sur les pixels joueurs qui ont été prélevés dans les images de la figure 1.19.

Ces différents travaux ne permettent pas de dégager un système de représentation de la couleur adapté à tous les traitements et à toutes les images. Bien au contraire, le choix d'un système est d'abord fonction de l'application considérée, c'est-à-dire du type d'images traitées. Ainsi de nombreux auteurs réalisent une étude préalable sur les images qu'ils doivent analyser en vue de déterminer le système de représentation le plus adéquat, puis ils se limitent à l'utilisation de ce système.

Notre contribution

Dans le cadre de la segmentation d'images couleur par classification de pixels, nous proposons de comparer les composantes couleur issues des différents systèmes de représentation de la couleur, afin de déterminer celles qui sont les mieux adaptées pour représenter et différencier les classes de pixels à retrouver.

L'originalité de notre approche consiste à regrouper les composantes couleur qui permettent la meilleure discrimination possible entre les classes de pixels en présence dans un *système hybride de représentation de la couleur*. Ces composantes peuvent être issues de différents systèmes de représentation de la couleur. Pour ce faire, nous proposons une méthodologie de sélection de ces composantes fondée sur des critères mesurant la dispersion et la compacité des classes.

Par exemple, le système hybride sélectionné par notre algorithme pour séparer au mieux, au sens d'un critère défini ultérieurement, les classes de pixels joueurs de la figure 1.19 est l'espace $(x, Ch_2, I3)$. La composante x est issue du système de représentation (x, y, z) de la CIE, Ch_2 du système de Carron (Y, Ch_1, Ch_2) et $I3$ du système d'Ohta $(I1, I2, I3)$. Pour illustrer le pouvoir discriminant de cet espace couleur hybride $(x, Ch_2, I3)$, nous y avons représenté dans la figure 1.22 les observations correspondant aux pixels joueurs prélevés des images de la figure 1.19. En comparant les figures 1.20 et 1.22, nous constatons que les observations forment des nuages de points mieux séparés et plus compacts dans l'espace couleur hybride que dans l'espace (R, G, B) .

Dans le deuxième chapitre, nous présentons une méthode d'analyse discriminante qui détermine les composantes couleur les plus pertinentes constituant un espace hybride de représentation de la couleur adapté à la discrimination entre les classes de pixels en présence. Nous indiquons comment notre approche s'applique à la segmentation de séquences d'images couleur afin de suivre automatiquement des footballeurs en mouvement. Les pixels ainsi classés forment des régions dont les contours sont utilisés afin de modéliser chaque joueur par un contour actif. Tous les aspects concernant le suivi des joueurs ne sont pas traités dans ce rapport car il s'agit de problèmes différents d'analyse de scènes dynamiques que nous avons exclus de notre champ d'investigation.

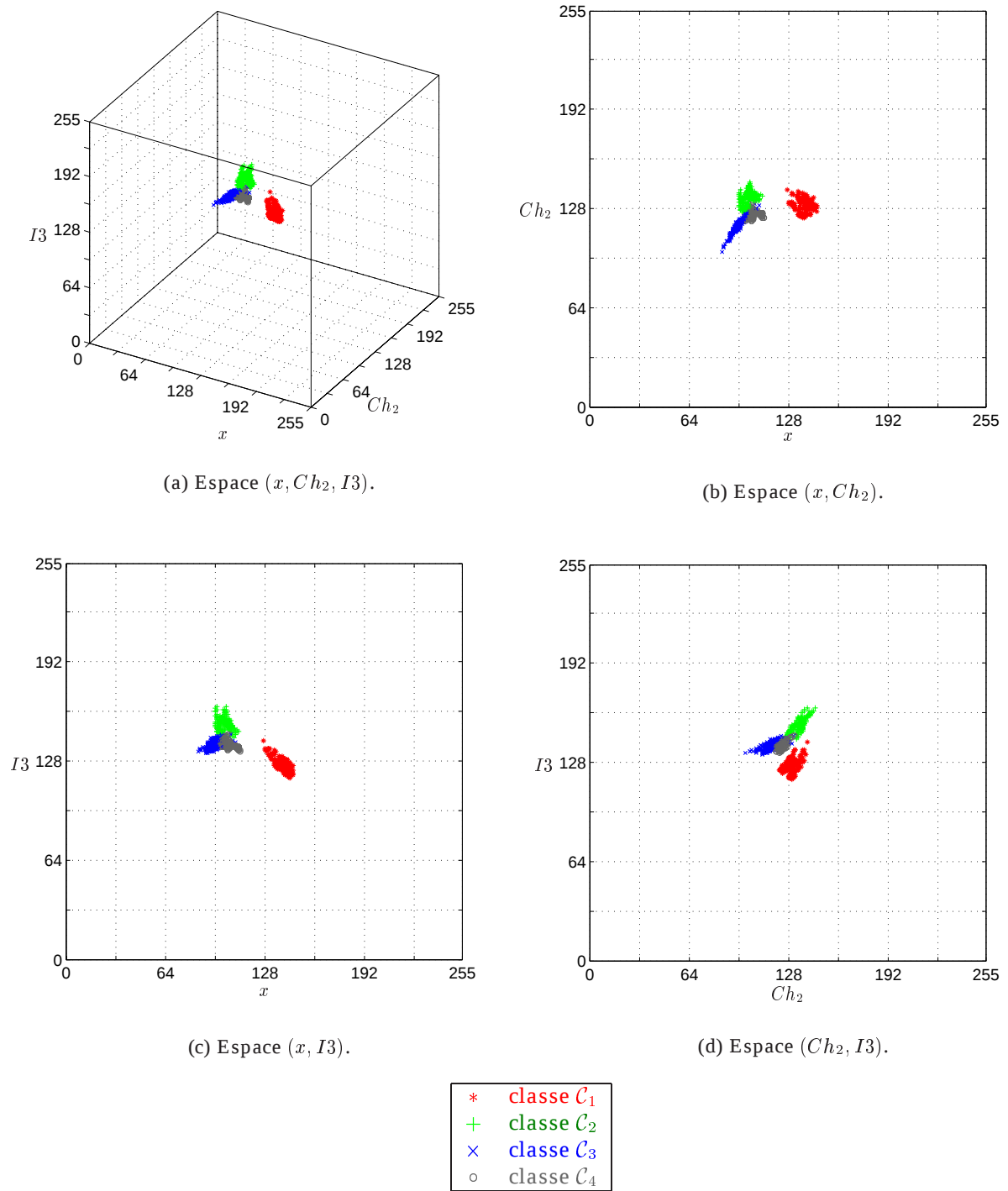


FIG. 1.22 – Représentation dans l'espace couleur hybride des moyennes des composantes des pixels joueurs calculées dans des fenêtres centrées sur les pixels joueurs qui ont été prélevés dans les images de la figure 1.19.

1.2.2 Régions de l'image et classes de pixels

La problématique

La problématique décrite ici porte sur la segmentation par classification de pixels. Généralement, les méthodes classiques de classification construisent les classes de pixels en n'analysant que les propriétés colorimétriques des pixels et ne prennent pas en compte la répartition spatiale des pixels ayant des couleurs similaires. Lors de la classification le contexte spatial est alors perdu, et des nuages de points-couleur dans l'espace couleur peuvent ne pas correspondre à des régions significatives dans l'image analysée.

Par exemple, bien qu'ils soient caractérisés par des couleurs vertes très proches, les deux carrés concentriques de l'image de la figure 1.23(a) peuvent être aisément et sans ambiguïté discriminés visuellement. Pour des raisons de simplicité de présentation, les niveaux de bleu des pixels sont mis à 0. Pourtant, les points-couleur représentant les pixels des deux carrés donnent naissance à un unique nuage dans le plan chromatique (R, G) (voir figure 1.23(b)). L'analyse de la distribution des points-couleur ne révélerait que très difficilement l'existence des deux carrés concentriques.

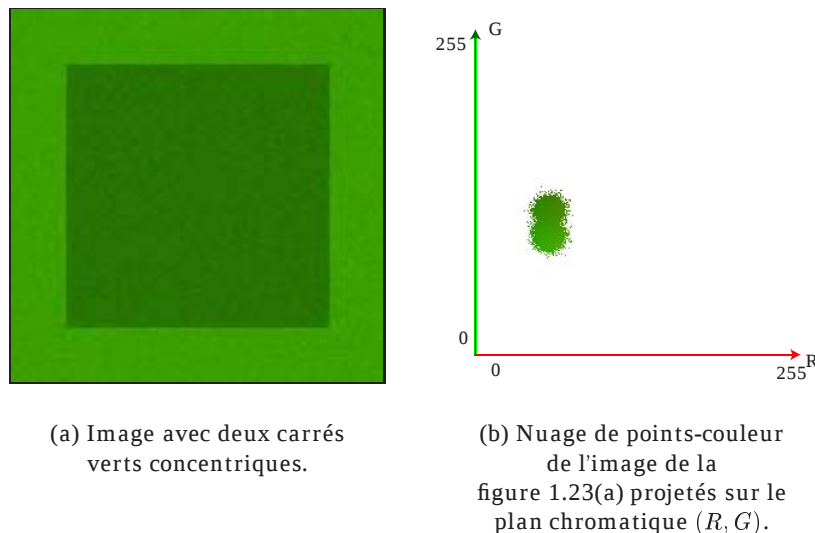


FIG. 1.23 – Deux carrés concentriques.

La prise en compte simultanée des propriétés colorimétriques des pixels et de l'arrangement spatial dans l'image des pixels ayant des couleurs similaires constitue donc une démarche naturelle pour construire les classes de pixels de telle sorte qu'elles correspondent effectivement aux régions de l'image.

Pourtant, peu de travaux introduisent l'information spatiale locale dès le début de la classification [CJSW01]. Parmi ceux-ci, les méthodes de construction des classes proposées par Orchard et par Balasubramanian minimisent la somme des distances euclidiennes entre les points-couleur représentant les pixels et les centres des classes auxquelles ils sont assignés, pondérées par une fonction qui dépend de la norme d'un gradient couleur [OB91, BAB94]. Les auteurs supposent que plus les couleurs d'une région constituée de pixels connexes assignés à une même classe sont homogènes, plus le poids de la distance associée à cette classe est faible.

Cheng et al. proposent d'analyser une mesure d'homogénéité des niveaux de chaque composante couleur pour construire les classes de pixels de chacune des trois images de composante [CJW02, CL03].

Cette procédure néglige le caractère multi-dimensionnel de l'information couleur et fournit souvent des images sur-segmentées. Les régions doivent alors être fusionnées à l'aide de l'analyse d'un graphe d'adjacence de régions. Cette approche très intéressante est sensible à l'ajustement des nombreux paramètres dont dépend le résultat.

Notre contribution

L'image de la figure 1.23(a) montre qu'il n'existe pas toujours une correspondance bijective entre les régions de l'image et les nuages de points-couleur dans l'espace couleur. Par conséquent, l'analyse des points-couleur ne permet pas toujours d'identifier les classes de pixels qui correspondent effectivement aux régions de l'image.

Plutôt que de supposer qu'il existe une correspondance entre chaque nuage de points-couleur et chaque classe de pixels, nous supposons qu'un nuage peut définir plusieurs classes. Cette hypothèse permet d'analyser les cas où les points-couleur représentatifs de différentes régions de l'image donnent naissance à un unique nuage dans l'espace couleur.

Pour construire les classes de pixels, nous proposons de considérer des parallélépipèdes rectangles de l'espace couleur. Appelés *domaines-couleur* de l'espace couleur, ils sont définis par leurs centres et leurs tailles dans l'espace couleur. Considérons par exemple la figure 1.24(b) qui montre la projection des points-couleur représentant les pixels de l'image de la figure 1.24(a) sur le plan chromatique (R, G) . Le rectangle dans la partie supérieure droite de ce plan couleur représente un domaine-couleur.

L'ensemble de tous les pixels dont les points-couleur se trouvent dans un domaine-couleur est appelé un *sous-ensemble-couleur*. Un domaine-couleur définit un unique sous-ensemble-couleur de pixels d'une image. La figure 1.24(c) affiche les pixels de l'image de la figure 1.24(a) qui appartiennent au sous-ensemble-couleur défini par le domaine-couleur présenté par la figure 1.24(b).

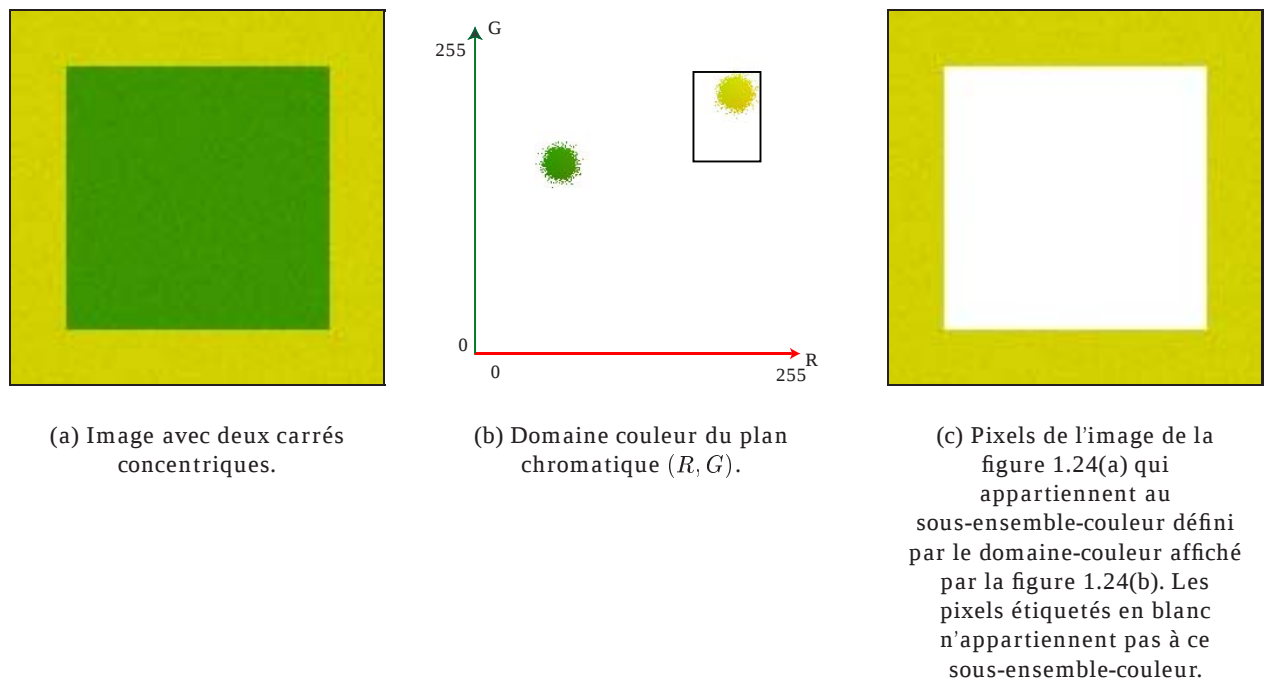


FIG. 1.24 – Un sous-ensemble-couleur de pixels défini par un domaine-couleur de l'espace couleur.

Examinons l'image de la figure 1.23(a) qui contient deux carrés verts concentriques, c'est à dire deux

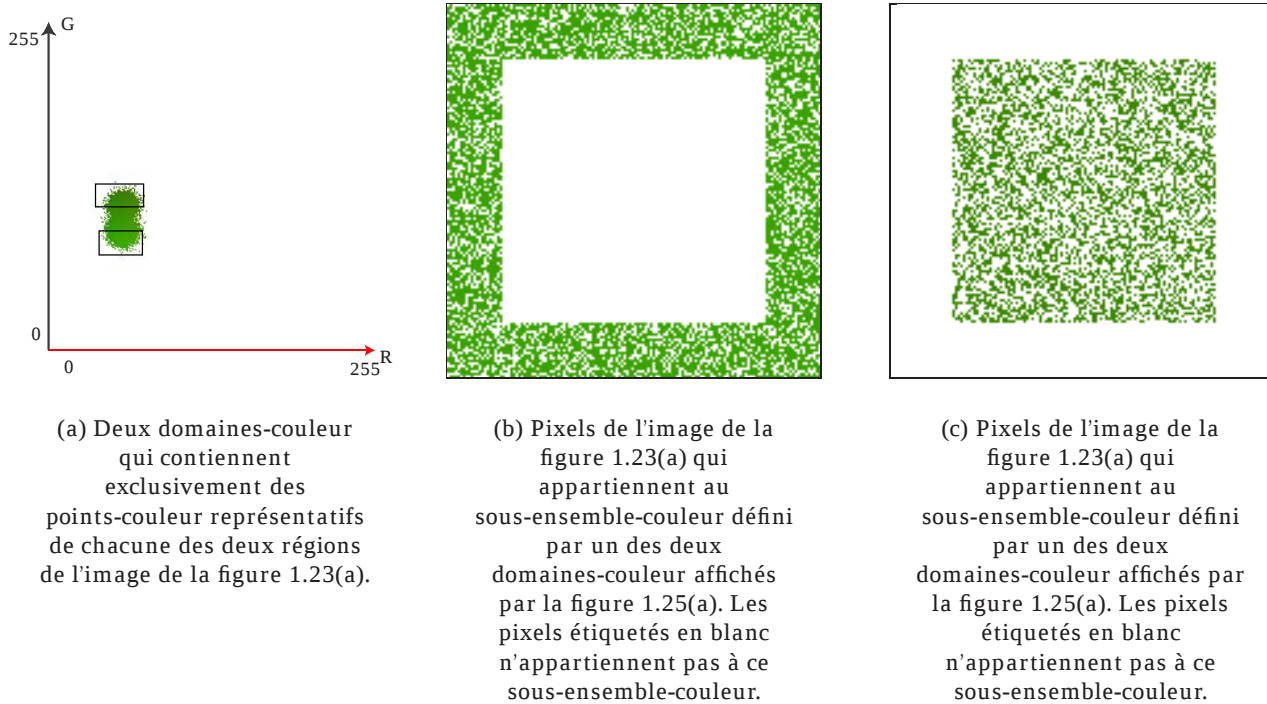


FIG. 1.25 – Deux sous-ensembles-couleur de l'image de la figure 1.23(a).

régions dont les points-couleur donnent naissance à un unique nuage de points-couleur dans le plan chromatique (R, G) (voir figure 1.23(b)). La figure 1.25(a) affiche deux domaines-couleur du plan chromatique (R, G) qui contiennent des points-couleur représentatifs des deux régions de l'image de la figure 1.23(a). Ces deux domaines-couleur ont été interactivement sélectionnés de sorte que chacun d'eux définisse un sous-ensemble-couleur qui contient un grand nombre de pixels appartenant exclusivement à l'une des deux régions. Comme les populations des points-couleur provenant des deux régions se chevauchent, ces deux domaines-couleur ne peuvent pas contenir la totalité des points-couleur représentant chacune d'elles. Les images des figures 1.25(b) and 1.25(c) montrent les pixels de l'image de la figure 1.23(a) qui constituent les sous-ensembles-couleur définis par ces deux domaines-couleur. Elles montrent que les pixels de chacun de ces deux sous-ensembles-couleur appartiennent exclusivement à chacune des deux régions, mais que tous les pixels de chaque région ne sont pas présents dans le sous-ensemble-couleur correspondant.

Nous considérons que les deux domaines-couleur affichés par la figure 1.25(a) définissent deux classes de pixels de l'image de la figure 1.23(a) qui correspondent effectivement aux deux régions de cette image. Lors de la dernière étape de la segmentation, les régions peuvent être reconstruites à partir des sous-ensembles-couleur définis par ces deux domaines-couleur grâce à des transformations morphologiques spécifiques.

Pour construire les deux classes de pixels correspondant aux deux régions de l'image de la figure 1.23(a), nous avons spécifié deux domaines-couleur pour définir deux sous-populations de points-couleur parmi l'ensemble de tous les points-couleur qui donnent naissance au nuage unique identifié dans l'espace couleur (voir figure 1.25(a)). Dans le contexte de la segmentation non supervisée par classification de pixels, nous ne disposons d'aucune information a priori pour spécifier les centres et tailles des domaines-couleur qui sont sélectionnés afin de définir les classes de pixels qui correspondent effectivement aux régions.

Pour sélectionner les domaines-couleur pertinents, nous proposons d'analyser les sous-ensembles-

couleur de pixels qu'il définissent, simultanément en termes de distribution des points-couleur dans l'espace couleur et de propriétés de connexité des pixels dans l'image.

Dans le troisième chapitre, nous décrivons la méthode proposée pour la construction spatio-colorimétrique des classes de pixels. Elle consiste à identifier des domaines-couleur de l'espace définissant des sous-ensembles-couleur constitués de pixels connexes dans l'image avec des couleurs les plus homogènes possible.

1.2.3 Reconnaissance d'objets sous éclairage non contrôlé

La problématique

Nous nous proposons d'aborder les problèmes soulevés par les modifications des images couleur consécutives à des changements d'illuminant pour la reconnaissance d'objets.

Les images considérées contiennent un seul objet par image placé sur un fond uniforme et sont acquises dans les conditions suivantes :

- les paramètres de réglage de la caméra ne sont pas modifiés entre les acquisitions,
- les modifications d'éclairage résultent de changements d'illuminant, c'est à dire d'une modification de la température et/ou de l'intensité de l'illuminant utilisé.

Examinons les cinq images de la figure 1.26 qui ont été acquises en respectant les conditions précédemment décrites. Dans cet exemple, les pixels de mêmes coordonnées spatiales dans ces images représentent les mêmes éléments de surface de l'objet.



FIG. 1.26 – Les images (a),..., (e) contiennent le même objet éclairé par un dispositif spécifique à chaque image. Ces images extraites de la base de l'Université Simon Fraser sont disponibles sur l'internet à l'adresse : http://www.cs.sfu.ca/~colour/image_db.

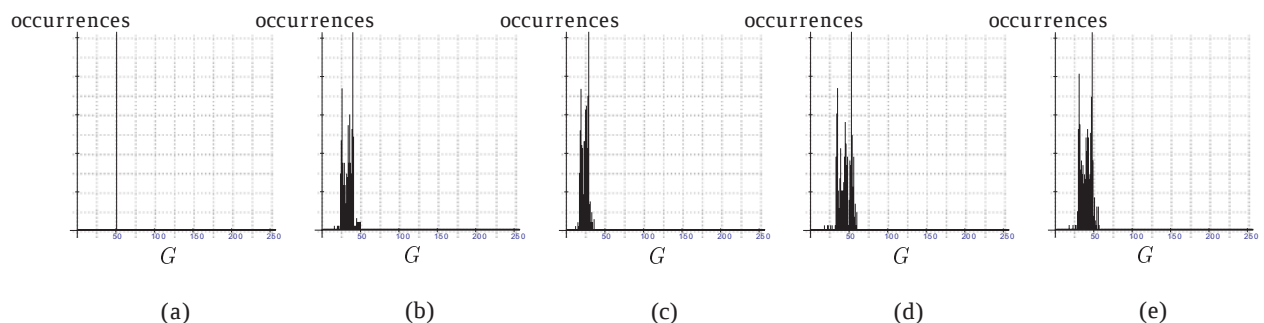


FIG. 1.27 – Histogrammes des niveaux de vert des pixels caractérisés par le niveau de vert égal à 50 dans l'image de la figure 1.26(a). Les histogrammes des figures 1.27(a),..., 1.27(e) représentent respectivement les occurrences de ces pixels dans les images des figures 1.26(a),..., 1.26(e).

Histogramme couleur et changements d'illuminant La figure 1.27 montre les histogrammes des niveaux de vert dans les images des figures 1.26(b),...,1.26(e) des pixels caractérisés par le niveau de vert égal à 50 dans l'image de la figure 1.26(a). Les niveaux de vert de ces pixels dans ces quatre images ne sont pas égaux à 50 mais sont répartis autour d'autres valeurs. Cet exemple montre que des changements d'illuminant provoquent des modifications des niveaux des composantes des pixels qui représentent les mêmes éléments de surface dans ces images.

Le modèle de formation de la couleur décrit par l'équation (1.4) indique que les coordonnées du point-couleur $I(P)$ représentant un pixel P ne dépendent pas seulement de la réflectance spectrale de l'élément de surface qui se projette sur P , mais également de la répartition spectrale relative d'énergie de l'illuminant et des fonctions de sensibilité des capteurs de la caméra. Ceci explique pourquoi les coordonnées du point-couleur $I(P)$ ne peuvent pas être considérées comme caractérisant les propriétés intrinsèques de l'élément de surface projeté sur le pixel P .

Plaçons nous dans le cadre de la reconnaissance d'objets et considérons que l'image de la figure 1.26(a) constitue l'image requête et que l'une des images des figures 1.26(b),...,1.26(e) soit l'image candidate.

Les points-couleur représentant les pixels sur lesquels se projettent les mêmes éléments de surface dans l'image requête et l'image candidate sont significativement différents. L'intersection entre les histogrammes couleur n'atteint pas alors une valeur élevée, même si les images contiennent les mêmes objets observés selon les mêmes conditions de prises de vue, mais éclairés avec des illuminants différents.

Cet exemple explique pourquoi l'intersection entre les histogrammes couleur fournit de mauvais résultats de reconnaissance d'objets quand différents illuminants de caractéristiques inconnues sont utilisés lors des acquisitions de l'image requête et des images candidates [FF95].

Rubner propose d'appliquer la distance EMD (*Earth Mover Distance*) [WPR85] pour comparer deux histogrammes couleur d'images acquises avec des illuminants différents [RTG98]. L'estimation de cette distance fait apparaître un problème d'optimisation sous contraintes, dont la résolution passe par des techniques d'optimisation linéaires classiques, telles que les méthodes de type simplex. La complexité introduite par l'algorithme d'optimisation nécessaire à l'évaluation de cette distance la rend très coûteuse en temps de calcul.

C'est pour cette raison que les images sont généralement comparées à l'aide de l'intersection entre histogrammes dits "invariants", qui sont aussi peu sensibles que possible à tout changement d'illuminant [tM03].

Ces histogrammes invariants sont soit les histogrammes des images couleur dites "invariantes" qui résultent de transformations des images couleur originales, soit les résultats de transformations des histogrammes couleur des images originales.

Histogramme d'image couleur invariante L'histogramme couleur de l'image couleur invariante $G(I)$ issue de la transformation G de l'image I , constitue l'histogramme couleur invariant $H_{inv}[I]$ de l'image I :

$$H_{inv}[I] = H[G(I)]. \quad (1.7)$$

Deux approches principales transforment l'image couleur I en une image invariante $G(I)$ où les pixels sont caractérisés par des points-couleur invariants. Elles consistent soit à estimer les points-couleur invariants de telle sorte qu'ils représentent la réflectance spectrale des surfaces observées, soit à évaluer les points-couleur invariants qui sont aussi peu sensibles que possible aux changements d'illuminant.

Les méthodes liées à la première approche déterminent les propriétés de réflectance des éléments de surface par analyse de l'image. Ceci nécessite d'estimer l'illuminant utilisé lors de l'acquisition [FHH02, FHH01]. Cette approche n'est pas adaptée lorsque nous ne disposons d'aucune information ni sur les caractéristiques de la caméra, ni sur la gamme d'illuminants utilisés.

Les méthodes relatives à la seconde approche supposent que les éléments de surface qui se projettent sur des pixels voisins sont éclairés selon les mêmes conditions d'éclairage [FF95, GS99a]. Sous l'hypothèse restrictive d'un changement d'intensité de l'illuminant, elles évaluent les coordonnées des points-couleur invariants représentant chaque pixel à partir des coordonnées des points-couleur associés à ses voisins.

La reconnaissance d'objets exploitant l'une de ces deux approches est divisée en deux étapes successives, le calcul de l'image invariante, puis la détermination de l'histogramme de cette image. Pour réduire les temps de calcul, de nombreux auteurs préfèrent calculer l'histogramme couleur invariant directement à partir de l'histogramme couleur de l'image couleur originale.

Histogramme couleur invariant L'histogramme couleur invariant $\mathbf{H}_{inv}[\mathbf{I}]$ de l'image $\mathbf{I}$ est alors obtenu par la transformation $\mathbf{T}_{inv}$ de son histogramme couleur $\mathbf{H}[\mathbf{I}]$:

$$\mathbf{H}_{inv}[\mathbf{I}] = \mathbf{T}_{inv}(\mathbf{H}[\mathbf{I}]). \quad (1.8)$$

La transformation d'un histogramme revient à modifier les valeurs de ses cellules, et peut donc s'implanter à l'aide d'une table de correspondance (look-up table). La transformation $\mathbf{T}_{inv}$ est associée à une table de correspondance $\mathbf{LUT}_{inv}[\mathbf{I}]$ de telle sorte que la valeur de la cellule $\mathbf{H}_{inv}[\mathbf{I}](\mathbf{x})$ soit exprimée par :

$$\mathbf{H}_{inv}[\mathbf{I}](\mathbf{x}) = \sum_{\mathbf{LUT}_{inv}[\mathbf{I}](\mathbf{c})=\mathbf{x}} \mathbf{H}[\mathbf{I}](\mathbf{c}). \quad (1.9)$$

La valeur de la cellule $\mathbf{H}[\mathbf{I}](\mathbf{c})$ de l'histogramme $\mathbf{H}[\mathbf{I}]$ associée au point-couleur $\mathbf{c}$ est cumulée dans la cellule $\mathbf{H}_{inv}[\mathbf{I}](\mathbf{x})$ de l'histogramme invariant $\mathbf{H}_{inv}[\mathbf{I}]$ associée au point-couleur $\mathbf{x}$, si la réponse de la table de correspondance $\mathbf{LUT}_{inv}[\mathbf{I}]$ au point-couleur $\mathbf{c}$ est égale à $\mathbf{x}$.

La définition de $\mathbf{T}_{inv}$ est basée soit sur des hypothèses portant sur la formation de la couleur, soit sur l'analyse de la distribution des points-couleur représentatifs des pixels de l'image.

Quand la définition de la table de correspondance est basée sur des hypothèses à propos de la formation de la couleur, elle ne dépend pas de l'image analysée. Les différentes tables de correspondance sont notamment obtenues par l'évaluation d'angles calculés à partir des coordonnées des points-couleur représentant les pixels [FCF95], ou par le passage de la représentation des couleurs dans l'espace (R, G, B) vers la représentation dans l'espace noté (l_1, l_2, l_3) , dont les composantes sont définies comme étant insensibles à un changement d'intensité d'un illuminant blanc [GS99b].

Les tables de correspondance obtenues par des analyses de la distribution des couleurs sont fournies par la normalisation "greyworld" [GJT88] qui divise les niveaux des composantes couleur des pixels par leurs moyennes dans l'image, la rotation des moments calculés avec les points-couleur représentant les pixels [LTM99], ainsi que la transformée de Karhunen-Loeve appliquée aux points-couleur [LWM02].

Pour illustrer la démarche de la comparaison d'images par l'intersection entre histogrammes couleur invariants, nous nous proposons d'examiner la figure 1.28. Elle montre la comparaison entre une image requête $\mathbf{I}_{req}$ et deux images candidates $\mathbf{I}_{cand1}$ et $\mathbf{I}_{cand2}$, par l'intersection entre les histogrammes couleur invariants calculés par la normalisation "greyworld". Les images $\mathbf{I}_{req}$ et $\mathbf{I}_{cand1}$ contiennent un

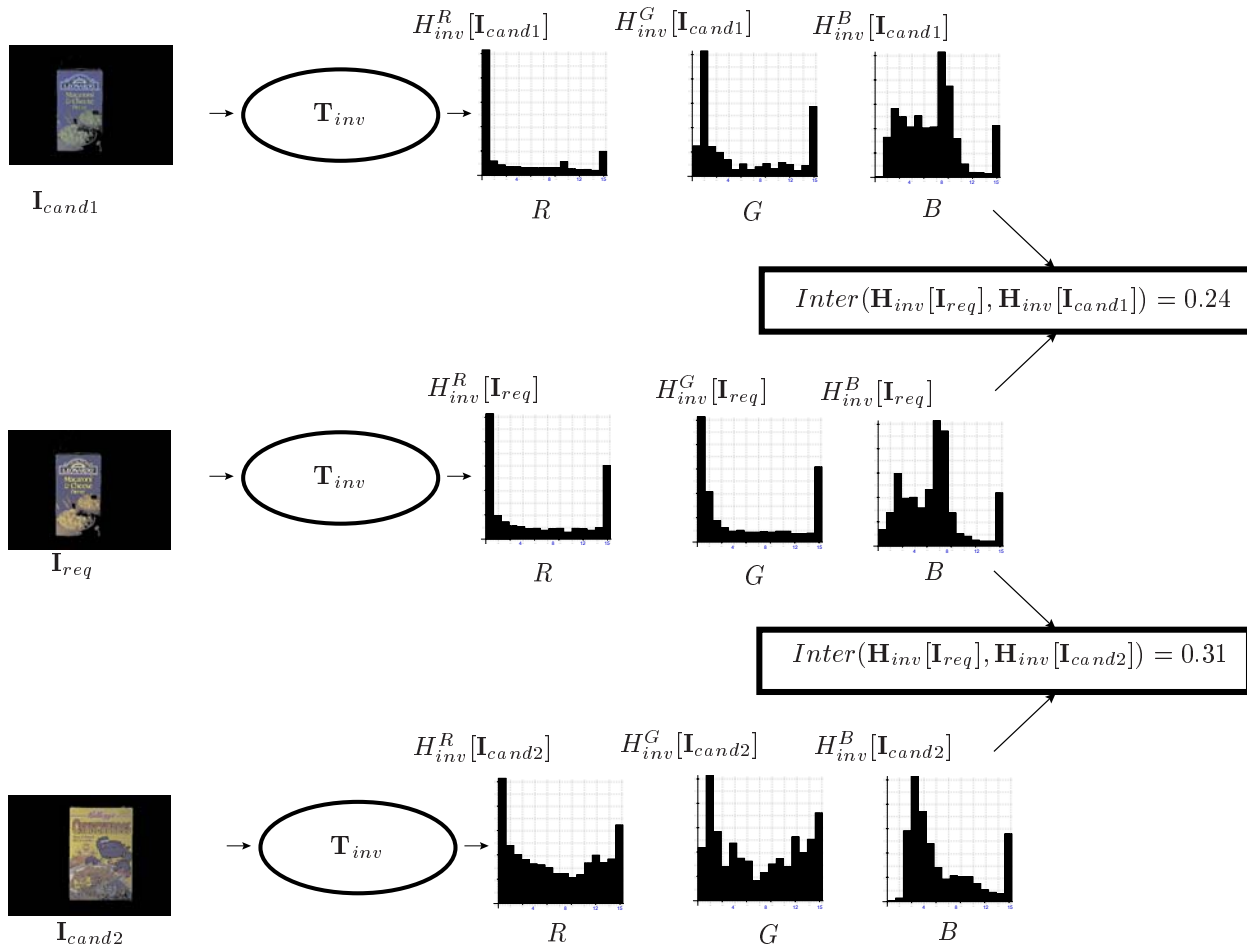


FIG. 1.28 – Comparaison entre l'image requête I_{req} et deux images candidates I_{cand1} et I_{cand2} par l'intersection entre les histogrammes couleur invariants obtenus par la normalisation "greyworld".

objet identique, différent de celui représenté par $\mathbf{I}_{cand2}$.

La figure 1.28 affiche les histogrammes mono-dimensionnels invariants des images, résultant des projections des histogrammes couleur invariants sur les composantes couleur. Cette figure montre que les histogrammes mono-dimensionnels invariants des images $\mathbf{I}_{req}$ et $\mathbf{I}_{cand1}$ ne sont pas identiques alors que les deux images contiennent le même objet. L'intersection entre leurs histogrammes couleur invariants ($\mathbf{H}_{inv}[\mathbf{I}_{req}], \mathbf{H}_{inv}[\mathbf{I}_{cand1}]$) n'atteint que la valeur 0.24. Alors que les images $\mathbf{I}_{req}$ et $\mathbf{I}_{cand2}$ représentent des objets différents, la valeur de l'intersection entre leurs histogrammes couleur invariants ($\mathbf{H}_{inv}[\mathbf{I}_{req}], \mathbf{H}_{inv}[\mathbf{I}_{cand2}]$) égale à 0.31, dépasse celle obtenue par la comparaison entre les deux images contenant le même objet.

Cet exemple montre que l'intersection entre les histogrammes couleur invariants ne permet pas toujours de discriminer les cas où les images requête et candidate contiennent le même objet des cas où les objets sont différents.

En effet, les fonctions permettant d'obtenir la plupart des histogrammes couleur invariants sont définies à partir de modèles, appelés modèles de changements d'illuminant, qui décrivent la modification de la couleur d'un pixel provoquée par un changement d'illuminant. Ces modèles ne sont pas directement issus du modèle de formation de la couleur décrit par l'équation (1.4), mais sont construits avec des hypothèses restrictives quant à la répartition spectrale d'énergie de l'illuminant et les fonctions de sensibilité des capteurs de la caméra.

Par exemple, la normalisation "greyworld" est basée sur le modèle diagonal qui suppose que les capteurs de la caméra ne sont sensibles que pour un intervalle très étroit de longueur d'onde. Les fonctions de sensibilité des capteurs et la répartition spectrale d'énergie de l'illuminant sont alors considérées comme constantes dans cet intervalle de longueur d'onde. Le niveau d'une composante couleur ne s'exprime plus par l'équation (1.4), mais par un simple produit de trois valeurs, chacune représentant respectivement la réflectance de l'élément de surface, l'illuminant et la sensibilité du capteur. Le modèle diagonal exprime la modification des coordonnées des points-couleur provoquée par un changement d'illuminant à l'aide d'une simple transformation linéaire. Or, les fonctions de sensibilité spectrale des capteurs qui tentent de s'approcher des fonctions colorimétriques définies par la CIE, ne sont pas sensibles à un seul intervalle de longueur d'onde (voir figure 1.3). Ce modèle se révèle être si grossier que les histogrammes invariants qui en découlent ne sont pas totalement insensibles à un changement d'illuminant.

Ceci constitue l'une des raisons pour laquelle les résultats expérimentaux obtenus en termes de reconnaissance d'objets par l'intersection entre les différents histogrammes couleur invariants ne sont pas très satisfaisants [FBM98, FS01].

Notre contribution

Nous pensons que l'exploitation de la couleur pour reconnaître des objets éclairés avec différents illuminants nécessite la mise en place d'une procédure spécifique de comparaison d'images. La démarche que nous proposons consiste à considérer chaque couple $(\mathbf{I}_{req}, \mathbf{I}_{cand})$ constitué de l'image requête $\mathbf{I}_{req}$ et de l'une des images candidates $\mathbf{I}_{cand}$.

La principale originalité de notre approche est de traiter le problème de la reconnaissance d'objets éclairés avec différents illuminants, non pas en analysant chaque image de la base indépendamment l'une de l'autre, mais en analysant chaque couple d'images $(\mathbf{I}_{req}, \mathbf{I}_{cand})$ construit durant la recherche des images candidates représentant un objet similaire à celui contenu dans l'image requête.

Plus précisément, nous proposons d'analyser chaque couple d'histogrammes couleur $(\mathbf{H}[\mathbf{I}_{req}],$

$\mathbf{H}[\mathbf{I}_{cand}]$) pour comparer le contenu de chaque couple d'images. Pour cela, la procédure détermine le couple d'histogrammes couleur dits "spécifiques" au couple d'histogrammes couleur considérés, de telle sorte que l'intersection entre les histogrammes couleur spécifiques soit élevée uniquement lorsque les deux objets contenus dans les deux images sont similaires.

Pour chaque comparaison entre l'image requête $\mathbf{I}_{req}$ et l'une des images candidates $\mathbf{I}_{cand}$, le couple d'histogrammes couleur spécifiques ($\mathbf{H}_{req,cand}[\mathbf{I}_{req}], \mathbf{H}_{req,cand}[\mathbf{I}_{cand}]$) est calculé à partir du couple d'histogrammes couleur ($\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}]$) par une transformation $\mathbf{T}_{req,cand}$:

$$(\mathbf{H}_{req,cand}[\mathbf{I}_{req}], \mathbf{H}_{req,cand}[\mathbf{I}_{cand}]) = \mathbf{T}_{req,cand}(\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}]). \quad (1.10)$$

$\mathbf{T}_{req,cand}$ est également associée à des tables de correspondance pour transformer le couple d'histogrammes couleur ($\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}]$) en un couple d'histogrammes couleur spécifiques ($\mathbf{H}_{req,cand}[\mathbf{I}_{req}], \mathbf{H}_{req,cand}[\mathbf{I}_{cand}]$).

Dans le quatrième chapitre, nous exposons un modèle original de changement d'illuminant à partir duquel la transformation $\mathbf{T}_{req,cand}$ est définie, de telle sorte que l'intersection entre les histogrammes couleur spécifiques permette de discriminer les cas où les objets contenus dans les images requête et candidate sont similaires des cas où ils sont différents.

Examinons la figure 1.29 qui montre la comparaison entre l'image requête $\mathbf{I}_{req}$ et les deux images candidates $\mathbf{I}_{cand1}$ and $\mathbf{I}_{cand2}$ identiques à celles de la figure 1.28, par l'intersection entre les histogrammes couleur spécifiques. Les images $\mathbf{I}_{req}$ et $\mathbf{I}_{cand1}$ contiennent le même objet, différent de celui représenté par l'image $\mathbf{I}_{cand2}$. Nous ne présentons pas cet exemple pour montrer l'apport de cette approche par rapport à celle s'appuyant sur les histogrammes couleur invariants, mais pour illustrer notre démarche.

Cette figure affiche les histogrammes mono-dimensionnels spécifiques qui sont les projections des histogrammes couleur spécifiques sur les composantes couleur. La figure montre que les histogrammes mono-dimensionnels spécifiques $H_{req,cand1}^i[\mathbf{I}_{req}], i = R, G, B$, de l'image requête calculés pour comparer les images $\mathbf{I}_{req}$ et $\mathbf{I}_{cand1}$ sont différents des histogrammes mono-dimensionnels spécifiques $H_{req,cand2}^i[\mathbf{I}_{req}], i = R, G, B$, de la même image requête calculés pour comparer les images $\mathbf{I}_{req}$ et $\mathbf{I}_{cand2}$. Elle met donc en évidence que l'histogramme couleur spécifique de l'image requête est calculé pour chaque comparaison avec une image candidate. Elle indique également que dans cet exemple, les valeurs d'intersection entre les histogrammes couleur spécifiques sont cohérentes avec les contenus des images.

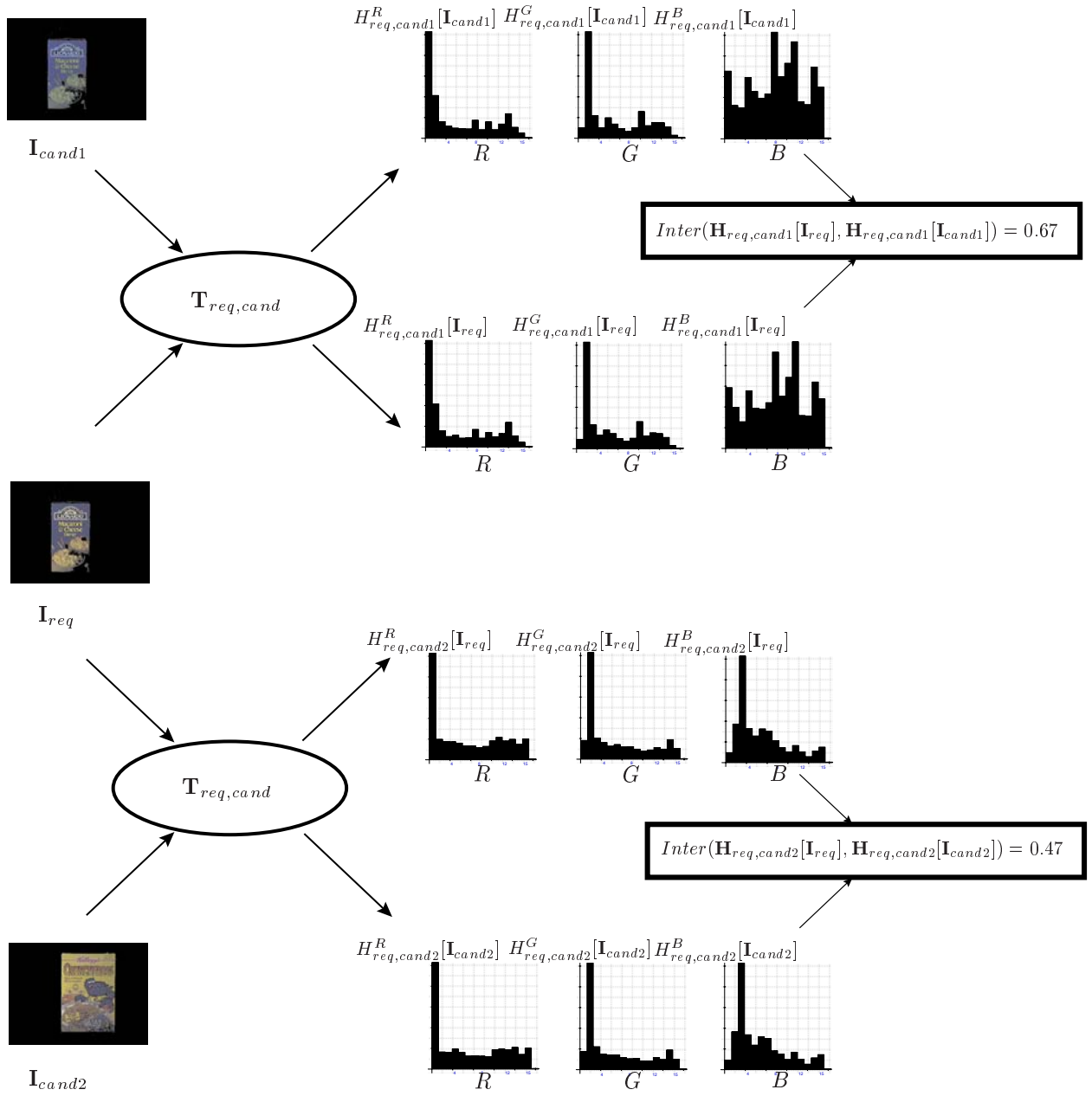


FIG. 1.29 – Comparaison entre l'image requête I_{req} et deux images candidates I_{cand1} et I_{cand2} par l'intersection entre les histogrammes couleur spécifiques.

1.3 Organisation du mémoire

Les trois chapitres suivants reprennent chacun l'un des trois thèmes précédemment décrits. L'ordre des chapitres respecte l'enchaînement des étapes d'une chaîne de traitements des images couleur.

Une fois l'image couleur acquise par la caméra, se pose alors la question du choix de l'espace de représentation de la couleur qui est le mieux adapté aux traitements à effectuer sur l'image. Dans le deuxième chapitre, nous décrivons la détermination d'espaces couleur hybrides adaptés à des problèmes spécifiques de segmentation. L'approche proposée analyse les distributions des points-couleur dans différents espaces de représentation de la couleur afin de constituer un espace hybride adapté aux classes de pixels à distinguer. Cette démarche est fondée sur l'analyse exclusive des propriétés colorimétriques des pixels pour la construction des classes de pixels.

Nous avons présenté un exemple (voir figure 1.23) qui montre qu'il n'existe pas toujours une correspondance bijective entre chaque nuage de points-couleur représentant les pixels dans un espace couleur et chaque région de l'image. Par conséquent, l'analyse exclusive des propriétés colorimétriques des pixels ne permet pas toujours de reconstruire les classes de pixels qui correspondent aux régions. Lors du troisième chapitre, nous détaillons une méthode de segmentation par classification de pixels qui tente de répondre à ce problème en analysant au même titre les propriétés colorimétriques des pixels et leur répartition spatiale dans l'image.

Une fois que les régions sont reconstruites, elles peuvent être analysées afin notamment de reconnaître les objets qu'elles représentent. Le quatrième chapitre est dédié à la reconnaissance d'objets éclairés avec différents illuminants. Nous y présentons un modèle original qui décrit les variations des couleurs des pixels provoquées par des changements d'illuminant. Ce modèle est utilisé afin de définir la fonction qui pour chaque comparaison entre l'image requête et une image candidate, transforme le couple constitué par les histogrammes couleur en un couple d'histogrammes couleur spécifiques. Ces histogrammes spécifiques sont calculés de telle sorte que leur intersection indique si les deux objets contenus dans les deux images sont semblables ou non, et ce quelque soient les sources d'éclairage utilisées lors des acquisitions des images.

Enfin, nous profitons du cinquième chapitre pour effectuer une synthèse des différents travaux présentés et tenter d'en extraire des perspectives.

1.4 Bibliographie

- [BAB94] R. Balasubramanian, J. P. Allebach et C. A. Bouman. "Color-image quantization with use of a fast binary splitting technique". *Journal of Optical Society of America*, 11(11):2777–2786, 1994.
- [BCS98] M. Borsotti, P. Campadelli et R. Schettini. "Quantitative evaluation of color image segmentation results". *Pattern Recognition Letters*, 19(8):741–747, 1998.
- [Car95] T. Carron. *Segmentation d'images couleur dans la base Teinte-Luminance-Saturatuion : approche numérique et symbolique*. Thèse de doctorat, Université de Savoie, décembre 1995.
- [CCL⁺01] L. Cinque, G. Ciocca, S. Levialdi, A. Pellicano et R. Schettini. "Color-based image retrieval using spatial-chromatic histograms". *Image and Vision Computing*, 19:979–986, 2001.
- [CJSW01] H. D. Cheng, X. H. Jiang, Y. Sun et J. Wang. "Color image segmentation : advances and projects". *Pattern Recognition*, 34(12):2259–2281, 2001.
- [CJW02] H. D. Cheng, X. Jiang et J. Wang. "Color image segmentation based on homogram thresholding and region merging". *Pattern Recognition*, 35(2):373–393, 2002.
- [CL02] T. Chen et Y. Lu. "Color image segmentation : an innovative approach". *Pattern Recognition*, 35(2):395–405, 2002.
- [CL03] H. D. Cheng et J. Li. "Fuzzy homogeneity and scale-space approach to color image segmentation". *Pattern Recognition*, 36(2):1545–1562, 2003.
- [Cou91] V. Coutance. *La couleur en vision par ordinateur. Application à la robotique*. Thèse de doctorat, Université Paul Sabatier de Toulouse, janvier 1991.

- [CP90] J. P. Cocquerez et S. Phillip. *Analyse d'images : filtrage et segmentation*. Masson, 1990.
- [Dai89] M. J. Daily. "Color image segmentation using markov random fields". In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 304–312, Palo Alo (Californie, USA), 1989.
- [FBM98] B. Funt, K. Barnard et L. Martin. "Is machine colour constancy good enough?". In *Proc. of the 5th European Conf. on Computer Vision*, pages 445–459, Freiburg (Allemagne), 1998.
- [FCF95] G. Finlayson, S. Chatterjee et B. Funt. "Color angle invariants for object recognition". In *Proc. of the IS&T/SID 3rd Color Imaging Conf.*, pages 44–47, Scottsdale (USA), 1995.
- [FF95] B. Funt et G. Finlayson. "Color constant color indexing". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(5):522–529, 1995.
- [FHH01] G. Finlayson, S. Hordley et P. Hubel. "Color by correlation : a simple, unifying framework for color constancy". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(11):1209–1221, 2001.
- [FHH02] G. Finlayson, S. Hordley et P. Hubel. "Illuminant estimation for object recognition". *Color Research and Application*, 27(4):260–270, 2002.
- [Fon01] M. Fontaine. *Segmentation non supervisée d'images couleur par analyse de la connexité des pixels*. Thèse de doctorat, Université Lille 1, decembre 2001.
- [FS01] G. Finlayson et G. Schaefer. "Colour indexing across devices and viewing conditions". In *Proc. of the 2nd Int. Workshop on Content-based MultiMedia Indexing*, pages 215–221, Brescia (Italie), 2001.
- [GJT88] R. Gershon, A. D. Jepson et J. K. Tsotsos. "From [r,g,b] to surface reflectance : computing color constant descriptors in images". *Perception*, pages 755–758, 1988.
- [GS99a] T. Gevers et A. W. M. Smeulders. "Color-based object recognition". *Pattern Recognition*, 32(3):453–464, 1999.
- [GS99b] T. Gevers et A. Smeulders. "Content-based image retrieval by viewpoint-invariant color indexing". *Image and Vision Computing*, 17:475–488, 1999.
- [GtBLP02] A. Gillet, **L. Macaire**, C. Botte-Lecocq et J. G. Postaire. Chapitre "Color image segmentation by analysis of 3D histogram with fuzzy morphological filters", pages 154–177, Dans *Fuzzy filters for image processing- Studies in fuzziness and soft computing*, Authors of the book : M. Nachtegael, D. Van der Weken, D. Van De Ville and E.E. Kerre. Springer-Verlag Editor, 2002.
- [HKM⁺97] J. Huang, S. R. Kumar, M. Mitra, W. Zhu et R. Zabih. "Image indexing using color correlogram". In *Proc. of the IEEE Int. Conf. on Computer Vision Pattern Recognition*, pages 762–768, San Juan (Porto Rico), 1997.
- [IK89] M. A. Ismail et M. S. Kamel. "Multidimensional data clustering utilizing hybrid search strategies". *Pattern Recognition*, 22(1):75–89, 1989.
- [Lam02] P. Lambert. *Etudes méthodologiques du filtrage et de la segmentation d'images multi-composantes*. Habilitation à Diriger des Recherches, Université de Savoie, Juillet 2002.
- [LC91] H. C. Lee et D. Cok. "Detection boundaries in a vector field". *IEEE Trans. on Signal Processing*, 39(5), 1991.
- [LC99] P. Lambert et T. Carron. "Symbolic fusion of luminance-hue-chroma features for region segmentation". *Pattern Recognition*, 32(11):1857–1872, 1999.
- [LCK94] J. H. Lee, B. H. Chang et S. D. Kim. "Comparison of colour transformations for image segmentation". *Electronics Letters*, 30(20):1660–1661, 1994.
- [Lez00] O. Lezoray. *Segmentation d'images couleur par morphologie mathématique et classification de données par réseaux de neurones : application à la classification de cellules en cytologie des séreuses*. Thèse de doctorat, Université de Caen, janvier 2000.
- [LR97] E. Littmann et H. Ritter. "Adaptive color segmentation. A comparison of neural and statistical methods". *IEEE Trans. on Neural Networks*, 8(1):175–185, 1997.
- [LRFt01] M. C. Larabi, N. Richard, C. Fernandez et **L. Macaire**. "Aide au diagnostic pour les cancers de peau basée sur une indexation par la couleur, la texture et la peau". Dans *Actes de ICISP'2001 Int. Conf. on Image and Signal Processing.*, pages 1055–1062, Agadir (Maroc), 2001.
- [Lt00] P. Lambert et **L. Macaire**. "Filtering and segmentation : the specificity of color images (keynote)". In *Proc. of the Int. Conf. on Color in Graphics and Image Processing*, volume 1, pages 57–64, Saint-Etienne, 2000.
- [LTM99] R. Lenz, L. Tran et P. Meer. "Moment based normalization of color images". In *Proc. of the IEEE Workshop on Multimedia Signal Processing*, pages 129–132, Copenhagen (Danemark), 1999.
- [LWM02] Z. Lin, J. Wang et K. Ma. "Using eigencolor normalization for illumination-invariant color object recognition". *Pattern Recognition*, 35(11):2629–2642, 2002.
- [MM96] B. Manjunath et W. Ma. "Texture features for browsing and retrieval of image data". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(8):837–842, 1996.

- [OB91] M. T. Orchard et C. A. Bouman. "Color quantization of images". *IEEE Trans. on Signal Processing*, 39(12) :2677–2690, 1991.
- [OKS80] Y. I. Ohta, T. Kanade et T. Sakai. "Color information for region segmentation". *Computer Graphics and Image Processing*, 13 :222–241, 1980.
- [OPR78] R. Ohlander, K. Price et D. R. Reddy. "Picture segmentation using a recursive region splitting method". *Computer Graphics and Image Processing*, 8 :313–333, 1978.
- [PY99] T. Pham et H. Yan. "Color image segmentation using fuzzy integral and mountain clustering". *Fuzzy Sets and Systems*, 107(2) :121–130, 1999.
- [PYL98] S. H. Park, I. D. Yun et S. U. Lee. "Color image segmentation based on 3D clustering : morphological approach". *Pattern Recognition*, 31(8) :1061–1076, 1998.
- [PZ99] G. Pass et R. Zabih. "Comparing images using joint histograms". *Multimedia Systems*, 7(3) :234–240, 1999.
- [Rak99] V. Rakotomalala. *Reconstruction bidimensionnelle de vaisseaux rétiniens par analyse d'images couleur du fond de l'oeil*. Thèse de doctorat, Université Lille 1, decembre 1999.
- [RTG98] Y. Rubner, C. Tomasi et L. Guibas. "A metric for distributions with applications to image databases". In *Proc. of the IEEE Int. Conf. on Computer Vision*, pages 59–66, Bombay, India, January 1998.
- [RtPV98] V. Rakotomalala, **L. Macaire**, J. G. Postaire et M. Valette. "Identification of retinal vessels by color image analysis". *Machine Graphics and Vision*, 7(4) :725–743, 1998.
- [SB91] M. J. Swain et D. H. Ballard. "Color indexing". *Int. Journal of Computer Vision*, 7(1) :11–32, 1991.
- [Sch97a] P. Scheunders. "A genetic c-means clustering algorithm applied to color image quantization". *Pattern Recognition*, 30(6) :859–866, 1997.
- [Sch97b] B. Schiele. *Object recognition using multidimensional receptive field histograms*. Phd thesis, Institut National Polytechnique de Grenoble, 1997.
- [SCZ01] R. Schettini, G. Ciocca et S. Zuffi. Chapter A Survey of methods for colour image indexing and retrieval in image databases, In *Color Imaging Science : Exploiting Digital Media*, Authors of the book : R. Luo and L. MacDonald. J. Wiley Eds., 2001.
- [SG00] A. Smeulders et T. Gevers. "Pictoseek : combining color and shape invariant features for image retrieval". *IEEE Trans. on Image Processing*, 9(1) :102–119, 2000.
- [SP97] L. Shafarenko et M. Petrou. "Automatic watershed segmentation of randomly textured color images". *IEEE Trans. on Image Processing*, 6(11) :1530–1543, 1997.
- [SPK98] L. Shafarenko, M. Petrou et J. V. Kittler. "Histogram based segmentation in a perceptually uniform color space". *IEEE Trans. on Image Processing*, 7(9) :1354–1358, 1998.
- [ST97] G. Sharma et H. J. Trussell. "Digital color imaging". *IEEE Trans. on Image Processing*, 6(7) :901–932, 1997.
- [Sèv96] R. Sève. *Physique de la couleur. De l'apparence colorée à la technique colorimétrique*. Masson, 1996.
- [TB97] A. Trémeau et N. Borel. "A region growing and merging algorithm to color segmentation". *Pattern Recognition*, 30(7) :1191–1203, 1997.
- [TC00] A. Trémeau et P. Colantoni. "Region adjacency graph applied to color image segmentation". *IEEE Trans. on Image Processing*, 9(4) :735–744, 2000.
- [TFMB04] A. Trémeau, C. Fernandez-Maloigne et P. Bonton. *Image couleur : de l'acquisition au traitement*. Editions Dunod, Paris, 2004.
- [tM03] **L. Macaire** et D. Muselet. "Invariants couleur - méthodes et applications". Dans *Actes de l'Ecole de printemps du GDR ISIS Operation Imagerie Couleur*, Dijon, avril 2003.
- [TNA95] K. Takahashi, H. Nakatani et K. Abe. "Color image segmentation using isodata clustering method". In *Proc. of the 2nd Asian Conf. on Computer Vision*, volume 1, pages 523–527, Singapore, 1995.
- [Tom92] S. Tominaga. "Color classification of natural color images". *Color Research and Application*, 17(4) :230–239, 1992.
- [TW03] S. Tabbone et L. Wendling. "Color and grey level object retrieval using a 3D representation of force histogram". *Image and Vision Computing*, 21(6) :483–495, 2003.
- [UA94] T. Uchiyama et M. A. Arbib. "Color image segmentation using competitive learning". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(12) :1197–1206, 1994.
- [Ult96] V. Ultré. *Contribution à la segmentation d'images de mosaïque couleur*. Thèse de doctorat, Université Lille 1, janvier 1996.
- [UtP96] V. Ultré, **L. Macaire** et J. G. Postaire. "Determining compatibility coefficients for color contour detection by relaxation". *Machine Graphics and Vision*, 5(1/2) :265–275, 1996.

- [Van00] N. Vandenbroucke. *Segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimétriques adaptés. Application à l'analyse d'images de football*. Thèse de doctorat, Université Lille 1, décembre 2000.
- [VtP98] N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color pixels classification in an hybrid color space". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 1, pages 176–180, Chicago (USA), 1998.
- [VtP00] N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color systems coding for color image processing". In *Proc. of the Int. Conf. on Color in Graphics and Image Processing*, volume 1, pages 180–185, Saint-Etienne, 2000.
- [VtP03] N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis.". *Computer Vision and Image Understanding*, 90(2) :190–216, 2003.
- [WPR85] M. Werman, S. Peleg et A. Rosenfeld. "A distance metric for multi-dimensional histograms". *Computer Vision Graphics and Image Processing*, 32 :328–336, 1985.
- [Zen86] S. D. Zenzo. "A note on the gradient of multi-image". *Computer Vision Graphics and Image Processing*, 33 :116–125, 1986.
- [Zuc76] S. Zucker. "Region growing : childhood and adolescence". *Computer Vision Graphics and Image Processing*, 5 :382–399, 1976.

Chapitre 2

Espaces couleur adaptés à la classification de pixels

2.1 Introduction

La performance d'une méthode de segmentation est connue pour dépendre de la sélection de l'espace couleur utilisé pour représenter les couleurs des pixels. De nombreux auteurs ont tenté de déterminer les espaces couleur qui sont les mieux appropriés pour résoudre leurs problèmes de segmentation des images couleur. Nous avons vu au paragraphe 1.2.1 de la page 58 qu'ils fournissent des réponses contradictoires sur la pertinence des espaces couleur dans ce contexte de segmentation [Van00].

Plutôt que de rechercher le meilleur espace couleur pour tous les problèmes de segmentation, nous proposons une approche originale afin d'améliorer les résultats de segmentation par classification de pixels. Nous définissons un nouveau type d'espaces couleur en sélectionnant un ensemble de composantes couleur qui proviennent des différents espaces couleur classiques présentés sur la figure 1.6 de la page 42. De tels espaces, qui ne reposent sur aucun fondement psycho-visuel ni physique, sont appelés des *espaces couleur hybrides*.

Nous proposons de classer les pixels dont les couleurs sont représentées dans un espace couleur hybride spécifiquement conçu pour permettre la meilleure discrimination possible entre les classes de pixels en présence dans l'image [VtP03]. Cet espace, appelé *l'espace couleur hybride adapté*, est construit grâce une procédure supervisée et séquentielle de sélection d'attributs décrite dans le premier paragraphe de ce chapitre. La dimension d'un tel espace n'est pas toujours égale à trois, comme c'est le cas pour les espaces couleur classiques.

La performance de notre approche est évaluée dans le deuxième paragraphe de ce chapitre où notre algorithme de segmentation est testé dans le contexte de l'analyse d'images couleur de matchs de football. L'équipe à laquelle appartient chaque joueur est caractérisée par les couleurs de sa tenue qui ne sont pas toujours homogènes. L'objectif de la procédure de segmentation consiste à extraire les régions significatives représentant les joueurs et à identifier leurs équipes.

Cette application est bien adaptée pour illustrer l'efficacité de notre méthode de segmentation, car de nombreux tests peuvent être menés avec un grand nombre de matchs de foot-ball [VtP97, VtV97, VtVP97, VtP98a, VtP00c, VtP00a, VtP00b]. Un des principaux problèmes est la détermination de l'espace couleur hybride adapté, c'est à dire l'espace qui permet d'obtenir le taux d'erreur de segmentation le plus faible. Comme les tailles des joueurs sont relativement petites dans les images, il est difficile de les séparer et d'identifier leurs équipes quand ils sont proches les uns des autres. L'étape de segmen-

tation des images étant une étape de bas-niveau intégrée à système automatique de suivi des joueurs, la qualité de la segmentation des images couleur est cruciale.

2.2 Construction de l'espace couleur hybride adapté

Dans ce paragraphe, nous présentons la construction de l'espace couleur hybride adapté à la discrimination entre les classes de pixels présentes dans une famille donnée d'images ou dans une séquence d'images. Cet espace couleur hybride adapté est déterminé dans un contexte supervisé. Dans la première partie de ce paragraphe, nous décrivons l'échantillon d'apprentissage constitué des pixels représentatifs des classes de pixels en présence.

Les composantes d'un espace couleur hybride proviennent de différents espaces couleur et peuvent alors être plus ou moins corrélées. Comme des composantes couleur corrélées ne sont pas a priori efficaces pour la discrimination entre les classes, une mesure de corrélation est introduite dans la seconde partie de ce paragraphe. Grâce à cette mesure, la construction de l'espace couleur hybride adapté ne considère que des espaces couleur hybrides candidats constitués de composantes couleur qui sont les moins corrélées possible.

Dans la troisième partie, nous présentons la procédure séquentielle de sélection des composantes couleur qui détermine l'espace couleur hybride adapté. Nous supposons que plus les nuages formés par les observations des classes de pixels sont séparés et compacts dans l'espace couleur hybride candidat, plus le pouvoir discriminant de l'espace considéré est élevé. Un critère pertinent est utilisé pour mesurer le pouvoir discriminant des espaces couleur hybrides candidats. Enfin, la détermination de la dimension de l'espace couleur hybride adapté est détaillée dans la dernière partie de ce paragraphe.

2.2.1 Échantillon d'apprentissage

Comme nous l'avons précédemment indiqué, l'approche proposée consiste à rechercher le meilleur espace couleur hybride pour discriminer les N_{class} classes de pixels en présence dans une famille donnée d'images.

Afin de déterminer l'espace couleur hybride adapté, nous définissons pour chaque classe de pixels $\mathcal{C}_j$, $j = 1, \dots, N_{class}$, un ensemble de N_ω pixels représentatifs $\omega_{i,j}$, $i = 1, \dots, N_\omega$. Nous montrerons dans le paragraphe suivant de ce chapitre comment ces pixels représentatifs peuvent être extraits des images par une procédure interactive conçue pour la segmentation des images d'une séquence d'un match de foot-ball.

Dans un espace couleur de dimension d , nous caractérisons chaque pixel $\omega_{i,j}$ représentatif de la classe $\mathcal{C}_j$, par une observation $\mathbf{X}_{i,j} = [x_{i,j}^1, \dots, x_{i,j}^k, \dots, x_{i,j}^d]^T$ où $x_{i,j}^k$ est le niveau de la k^{eme} composante couleur. Nous représentons l'ensemble des observations sous la forme d'une matrice $\mathbf{X}$:

$$\mathbf{X} = \begin{bmatrix} x_{1,1}^1 & \cdots & x_{N_\omega,1}^1 & \cdots & x_{i,j}^1 & \cdots & x_{1,N_{class}}^1 & \cdots & x_{N_\omega,N_{class}}^1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{1,1}^k & \cdots & x_{N_\omega,1}^k & \cdots & x_{i,j}^k & \cdots & x_{1,N_{class}}^k & \cdots & x_{N_\omega,N_{class}}^k \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{1,1}^d & \cdots & x_{N_\omega,1}^d & \cdots & x_{i,j}^d & \cdots & x_{1,N_{class}}^d & \cdots & x_{N_\omega,N_{class}}^d \end{bmatrix}. \quad (2.1)$$

Les lignes de la matrice $\mathbf{X}$ correspondent aux $(N_\omega \times N_{class})$ pixels représentatifs, tandis que les colonnes correspondent aux niveaux des d composantes couleur de chaque pixel représentatif.

Nous pouvons également exprimer la matrice $\mathbf{X}$ par :

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}^1 \\ \vdots \\ \mathbf{X}^k \\ \vdots \\ \mathbf{X}^d \end{bmatrix}, \quad (2.2)$$

où $\mathbf{X}^k = [x_{1,1}^k, \dots, x_{N_\omega,1}^k, x_{1,2}^k, \dots, x_{i,j}^k, \dots, x_{1,N_{class}}^k, \dots, x_{N_\omega,N_{class}}^k]$ contient les niveaux de la k^{eme} composante couleur des $(N_\omega \times N_{class})$ pixels représentatifs, ou par :

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{1,1} & \cdots & \mathbf{X}_{N_\omega,1} & \mathbf{X}_{1,2} & \cdots & \mathbf{X}_{i,j} & \cdots & \mathbf{X}_{1,N_{class}} & \cdots & \mathbf{X}_{N_\omega,N_{class}} \end{bmatrix}, \quad (2.3)$$

où $\mathbf{X}_{i,j}$ représente l'observation associée à chaque pixel représentatif $\omega_{i,j}$.

2.2.2 Espace couleur hybride candidat

La construction de l'espace couleur hybride adapté est basée sur une procédure de sélection des composantes couleur qui procède pas à pas [DL97]. A chaque itération de rang d , cette procédure séquentielle sélectionne une nouvelle composante couleur qui est associée aux $(d - 1)$ composantes couleur précédemment sélectionnées, afin de constituer *l'espace couleur hybride de dimension d le plus discriminant*, noté $\mathbf{S}^d$. La procédure est arrêtée à l'itération de rang $d = D$ quand le critère d'arrêt décrit dans la quatrième partie de ce paragraphe est vérifié. L'espace couleur hybride de dimension D le plus discriminant constitue l'espace couleur hybride adapté qui est alors utilisé pour classer les pixels.

Soit N_{comp} le nombre de composantes couleur examinées, c'est à dire le nombre de composantes couleur différentes définissant tous les espaces couleur classiques décrits par la figure 1.6 de la page 42.

A chaque itération de rang d ($d \geq 2$) de la procédure séquentielle de sélection, nous associons chacune des $(N_{comp} - d + 1)$ composantes couleur restantes aux $(d - 1)$ composantes couleur précédemment sélectionnées afin de constituer $(N_{comp} - d + 1)$ espaces couleur hybrides de dimension d , notés HS_a^d , avec $a = 1, \dots, (N_{comp} - d + 1)$.

Parmi ces composantes, plusieurs sont plus ou moins corrélées avec les $(d - 1)$ composantes couleur précédemment sélectionnées. Ohta et Tominaga montrent que lorsque les couleurs des pixels sont représentées dans un espace résultant de la transformation de Karhunen-Loève où les composantes sont non corrélées, la classification de ces pixels fournit des résultats satisfaisants [OKS80, Tom92]. Par ailleurs, Kouassi considère que les composantes couleur doivent être décorrélées de sorte que chacune d'elles puisse être analysée séparément [KGP01]. Comme nous recherchons l'espace couleur hybride le plus discriminant parmi un nombre important d'espaces candidats, nous éliminons dans un premier temps ceux qui contiennent des composantes couleur corrélées.

Lors de chaque itération de rang d de la procédure séquentielle de sélection, nous mesurons la corrélation entre les composantes couleur de chacun des $(N_{comp} - d + 1)$ espaces couleur hybrides HS_a^d , avec $a = 1, \dots, (N_{comp} - d + 1)$.

L'espace couleur hybride HS_a^d contient les $(d - 1)$ composantes couleur qui constituent l'espace couleur hybride de dimension $(d - 1)$ le plus discriminant. Le vecteur $\mathbf{X}^k$, $k = 1, \dots, d - 1$, défini par l'équation (2.2), est associé à chacune des composantes. La d^{eme} composante couleur de HS_a^d est l'une des $(N_{comp} - d + 1)$ composantes couleur restantes. Nous notons $\mathbf{X}^d$ le vecteur qui contient les niveaux de cette d^{eme} composante couleur pour les $(N_\omega \times N_{class})$ pixels représentatifs.

Pour mesurer le niveau de corrélation entre les composantes couleur de HS_a^d , nous considérons les $(d - 1)$ couples de vecteurs constitués de $\mathbf{X}^d$ et de l'un des vecteurs $\mathbf{X}^k$, $k = 1, \dots, d - 1$.

Soit $\text{cor}(\mathbf{X}^k, \mathbf{X}^d)$ la mesure de la corrélation entre la k^{eme} composante couleur, $k = 1, \dots, d - 1$, et la d^{eme} composante couleur, définie par :

$$\text{cor}(\mathbf{X}^k, \mathbf{X}^d) = \frac{\text{cov}(\mathbf{X}^k, \mathbf{X}^d)}{\sigma^k \times \sigma^d}, \quad (2.4)$$

où la co-variance $\text{cov}(\mathbf{X}^k, \mathbf{X}^d)$ entre la k^{eme} et la d^{eme} composantes couleur s'exprime sous la forme :

$$\text{cov}(\mathbf{X}^k, \mathbf{X}^d) = \frac{1}{N_{\text{class}} \times N_{\omega}} \times \sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} (x_{i,j}^k - m^k) \times (x_{i,j}^d - m^d). \quad (2.5)$$

m^k , m^d , σ^k et σ^d sont respectivement les moyennes et les écarts-type des k^{eme} et d^{eme} composantes couleur. Pour la k^{eme} couleur composante, il sont définis par :

$$m^k = \frac{1}{N_{\text{class}} \times N_{\omega}} \times \sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} x_{i,j}^k, \quad (2.6)$$

et :

$$\sigma^k = \frac{1}{\sqrt{N_{\text{class}} \times N_{\omega}}} \times \sqrt{\sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} (x_{i,j}^k - m^k)^2}. \quad (2.7)$$

Les niveaux de corrélation varient entre 0 et 1. Plus le niveau de corrélation est proche de 1, plus les composantes couleur sont considérées comme étant corrélées.

Le maximum des niveaux de corrélation entre la d^{eme} composante couleur et toutes les autres $(d - 1)$ composantes couleur est considéré comme une mesure de corrélation, notée $J_{\text{cor}}(HS_a^d)$, entre les d composantes couleur de l'espace couleur hybride HS_a^d :

$$J_{\text{cor}}(HS_a^d) = \max_{k=1}^{d-1} (\text{cor}(\mathbf{X}^d, \mathbf{X}^k)). \quad (2.8)$$

Nous considérons que seuls les espaces couleur hybrides de dimension d pour lesquels la mesure de corrélation est inférieure à un seuil donné sont les espaces candidats qui seront pris en compte par la procédure de construction de l'espace couleur hybride le plus discriminant. Grâce à cette étape de seuillage effectuée à chaque itération de rang d de la procédure de sélection, nous considérons seulement un nombre limité N_{cand}^d d'espaces couleur hybrides candidats de dimension d , notés S_l^d , avec $l = 1, \dots, N_{\text{cand}}^d$, parmi tous les $(N_{\text{comp}} - d + 1)$ espaces couleur hybrides HS_a^d possibles.

2.2.3 Construction de l'espace couleur hybride adapté

Comme il ne serait pas réaliste d'évaluer le pouvoir discriminant de toutes les combinaisons possibles de composantes couleur, nous employons un algorithme sous-optimal de sélection qui procède pas à pas pour déterminer l'espace couleur hybride adapté [FHPZ96]. Etant donné un ensemble d'attributs (ici, les composantes couleur) évalués sur une population d'individus (ici, les observations correspondant aux pixels représentatifs), une telle procédure consiste à sélectionner le meilleur attribut, puis l'associer à chacun des attributs restants et retenir le meilleur couple d'attributs, puis l'associer à chacun des attributs restants et ne conserver que le meilleur triplet d'attributs, jusqu'à atteindre la

dimension souhaitée. Dans ce cas, la procédure commence par un ensemble vide mais il est également possible de l'initialiser en partant de l'espace composé de tous les attributs, puis d'éliminer successivement les moins bons. D'autres stratégies plus sophistiquées alternent ajout et suppression d'attributs [Lez00, SPNP99]. Une comparaison des performances atteintes par différents algorithmes de sélection d'attributs montre que les résultats obtenus par différentes procédures séquentielles de recherche sont généralement similaires pour les cinq premiers attributs sélectionnés [PO01]. Par conséquent, nous avons retenu la procédure séquentielle de recherche la plus simple.

A chaque itération de rang d de la procédure de construction, nous considérons l'ensemble des N_{cand}^d espaces couleur hybrides candidats S_l^d de dimension d , $l = 1, \dots, N_{cand}^d$. Nous évaluons leurs pouvoirs discriminants avec un critère pertinent, et sélectionnons le meilleur qui est l'espace couleur hybride de dimension d le plus discriminant, noté S^d . Ce critère, noté $J_{dis}(S_l^d)$, mesure la dispersion des observations associées aux pixels représentatifs dans chaque espace couleur hybride candidat. Il permet de trouver un espace couleur hybride dans lequel les points-couleur représentant les pixels qui appartiennent à chaque région de l'image sont proches les uns des autres, mais aussi dans lequel les points-couleur représentant les pixels appartenant à des régions distinctes sont éloignés les uns des autres.

À la première itération ($d = 1$), nous considérons les espaces candidats mono-dimensionnels définis par chacune des N_{comp} composantes couleur disponibles. L'espace mono-dimensionnel sélectionné S^1 est celui qui maximise le pouvoir discriminant, à savoir le critère $J_{dis}(S_l^1)$. Lors de la seconde itération de la procédure ($d = 2$), les espaces couleur hybrides bi-dimensionnels sont constitués en associant la composante couleur de S^1 à chacune des $(N_{comp} - 1)$ composantes couleur restantes. Parmi ces N_{cand}^2 espaces couleur hybrides bi-dimensionnels S_l^2 , $l = 1, \dots, N_{cand}^2$, qui satisfont le critère basé sur la mesure de corrélation, nous sélectionnons celui qui maximise $J_{dis}(S_l^2)$ car il est le meilleur en termes de pouvoir discriminant entre les classes de pixels en présence. Cet espace sélectionné est l'espace couleur hybride le plus discriminant S^2 . La procédure est itérée jusqu'à ce qu'un critère d'arrêt, noté J_{stop}^d et détaillé ultérieurement, soit vérifié. Le rang $d = D$ de la dernière itération indique la dimension D de l'espace couleur hybride adapté S^D . La figure 2.1 illustre la procédure de construction de l'espace couleur hybride adapté.

Le pouvoir discriminant d'un espace couleur hybride candidat est évalué grâce à une mesure de compacité et de séparabilité des classes en présence. Nous utilisons la matrice de dispersion intra-classe $\mathbf{W}(S_l^d)$ pour évaluer la compacité de chaque classe $\mathcal{C}_j$ dans chaque espace couleur hybride candidat S_l^d , $l = 1, \dots, N_{cand}^d$:

$$\mathbf{W}(S_l^d) = \frac{1}{N_\omega \times N_{class}} \times \sum_{j=1}^{N_{class}} \sum_{i=1}^{N_\omega} (\mathbf{X}_{i,j} - \mathbf{M}_j)(\mathbf{X}_{i,j} - \mathbf{M}_j)^T, \quad (2.9)$$

où l'observation $\mathbf{X}_{i,j}$ est définie par l'équation (2.3) et où $\mathbf{M}_j = [m_j^1, \dots, m_j^k, \dots, m_j^d]^T$ est le vecteur des centres de gravité des N_ω observations $\mathbf{X}_{i,j}$ utilisées pour définir la classe $\mathcal{C}_j$, avec :

$$m_j^k = \frac{1}{N_\omega} \sum_{i=1}^{N_\omega} x_{i,j}^k. \quad (2.10)$$

Pour mesurer la séparabilité des classes, nous utilisons la matrice de dispersion inter-classe $\mathbf{B}(S_l^d)$:

$$\mathbf{B}(S_l^d) = \frac{1}{N_{class}} \times \sum_{j=1}^{N_{class}} (\mathbf{M}_j - \mathbf{M})(\mathbf{M}_j - \mathbf{M})^T, \quad (2.11)$$

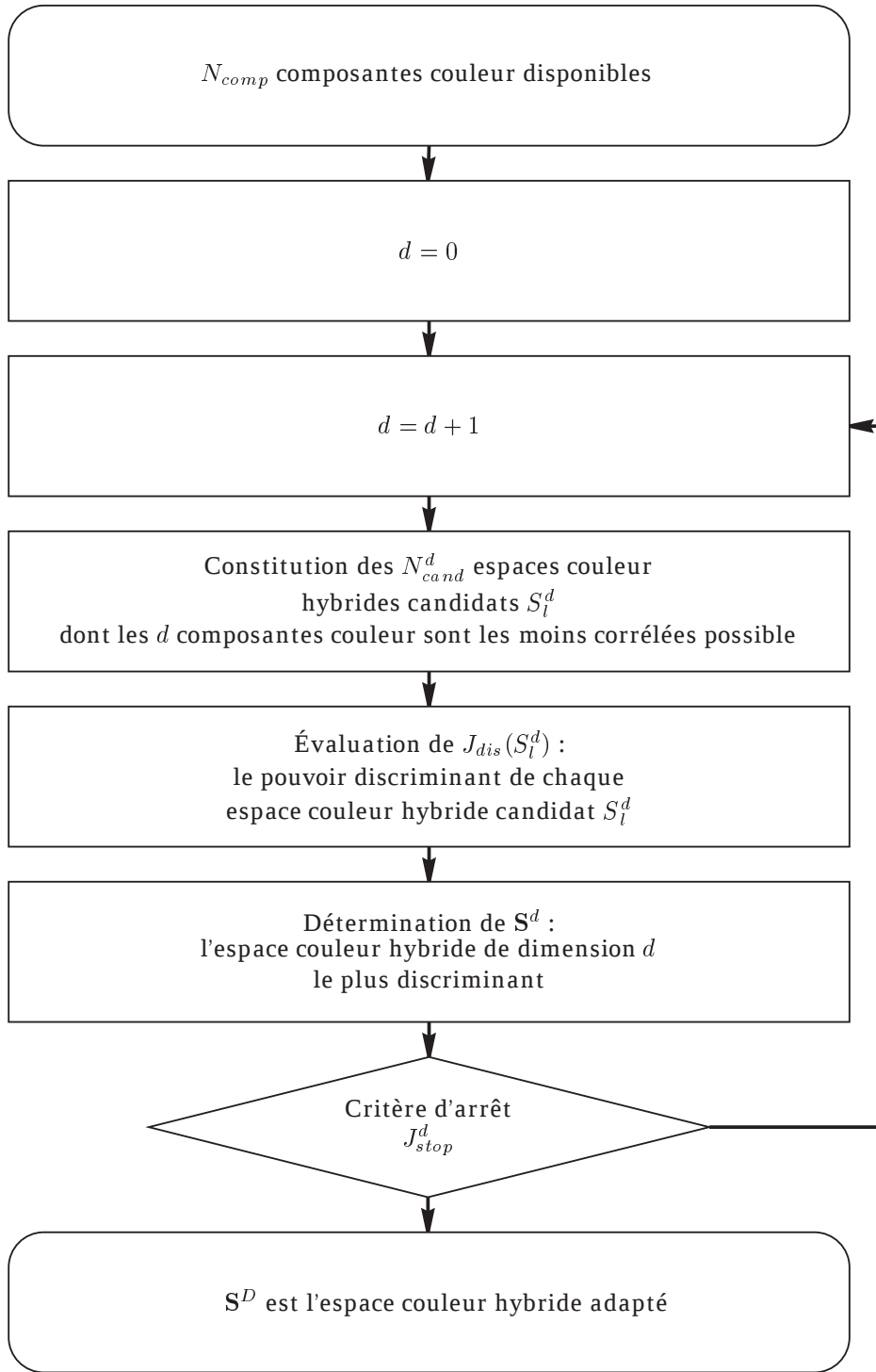


FIG. 2.1 – Procédure itérative de construction de l'espace couleur hybride adapté.

où $\mathbf{M} = [m^1, \dots, m^k, \dots, m^d]^T$ est le vecteur des centres de gravité des observations de toutes les classes, avec m^k défini par l'équation (2.6).

Parmi les nombreux critères proposés dans la littérature [HATB98] pour mesurer le pouvoir discriminant $J_{dis}(S_l^d)$ de chaque espace couleur hybride candidat S_l^d , nous avons retenu le critère de la trace. Dans ces conditions, $J_{dis}(S_l^d)$ peut être écrit comme :

$$J_{dis}(S_l^d) = \text{trace}\left(\left(\mathbf{W}(S_l^d) + \mathbf{B}(S_l^d)\right)^{-1} \mathbf{W}(S_l^d)\right) = \text{trace}\left(\left(\mathbf{T}(S_l^d)\right)^{-1} \mathbf{W}(S_l^d)\right), \quad (2.12)$$

où $\mathbf{T}(S_l^d) = \mathbf{W}(S_l^d) + \mathbf{B}(S_l^d)$ est la matrice de dispersion totale.

A chaque itération de rang d de la procédure de construction, l'espace couleur hybride le plus discriminant $\mathbf{S}^d$ est celui qui maximise le critère J_{dis} :

$$\mathbf{S}^d = \left\{ S_{l_0}^d \mid J_{dis}(S_{l_0}^d) = \max_{l=1}^{N_{cand}^d} (J_{dis}(S_l^d)) \right\}, \quad (2.13)$$

où l_0 indique le numéro de l'espace couleur hybride candidat sélectionné.

2.2.4 Critère d'arrêt

Le critère décrit par l'équation (2.12) permet de comparer les pouvoirs discriminants des différents ensembles de composantes couleur. Cependant, il n'indique pas si l'espace couleur hybride $\mathbf{S}^d$ qui est sélectionné à l'itération de rang d , est plus discriminant que l'espace couleur hybride $\mathbf{S}^{d-1}$ sélectionné lors de l'itération de rang $(d-1)$. Nous proposons d'exploiter un critère informationnel de discrimination afin d'arrêter la procédure de sélection quand la d^{eme} composante couleur de $\mathbf{S}^d$, sélectionnée à l'itération de rang d , n'améliore pas significativement la discrimination.

Ce critère mesure la dispersion des observations projetées sur la d^{eme} composante de $\mathbf{S}^d$, étant donnée la dispersion des observations dans l'espace le plus discriminant $\mathbf{S}^{d-1}$ de dimension $(d-1)$ [Rao73].

Rao définit respectivement les dispersions conditionnelles intra-classe et totale $\varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1})$ et $\varsigma_T(\mathbf{S}^d/\mathbf{S}^{d-1})$ évaluées en ajoutant la d^{eme} composante couleur de $\mathbf{S}^d$ aux $(d-1)$ composantes couleur de $\mathbf{S}^{d-1}$. Pour cela, il divise les d composantes de $\mathbf{S}^d$ en deux groupes, à savoir les $(d-1)$ composantes de $\mathbf{S}^{d-1}$ et la d^{eme} composante associée à $\mathbf{S}^d$. La matrice de dispersion intra-classe $\mathbf{W}(\mathbf{S}^d)$ est une matrice carrée de dimension $(d \times d)$ qui peut être écrite de la manière suivante :

$$\mathbf{W}(\mathbf{S}^d) = \frac{1}{N_{class} \times N_\omega} \times \begin{bmatrix} \sum_{j=1}^{N_{class}} \sum_{i=1}^{N_\omega} (x_{i,j}^1 - m_j^1)^2 & \cdots & \sum_{j=1}^{N_{class}} \sum_{i=1}^{N_\omega} (x_{i,j}^1 - m_j^1) \times (x_{i,j}^d - m_j^d) \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{N_{class}} \sum_{i=1}^{N_\omega} (x_{i,j}^1 - m_j^1) \times (x_{i,j}^d - m_j^d) & \cdots & \sum_{j=1}^{N_{class}} \sum_{i=1}^{N_\omega} (x_{i,j}^d - m_j^d)^2 \end{bmatrix}. \quad (2.14)$$

Les éléments de la dernière ligne et de la dernière colonne de cette matrice sont associés à la d^{eme} composante couleur sélectionnée. La matrice $\mathbf{W}(\mathbf{S}^d)$ peut donc être écrite comme :

$$\mathbf{W}(\mathbf{S}^d) = \begin{bmatrix} \mathbf{W}(\mathbf{S}^{d-1}) & \mathbf{W}_{d-1,1}(\mathbf{S}^d) \\ \mathbf{W}_{1,d-1}(\mathbf{S}^d) & W_{1,1}(\mathbf{S}^d) \end{bmatrix}, \quad (2.15)$$

où $\mathbf{W}(\mathbf{S}^{d-1})$ est la matrice de dispersion intra-classe, carrée de dimension $(d-1) \times (d-1)$ et évaluée à partir des observations dans l'espace $\mathbf{S}^{d-1}$. Les deux vecteurs $\mathbf{W}_{1,d-1}(\mathbf{S}^d)$ et $\mathbf{W}_{d-1,1}(\mathbf{S}^d)$, de dimension $(d-1)$, ainsi que le scalaire $W_{1,1}(\mathbf{S}^d)$ fournissent des mesures de la contribution de la d^{eme} composante couleur en termes de pouvoir discriminant. Une décomposition similaire peut être appliquée à la matrice de dispersion totale $\mathbf{T}(\mathbf{S}^d)$.

Grâce à cette décomposition, la dispersion conditionnelle intra-classe $\varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1})$ est définie par :

$$\varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1}) = W_{1,1}(\mathbf{S}^d) - \mathbf{W}_{1,d-1}(\mathbf{S}^d) (\mathbf{W}(\mathbf{S}^{d-1}))^{-1} \mathbf{W}_{d-1,1}(\mathbf{S}^d). \quad (2.16)$$

De manière similaire, la dispersion conditionnelle totale $\varsigma_T(\mathbf{S}^d/\mathbf{S}^{d-1})$ s'écrit comme :

$$\varsigma_T(\mathbf{S}^d/\mathbf{S}^{d-1}) = T_{1,1}(\mathbf{S}^d) - \mathbf{T}_{1,d-1}(\mathbf{S}^d)(\mathbf{T}(\mathbf{S}^{d-1}))^{-1}\mathbf{T}_{d-1,1}(\mathbf{S}^d). \quad (2.17)$$

Comme la différence entre les dimensions de $\mathbf{S}^d$ et $\mathbf{S}^{d-1}$ est égale à un, les dispersions conditionnelles intra-classe et totale sont des scalaires. Le critère d'arrêt qui a été proposé par [Rom73] est la différence relative entre ces dispersions conditionnelles $\varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1})$ et $\varsigma_T(\mathbf{S}^d/\mathbf{S}^{d-1})$:

$$J_{stop}^d = \frac{\varsigma_T(\mathbf{S}^d/\mathbf{S}^{d-1}) - \varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1})}{\varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1})}. \quad (2.18)$$

Avec ce critère, il est possible de quantifier la contribution de l'espace couleur hybride $\mathbf{S}^d$ sélectionné au rang d par rapport à l'espace $\mathbf{S}^{d-1}$ précédemment sélectionné. Tant que la valeur du critère J_{stop}^d augmente, l'espace couleur hybride le plus discriminant devrait améliorer la discrimination entre les classes de pixels en présence. L'évolution des valeurs de J_{stop}^d est analysée afin de détecter toute phase de décroissance qui indique que le nouvel espace couleur hybride le plus discriminant $\mathbf{S}^d$ devrait conduire à de plus faibles performances en termes de discrimination entre les classes de pixels en présence que l'espace $\mathbf{S}^{d-1}$ précédemment sélectionné. Par conséquent, la procédure est arrêtée quand la valeur du critère diminue. Nous notons D le rang de l'itération qui précède la première phase de décroissance du critère d'arrêt. Cette stratégie fournit la dimension D de l'espace couleur hybride adapté $\mathbf{S}^D$.

Afin d'évaluer les performances obtenues par notre approche, nous proposons d'analyser des images couleur de foot-ball. L'objectif de la procédure est d'extraire les joueurs et d'identifier leurs équipes par l'analyse des couleurs de leurs tenues qui peuvent être hétérogènes. Cette application a été choisie afin de montrer comment des régions dont les couleurs sont hétérogènes dans l'espace couleur (R, G, B) peuvent être représentées par des sous-ensembles de pixels partageant des propriétés similaires dans l'espace couleur hybride adapté.

2.3 Application à la segmentation d'images de foot-ball

Dans le contexte du suivi de joueurs de foot-ball par analyse de séquences d'images, de nombreux auteurs ont tenté d'isoler des joueurs de foot-ball dans les différentes images constituant ces séquences. Ils appliquent notamment une technique de multi-seuillage basée sur l'analyse des histogrammes des composantes tri-chromatiques [KYA94, SCKH97, NFM⁺00] ou sur la différence entre l'image courante et une image de référence qui ne représente que le terrain [BB98, SO99]. Bien que ces techniques de segmentation fournissent de bons résultats, elles ont tendance à échouer en cas d'occultation des joueurs ou quand les joueurs sont très proches les uns des autres dans l'image.

Nous proposons d'appliquer notre méthode de classification de pixels dans des espaces couleur hybrides afin d'identifier les équipes auxquelles appartiennent les joueurs extraits de séquences d'images couleur acquises par une caméra fixe. Cette approche a été développée pour être intégrée à un système automatique de suivi des joueurs [VtPV04].

Examinons les quatre images de la figure 2.2 provenant d'une séquence à analyser. L'objectif de l'analyse de ces images est d'extraire les joueurs et l'arbitre afin de les suivre durant le match. Dans ce contexte, les *pixels-joueurs* représentant les joueurs de chacune des deux équipes et l'arbitre sont à classer en trois classes $\mathcal{C}_j$, avec $j = 1, \dots, 3$.

FIG. 2.2 – Images de foot-ball (125×125 pixels).

Les images sont pré-traitées afin d'éliminer le terrain grâce à une procédure qui extrait les pixels-joueurs (voir figure 2.3). Cette procédure, basée sur la technique de multi-seuillage de Ohlander [OPR78], est détaillée dans [Van00].



FIG. 2.3 – Pixels-joueurs extraits des images de la figure 2.2.

L'algorithme de reconnaissance des équipes des joueurs de foot-ball est divisé en deux étapes qui sont décrites dans les deux premières parties de ce chapitre. D'abord, une procédure d'apprentissage supervisé "off-line" construit l'espace couleur hybride le plus discriminant à partir d'un échantillon d'apprentissage qui a été sélectionné interactivement.

Ensuite, une procédure de classification de pixels-joueurs "on-line" est appliquée afin de segmenter en temps réel les images successives dont la couleur est représentée dans l'espace couleur hybride construit à l'étape précédente. Pour cela, chaque pixel-joueur est assigné à la classe dont le centre est le plus proche du point-couleur qui le représente dans l'espace couleur hybride le plus discriminant. Dans la dernière partie de ce paragraphe, nous montrons l'efficacité de notre approche avec des résultats expérimentaux.

2.3.1 Apprentissage supervisé

Afin de déterminer le meilleur espace couleur hybride pour la classification des pixels-joueurs, il est nécessaire de construire un échantillon d'apprentissage constitué de N_ω pixels-joueurs $\omega_{i,j}$ représentatifs de chacune des N_{class} classes $\mathcal{C}_j$ de pixels-joueurs, avec $i = 1, \dots, N_\omega$ et $j = 1, \dots, N_{class}$.

Nous considérons les images d'apprentissage issues des différentes images constituant la séquence à analyser. Ces images d'apprentissage sont pré-traitées afin d'en éliminer le terrain (voir figure 2.3). Des *fenêtres-joueurs* sont sélectionnées interactivement dans chacune de ces images d'apprentissage. Chacune de ces fenêtres-joueurs ne contient que les pixels-joueurs d'un seul joueur (voir figure 2.4). Ces fenêtres-joueurs sont extraites de séries d'images où les joueurs effectuent différents types d'actions (courir, sauter, shooter, ...) et dans différentes positions par rapport à la caméra (de dos, de face, de

côté, ...).

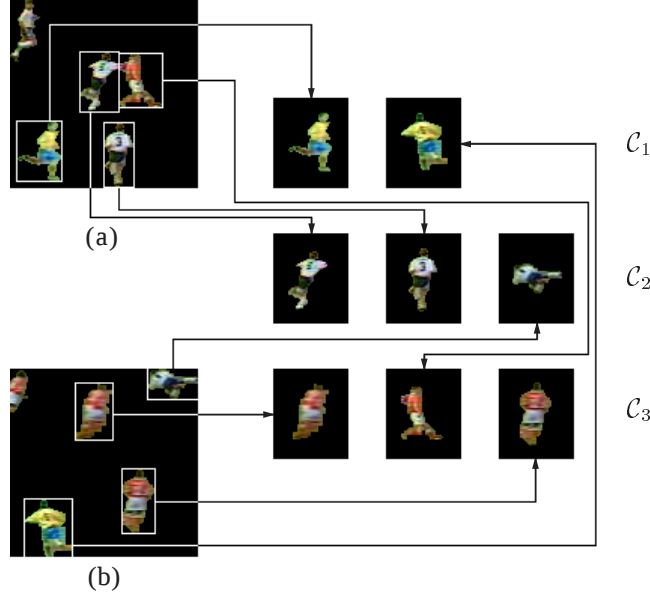


FIG. 2.4 – Fenêtres-joueurs interactivement extraites des images d'apprentissage.

Nous extrayons le même nombre N_W de fenêtres-joueurs pour chaque classe (voir figure 2.5). Afin de constituer l'ensemble des N_ω pixels-joueurs représentatifs de chaque classe C_j , $j = 1, \dots, N_{class}$, nous sélectionnons aléatoirement $\frac{N_\omega}{N_W}$ pixels-joueurs dans chacune des N_W fenêtres-joueurs (voir figure 2.6).

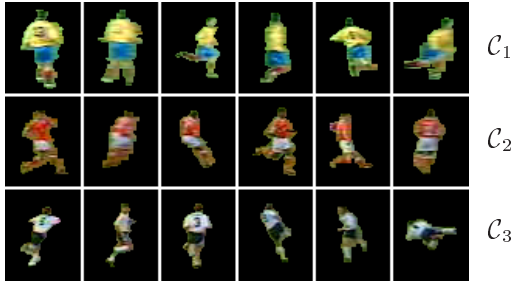


FIG. 2.5 – Fenêtres joueurs extraites.

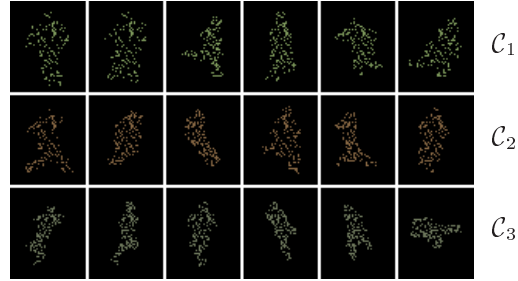


FIG. 2.6 – Pixels-joueurs représentatifs sélectionnés.

Afin d'analyser un espace couleur hybride candidat de dimension d par la procédure décrite au paragraphe précédent, nous associons une observation $\mathbf{X}_{i,j} = [x_{i,j}^1, \dots, x_{i,j}^k, \dots, x_{i,j}^d]^T$ à chaque pixel-joueur $\omega_{i,j}$ représentatif de la classe C_j .

Si nous considérons seulement les niveaux des d composantes couleur du pixel-joueur représentatif $\omega_{i,j}$, les observations correspondant à un même joueur risquent d'être sensibles aux variations, même insignifiantes, de la couleur de sa tenue. Afin de réduire les effets de ces variations, la k^{eme} coordonnée de l'observation $\mathbf{X}_{i,j}$ est la moyenne des niveaux de la k^{eme} composante couleur des pixels-joueurs appartenant à un certain voisinage de $\omega_{i,j}$.

Comme le champ d'observation de la caméra utilisée lors de nos expériences couvre une zone limitée du terrain de foot-ball, nous considérons que la taille des joueurs dans les images est approximativement constante. Dans ces conditions, les tailles verticale et horizontale du voisinage considéré de chaque pixel-joueur dépendent respectivement des hauteurs et largeurs moyennes des joueurs dans

les images [VtP98b].

La sélection de l'espace couleur hybride le plus discriminant est basée sur l'analyse de cet ensemble d'apprentissage, qui est constitué des $(N_\omega \times N_{class})$ observations $X_{i,j}$ associées aux pixels-joueurs représentatifs affichés sur la figure 2.6. Pour cet exemple, l'espace couleur hybride le plus discriminant S^D est le plan ($D = 2$) défini par la composante couleur a^* de l'espace couleur (L^*, a^*, b^*) et par la composante couleur C_{UV} de l'espace couleur (Y', U', V') .

La figure 2.7 montre l'évolution de la valeur du critère d'arrêt utilisé par la procédure séquentielle de construction de l'espace couleur hybride adapté. La valeur de ce critère diminue après la seconde itération, indiquant ainsi que l'espace couleur hybride de dimension deux est le plus discriminant.

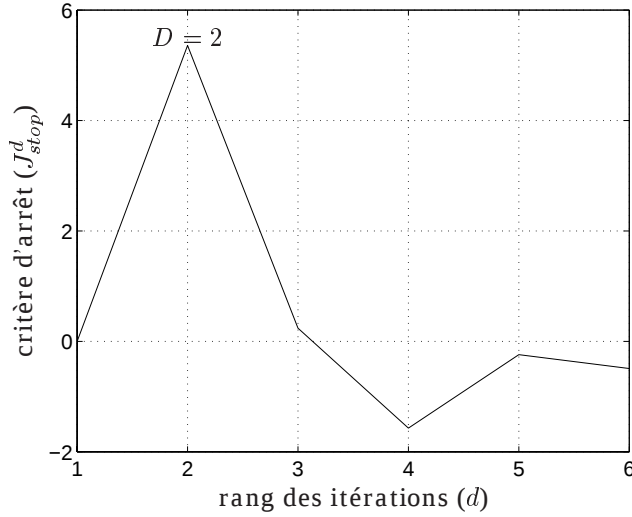


FIG. 2.7 – Évolution de la valeur du critère d'arrêt utilisé pour construire l'espace couleur hybride adapté à partir des pixels-joueurs représentatifs de la figure 2.6.

2.3.2 Classification des pixels-joueurs

Une fois que la procédure "off-line" d'apprentissage supervisé a fourni l'espace couleur hybride le plus discriminant, la segmentation d'une séquence d'images couleur peut être exécutée en temps réel par la procédure de classification. Afin de classer les pixels-joueurs extraits par la phase de pré-traitement des images, nous calculons un vecteur d'attributs $\mathbf{X}_P = [x_P^1, \dots, x_P^D]^T$ pour chaque pixel-joueur P dans l'espace couleur hybride S^D précédemment déterminé. Comme indiqué précédemment, pour réduire les variations non significatives de la couleur, les coordonnées du vecteur $\mathbf{X}_P$ sont les moyennes des niveaux des composantes couleur des pixels-joueurs appartenant à un certain voisinage du pixel-joueur P (voir figure 2.8).

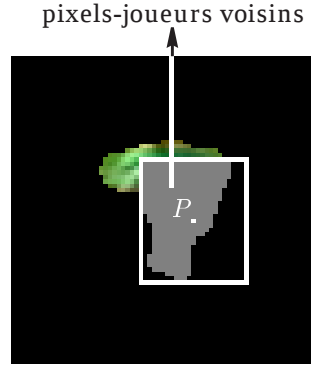


FIG. 2.8 – Voisinage d'un pixel-joueur.

Le pixel-joueur P est assigné à la classe dont le centre est le plus proche du point-couleur le représentant dans l'espace couleur hybride le plus discriminant. La distance euclidienne D_j^P entre $\mathbf{X}_P$ et le centre de chaque classe $\mathcal{C}_j$ défini par le vecteur moyenne $\mathbf{M}_j = [m_j^1, \dots, m_j^D]^T$ dans l'espace couleur hybride $\mathbf{S}^D$, est évaluée par :

$$D_j^P = \|\mathbf{X}_P - \mathbf{M}_j\| = \sqrt{\sum_{d=1}^D (x_P^d - m_j^d)^2}. \quad (2.19)$$

Le pixel-joueur est assigné à la classe $\mathcal{C}_l$ associée à la distance la plus faible :

$$\{\mathcal{C}_l \mid D_l^P = \min_{j=1}^{N_{classes}} (D_j^P)\}. \quad (2.20)$$

Comme notre objectif se concentre sur la comparaison des résultats obtenus par une même procédure de classification des observations dans différents espaces couleur, nous n'avons pas utilisé des procédures de classification automatique plus sophistiquées, telles que par exemple, des procédures basées sur des réseaux de neurones [LR97].

2.3.3 Résultats

Afin de comparer la classification des pixels-joueurs effectuée dans l'espace couleur hybride le plus discriminant à la classification dans des espaces couleur classiques, nous nous proposons d'examiner les deux exemples suivants.

Exemple 1

La figure 2.9 affiche quatre images couleur extraites d'une séquence contenant deux joueurs opposés et un arbitre.

Les images de la figure 2.9 constituent l'échantillon test extrait de la même séquence que les images de la figure 2.2 qui ont été utilisées pour construire l'échantillon d'apprentissage. Ces images différentes de celles de la figure 2.2 ne sont pas très simples à segmenter car chacune d'elles contient aux moins deux joueurs très proches les uns des autres.

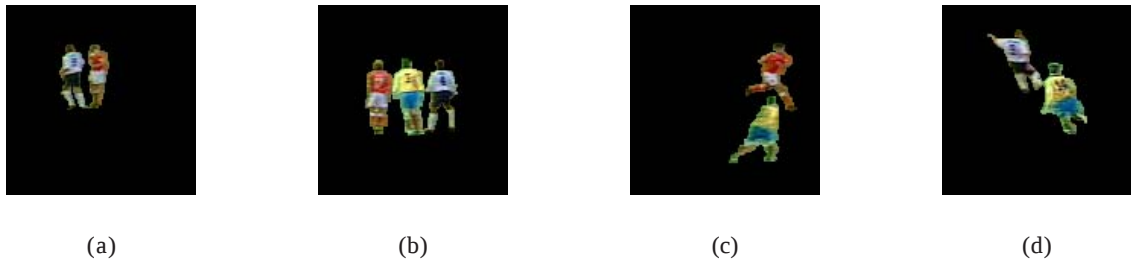
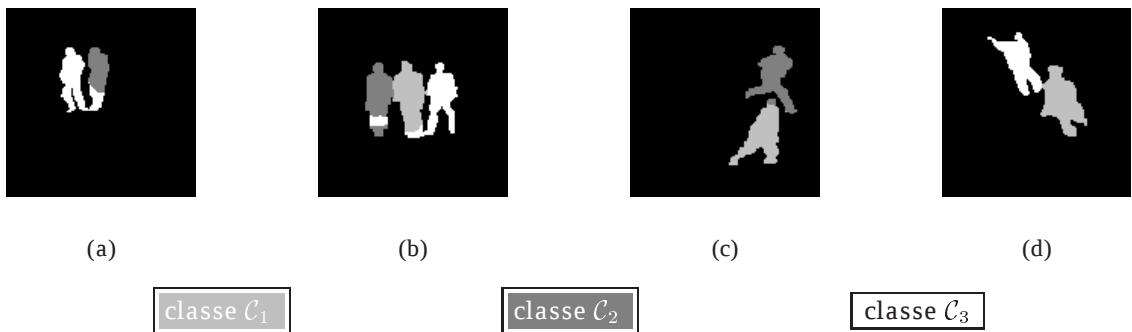
FIG. 2.9 – Images couleur de foot-ball de l'exemple 1 (125×125 pixels).

FIG. 2.10 – Pixels-joueurs extraits des images de la figure 2.9.

Les images de la figure 2.10 contiennent les pixels-joueurs qui sont extraits des images de la figure 2.9 par le pré-traitement basé sur la procédure de multi-seuillage de Ohlander.

Les images de la figure 2.11 montrent comment les pixels-joueurs sont classés dans l'espace couleur hybride le plus discriminant (a^* , C_{UV}) précédemment déterminé. Les étiquettes des pixels-joueurs sont spécifiques aux classes auxquelles ils sont assignés. Dans les images de la figure 2.11, les pixels-joueurs connexes avec les mêmes étiquettes forment des régions qui correspondent effectivement aux joueurs.

FIG. 2.11 – Pixels-joueurs de la figure 2.10 classés dans l'espace couleur hybride (a^* , C_{UV}).

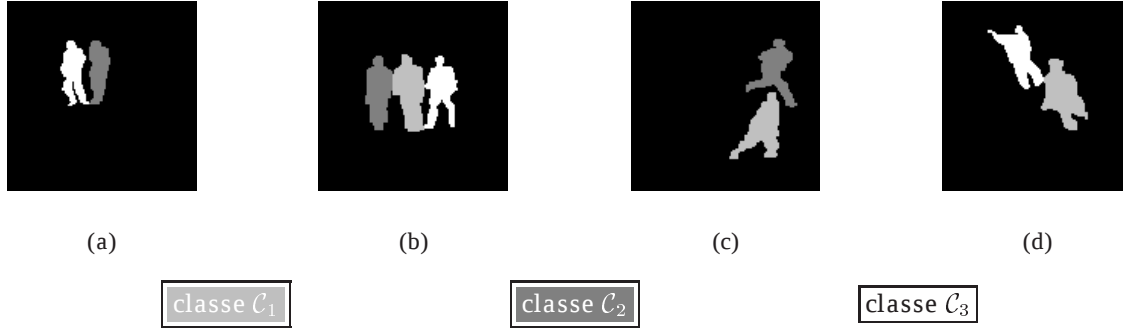


FIG. 2.12 – Vérité-terrain : images des pixels-joueurs des images de la figure 2.10 classés interactive-ment.

Nous constatons que dans les images (a) et (b) de la figure 2.11, les seuls pixels-joueurs mal classés sont ceux qui représentent les jambes ou les chaussettes des joueurs. Comme ces pixels-joueurs forment de petites régions, ils peuvent être facilement éliminés par un post-traitement.

Afin d'illustrer les propriétés discriminantes de l'espace couleur hybride sélectionné, nous comparons les résultats de la segmentation des images de la figure 2.11 avec ceux obtenus par une classification dans les vingt-deux espaces couleur classiques suivants :

- les quatre espaces de primaires : (R, G, B) , (r, g, b) , (X, Y, Z) , (x, y, z) ,
- l'espace d'axes indépendants : $(I1, I2, I3)$,
- les huit espaces luminance-chrominance : (L^*, u^*, v^*) , (L^*, a^*, b^*) , (bw, rg, by) , $(A, C1, C2)$, (Y', U', V') , (Y', I', Q') , (Y, Ch_1, Ch_2) , (I, r, g) ,
- et les neuf espaces perceptuels incluant les huit espaces de coordonnées polaires (L, C, H) déduits des huit espaces luminance-chrominance et l'espace de coordonnées perceptuelles (I, S, T) .

Pour cela, nous proposons de mesurer le taux d'erreur de classification, noté ε , obtenu par comparaison entre les images segmentées automatiquement et les images segmentées interactivement qui constituent la "vérité-terrain" (voir figure 2.12). Bien que l'assignation interactive des pixels-joueurs aux différentes classes soit fastidieuse et nécessite une attention très soutenue de la part de l'opérateur, il n'existe aucune autre solution pour estimer la qualité des résultats.

Les espaces (r, g, b) , (I, S, T) et (Y', C_{UV}, h_{UV}) sont les trois espaces couleur classiques pour lesquels les taux d'erreur sont les plus faibles. Le tableau 2.1 indique ces taux d'erreur de classification pour toutes les images de la figure 2.10. Ils sont à comparer avec le taux d'erreur obtenu avec l'espace (a^*, C_{UV}) qui est l'espace couleur hybride le plus discriminant. La classification des pixels-joueurs effectuée dans cet espace couleur hybride fournit les taux d'erreur les plus faibles pour les images 2.10(b), 2.10(c) et 2.10(d). Pour l'image 2.10(a), le taux d'erreur de classification dans l'espace couleur hybride le plus discriminant reste proche du taux le plus faible obtenu avec l'espace (I, S, T) . Par ailleurs, le tableau 2.1 montre que les taux d'erreur obtenus avec l'espace (R, G, B) sont bien plus élevés que ceux obtenus en utilisant l'espace couleur hybride le plus discriminant.

Enfin, comme le taux moyen d'erreur de classification obtenu avec l'espace couleur hybride le plus discriminant est le plus faible de tous ceux présents dans le tableau 2.1, nous concluons que l'espace couleur hybride permet d'obtenir des meilleurs résultats de classification que les trois espaces couleur classiques les mieux adaptés à cet exemple.

Espace couleur	Taux d'erreur de classification ε				Taux d'erreur moyen $\bar{\varepsilon}$
	image (a)	image (b)	image (c)	image (d)	
(r, g, b)	32,63	10,01	5,19	0,65	10,24
(I, S, T)	8,63	27,77	11,63	1,81	14,39
(Y', C_{UV}, h_{UV})	12,28	17,15	23,49	6,14	15,08
(a^*, C_{UV})	11,5	7,35	1,86	0,29	4,98
(R, G, B)	24,12	29,25	44,57	48,23	36,73

TAB. 2.1 – Taux d'erreur de classification des pixels-joueurs des images de la figure 2.10.

Exemple 2

Afin de mettre en évidence les performances atteintes par notre méthode dans des situations non triviales, nous nous proposons de segmenter six images couleur extraites d'une séquence avec deux joueurs adverses, un gardien de but et l'arbitre (voir figure 2.13). Ce second exemple a été choisi pour donner un aperçu du comportement de notre méthode de segmentation dans des situations complexes, par exemple des occultations de joueurs dans les images 2.13(c) et 2.13(d) et des joueurs en contact dans les images 2.13(a), 2.13(b), 2.13(e) et 2.13(f).

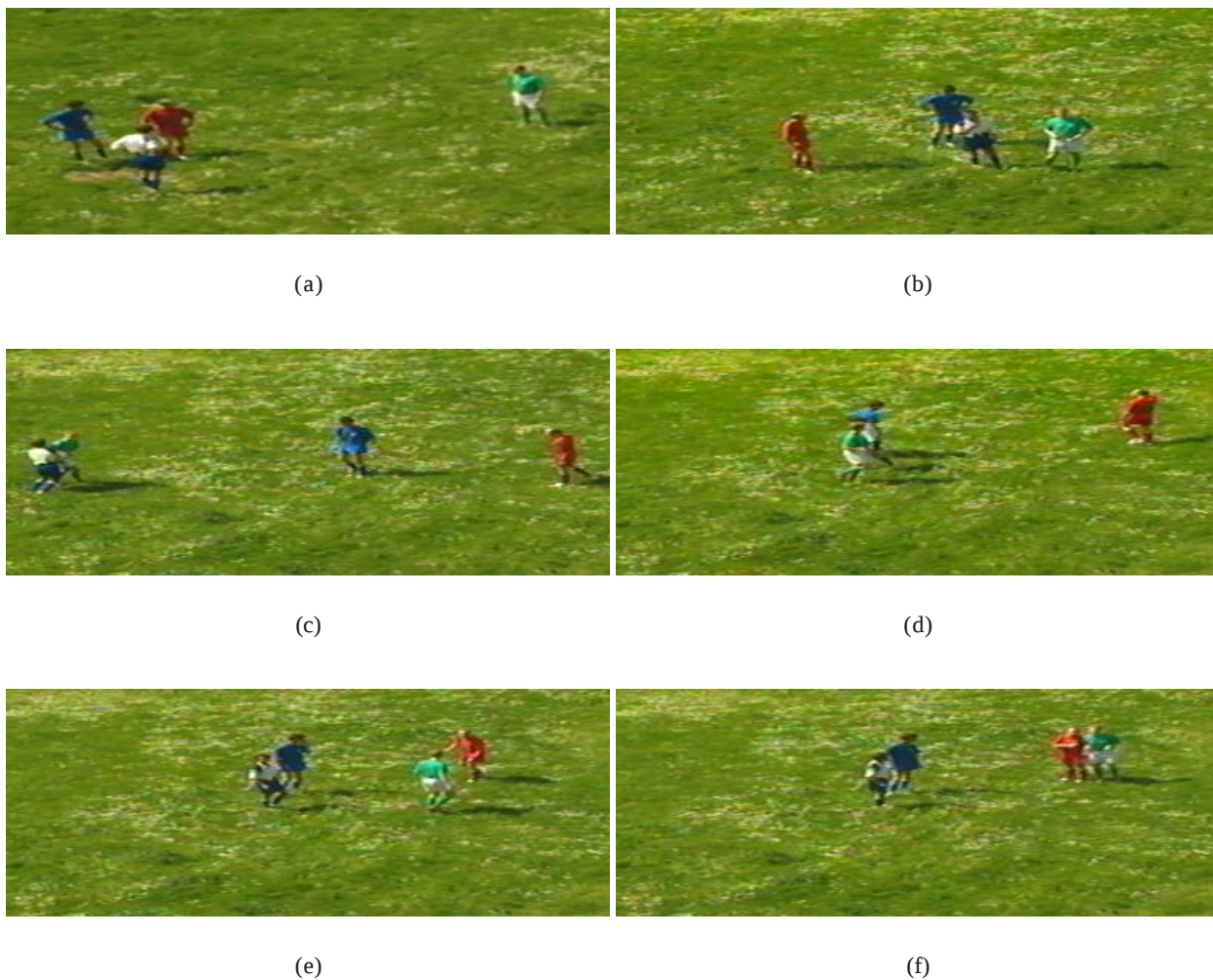


FIG. 2.13 – Images couleur de foot-ball de l'exemple 2 (400×150 pixels).

Les images de la figure 2.14 montrent les pixels-joueurs qui sont extraits des images de la figure 2.13 par le pré-traitement basé sur le multi-seuillage de Ohlander.

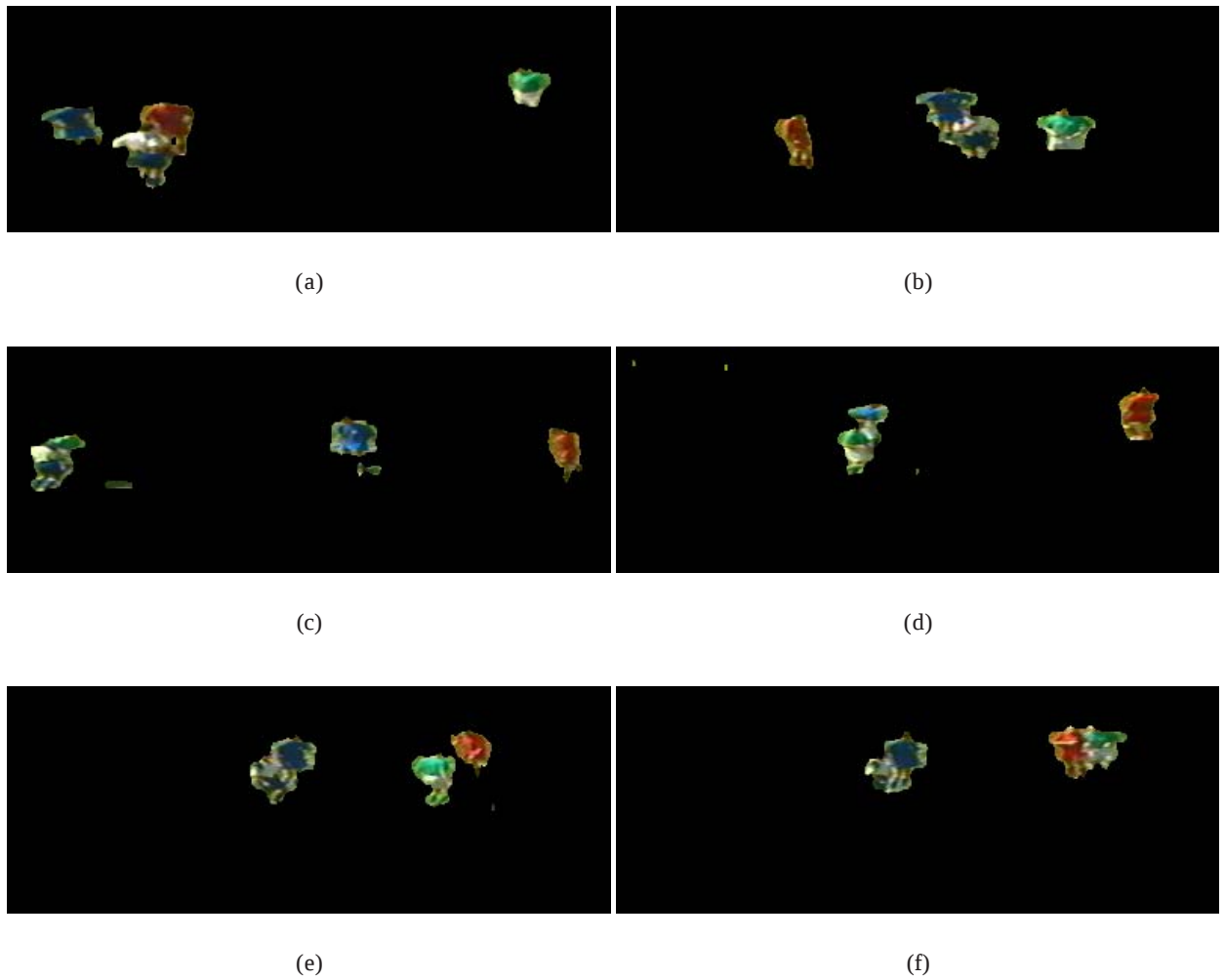


FIG. 2.14 – Pixels-joueurs extraits des images de la figure 2.13.

Durant l'étape d'apprentissage supervisé, nous avons sélectionné l'ensemble des différentes fenêtres-joueurs présentées sur la figure 2.15. Ces fenêtres-joueurs sont sélectionnées dans des images du même match qui diffèrent de celles constituant l'échantillon test de la figure 2.13. La figure 2.16 contient l'ensemble des pixels-joueurs représentatifs des quatre classes correspondant aux joueurs des deux équipes, au gardien de but et à l'arbitre.

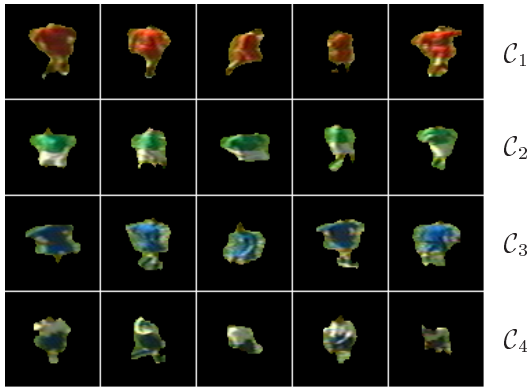


FIG. 2.15 – Fenêtres-joueurs sélectionnées.

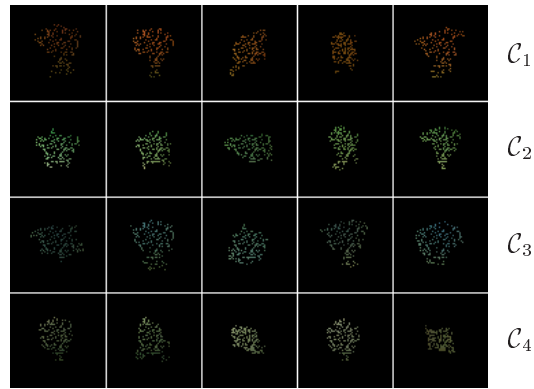


FIG. 2.16 – Pixels-joueurs représentatifs.

L'espace couleur hybride le plus discriminant résultant de l'analyse de ces pixels-joueurs représentatifs est le plan défini par la coordonnée normalisée x du système de représentation (x, y, z) de la CIE et par la composante couleur Ch_2 de l'espace couleur (Y, Ch_1, Ch_2) de Carron [LC99].

Les images de la figure 2.17 montrent comment sont classés les pixels-joueurs des images de la figure 2.14 dans cet espace couleur hybride. Nous constatons que trois joueurs sont identifiés sans ambiguïté dans l'image 2.17(d) alors qu'ils sont en contact dans l'image 2.13(d) correspondante et qu'il est donc très difficile de les distinguer visuellement. Par ailleurs, les joueurs qui sont proches les uns des autres dans les images sont effectivement représentés par des régions distinctes (voir images 2.17(a), 2.17(b), 2.17(c) et 2.17(e)). Une seule erreur importante apparaît : la région blanche comprise entre les régions rouge et verte ne correspond à aucun joueur de foot-ball dans l'image 2.17(f).

Comme pour le premier exemple, nous comparons les résultats de classification des pixels-joueurs effectuée dans l'espace couleur hybride le plus discriminant à ceux obtenus avec différents espaces couleur classiques. Pour constituer la "vérité-terrain" de ce problème de segmentation, les pixels-joueurs sont interactivement assignés aux différentes classes (voir figure 2.18).

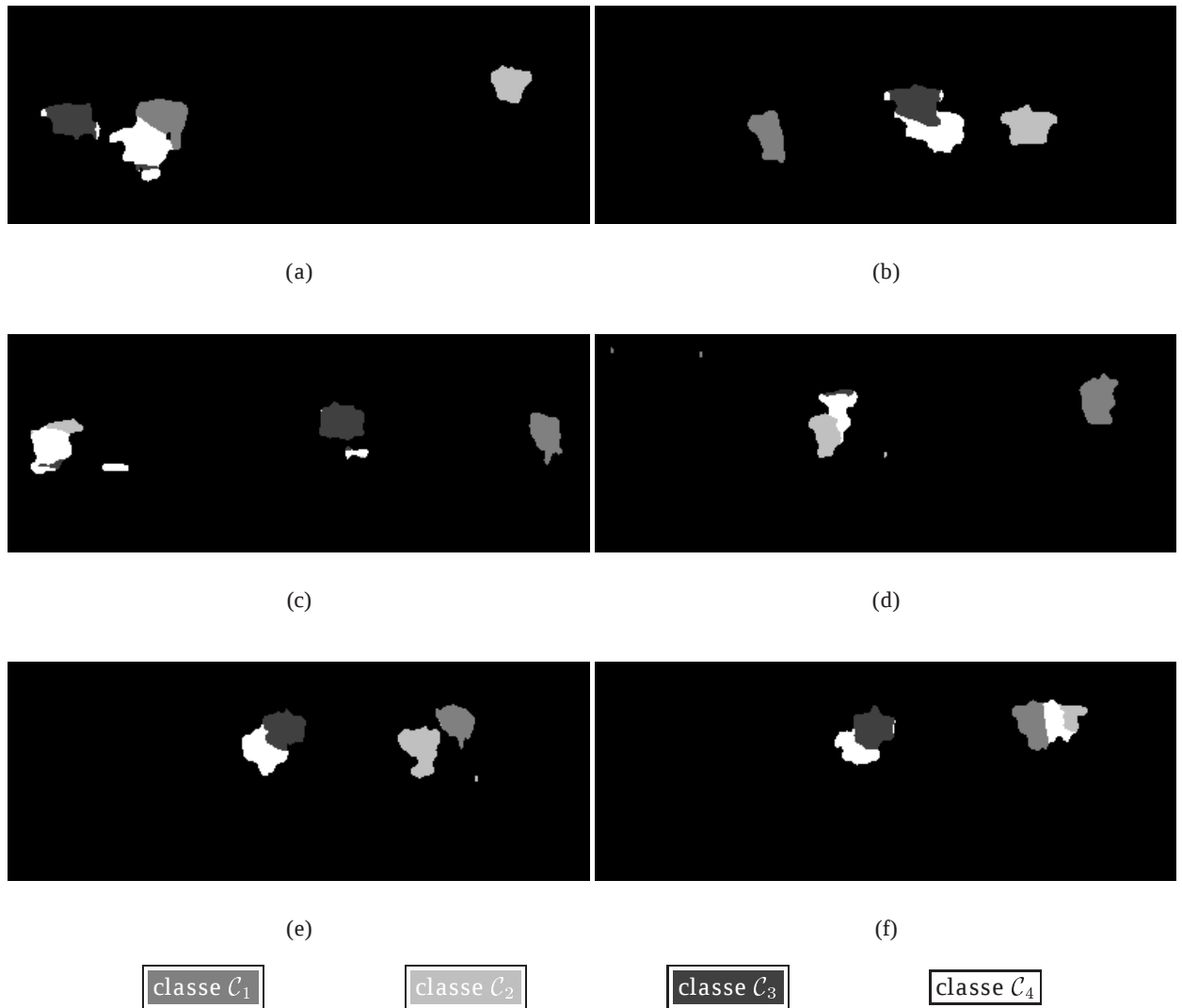


FIG. 2.17 – Pixels-joueurs de la figure 2.14 classés dans l'espace couleur hybride le plus discriminant (x, Ch_2) .

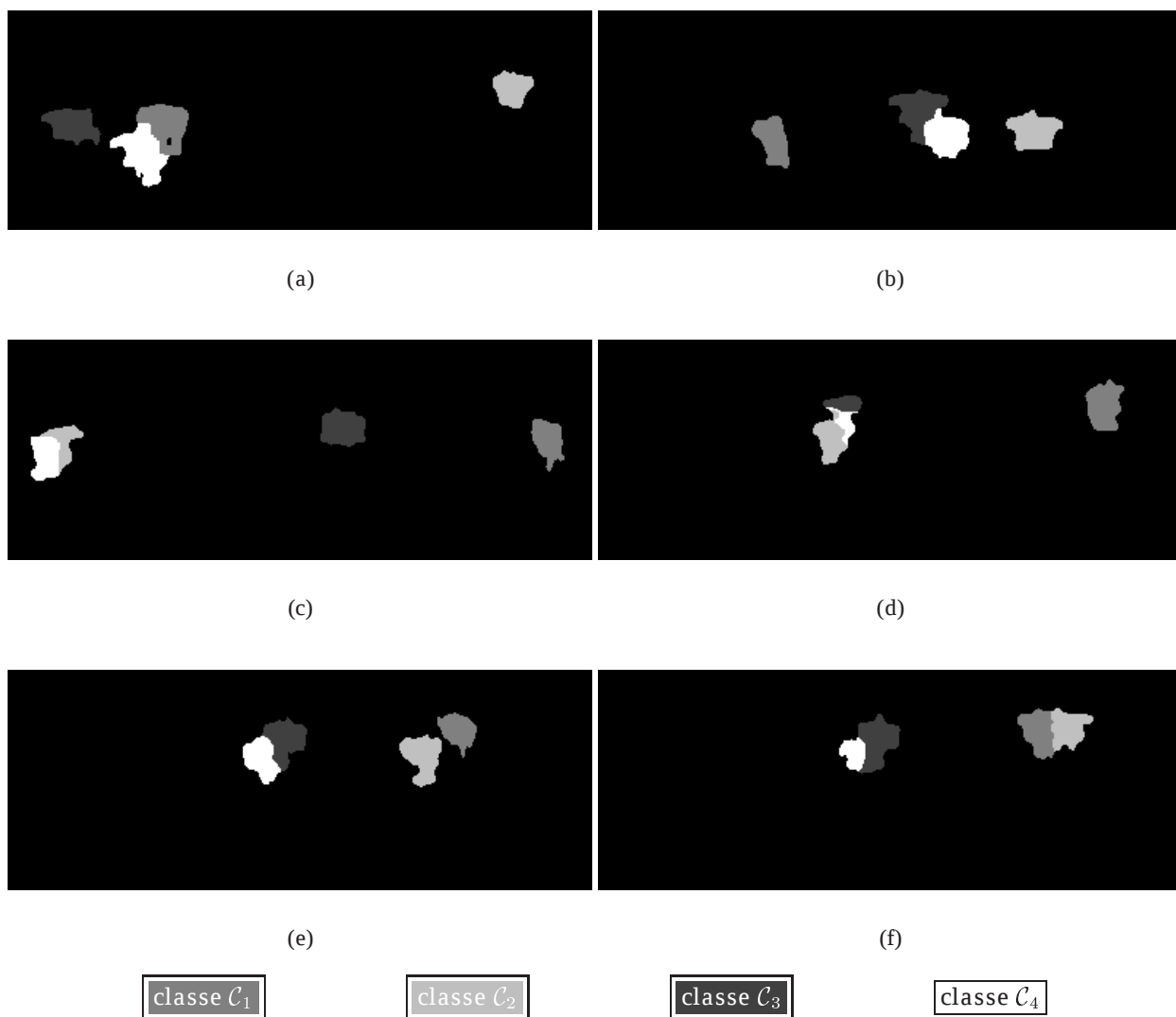


FIG. 2.18 – Vérité-terrain : images des pixels-joueurs des images de la figure 2.14 classés interactivement.

Pour toutes les images de la figure 2.14, la classification est effectuée dans tous les espaces couleur classiques utilisés pour le premier exemple (voir paragraphe 2.3.3 de la page 90). Les espaces (L^*, C_{ab}^*, h_{ab}) , (L^*, C_{uv}^*, h_{uv}) et (L^*, a^*, b^*) sont les trois espaces couleur classiques qui fournissent les taux d'erreur les plus faibles. Les taux d'erreur de classification associés à ces espaces pour toutes les images de la figure 2.14 sont présentés dans le tableau 2.2. Pour toutes les images de la figure 2.14, l'espace couleur le plus discriminant fournit de meilleurs résultats de classification que l'espace (R, G, B) .

Les taux d'erreur obtenus par la classification dans ces trois espaces couleur classiques sont maintenant comparés avec ceux obtenus par la classification dans l'espace couleur hybride le plus discriminant (x, Ch_2) . Les taux d'erreur de classification dans l'espace couleur hybride le plus discriminant sont plus élevés pour les images 2.14(b) et 2.14(f), similaires pour l'image 2.14(e), et plus faibles pour les images 2.14(a), 2.14(c) et 2.14(d) que ceux associés aux trois espaces couleur classiques les mieux adaptés. Ainsi, l'utilisation de l'espace couleur hybride le plus discriminant améliore les résultats de classification pour trois images parmi les six de la figure 2.14. Examinons de plus près les images 2.14(b), 2.14(e) et 2.14(f), pour lesquelles l'espace couleur hybride ne fournit pas les meilleurs taux d'erreur. Ces trois images sont si difficiles à segmenter qu'il n'existe pas un unique espace couleur classique adapté à ces images. En effet, l'espace (L^*, C_{uv}^*, h_{uv}) est le meilleur pour les images 2.14(b) et 2.14(f), tandis que l'espace (L^*, C_{ab}^*, h_{ab}) fournit les meilleurs résultats pour l'image 2.14(e).

Espace couleur	Taux d'erreur de classification ε						Taux d'erreur moyen $\bar{\varepsilon}$
	image (a)	image (b)	image (c)	image (d)	image (e)	image (f)	
$(L^*, C_{ab}^*, h_{ab}^*)$	7, 44	5, 19	16, 82	22, 14	4, 58	12, 00	11, 36
$(L^*, C_{uv}^*, h_{uv}^*)$	12, 59	4, 65	16, 92	24, 53	4, 71	7, 97	11, 89
(L^*, a^*, b^*)	8, 41	9, 43	13, 30	16, 72	4, 98	21, 79	12, 44
(x, Ch_2)	7, 09	11, 13	11, 40	14, 01	4, 76	25, 07	12, 24
(R, G, B)	23, 37	16, 78	32, 03	23, 82	10, 69	33, 60	23, 38

TAB. 2.2 – Taux d'erreur de classification des pixels-joueurs des images de la figure 2.14.

Dimension de l'espace couleur hybride le plus discriminant

Pour ces deux exemples, les deux espaces couleur les plus discriminants qui ont été sélectionnés par la procédure sont des plans. Ces deux exemples montrent donc que la dimension de l'espace couleur hybride le plus discriminant n'est pas toujours égale à trois, comme pour les espaces couleur classiques.

Comme précédemment indiqué, la procédure séquentielle de construction de l'espace couleur hybride le plus discriminant est arrêtée quand la valeur du critère d'arrêt commence à diminuer. Cette stratégie fournit la dimension D de l'espace couleur hybride sélectionné. Afin d'évaluer la pertinence de cette dimension, les images des deux exemples sont segmentées en utilisant les six premiers espaces les plus discriminants sélectionnés par la procédure séquentielle de construction.

Les figures 2.19 et 2.20 montrent, pour les exemples 1 et 2, l'évolution des valeurs (a) du critère d'arrêt et (b) du taux moyen d'erreur de classification en fonction de d , la dimension de l'espace couleur hybride sélectionné à l'itération de rang d de la procédure de construction.

Ces deux figures montrent la correspondance entre l'évolution des valeurs du critère d'arrêt et celle du taux moyen d'erreur de classification. Dans les figures 2.19 et 2.20, l'évolution des valeurs du critère d'arrêt est liée à l'évolution du taux d'erreur de classification. En effet, quand la valeur du critère d'arrêt J_{stop}^d commence à diminuer, le taux moyen d'erreur de classification $\bar{\varepsilon}$ commence à augmenter.

Dans le premier exemple, la valeur du critère d'arrêt atteint un maximum local lors de la seconde

itération de la procédure de construction. Ce critère fournit donc un espace couleur hybride le plus discriminant à deux dimensions (voir figure 2.19(a)). La figure 2.19(b) montre que le taux d'erreur moyen associé à cet espace couleur hybride est le plus faible. Cet exemple illustre la pertinence du critère d'arrêt utilisé pour arrêter la procédure de sélection des composantes couleur.

Dans le second exemple, la valeur du critère d'arrêt décroît après la seconde itération de la procédure de construction, mais elle atteint un maximum à la quatrième itération (voir figure 2.20(a)). La figure 2.20(b) indique que le taux d'erreur dans l'espace couleur hybride le plus discriminant de dimension deux est légèrement plus élevé que celui dans l'espace de dimension quatre. Dans le contexte d'une implantation en temps réel de la procédure de segmentation, la classification dans un espace de dimension quatre est plus coûteuse en temps de calcul que la segmentation effectuée avec un espace couleur hybride de dimension deux. Ainsi, dans cet exemple, l'amélioration des résultats de la classification obtenue avec l'espace couleur hybride le plus discriminant de dimension quatre n'est pas justifiée en termes de temps de calcul.

Les espaces couleur hybrides les plus discriminants qui sont sélectionnés dans ces deux exemples sont des plans et fournissent des résultats de segmentation similaires, voire meilleurs, que ceux obtenus avec les espaces couleur classiques de dimension trois. Pour une implantation en temps réel de la segmentation des images couleur successives, il est important de noter que la classification des pixels-joueurs qui s'effectue dans l'espace couleur hybride le plus discriminant de dimension deux est moins coûteuse en temps de calcul que la classification opérant dans un espace couleur classique à trois dimensions.

D'autres expériences ont montré que l'espace couleur hybride le plus discriminant coïncide rarement avec l'un des espaces couleur classiques même quand sa dimension est égale à trois [VtVP97, VtP97, VtP98a, VtP03]. En effet, l'espace couleur hybride le plus discriminant est déterminé par une analyse des propriétés de discrimination entre les classes de pixels en présence tandis que les espaces couleur classiques respectent des propriétés psycho-visuelles. Ceci explique que les composantes couleur d'un espace couleur hybride le plus discriminant ne proviennent pas nécessairement du même espace couleur classique.

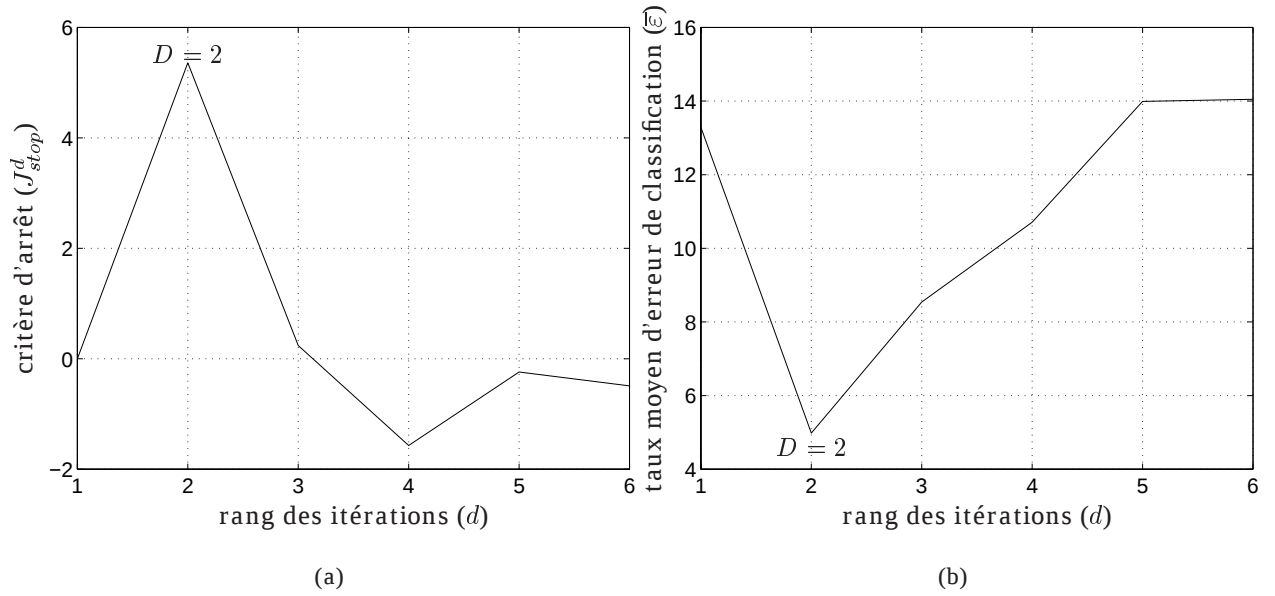


FIG. 2.19 – Évolution du critère d'arrêt et du taux moyen d'erreur de classification de l'exemple 1.

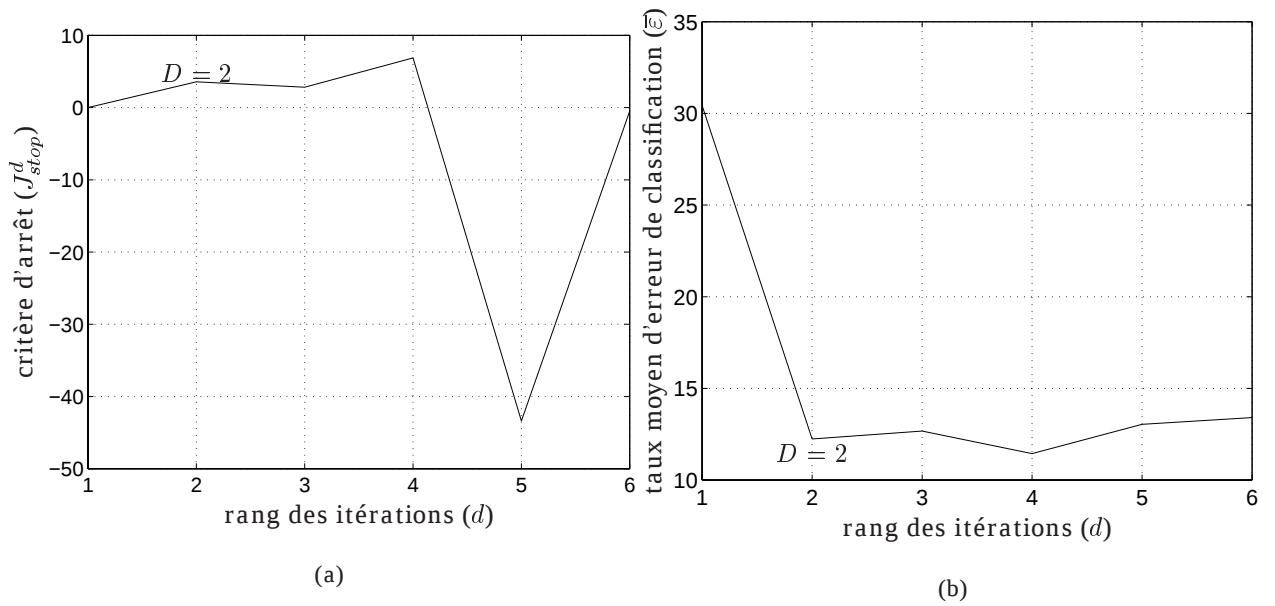


FIG. 2.20 – Évolution du critère d'arrêt et du taux moyen d'erreur de classification de l'exemple 2.

2.4 Conclusion

Dans ce chapitre, nous avons souligné l'influence des espaces de représentation de la couleur sur les résultats de segmentation des images couleur. Comme il n'existe aucun espace couleur classique qui fournit des résultats satisfaisants pour la segmentation de toutes les familles d'images, nous avons proposé de classer les pixels dans des espaces couleur hybrides qui sont spécifiquement conçus pour fournir la meilleure discrimination possible entre les classes de pixels en présence. Ces espaces couleur hybrides les plus discriminants sont construits par une procédure séquentielle de sélection des composantes dans un contexte supervisé. Ils contiennent un nombre optimal de composantes couleur au sens d'un critère informationnel. Ces composantes sont les moins corrélées possible et peuvent provenir de différents espaces couleur classiques.

Afin d'évaluer les performances atteintes par notre approche, nous avons utilisé des espaces couleur hybrides pour segmenter des images couleur. Les exemples présentés montrent qu'en utilisant l'espace couleur hybride le plus discriminant, une procédure de classification très simple fournit des résultats de segmentation très satisfaisants.

Les exemples d'images de foot-ball que nous avons présentés montrent également qu'il n'existe aucun espace couleur classique qui soit bien adapté à la discrimination entre toutes les classes de pixels-joueurs. La sélection de l'espace couleur hybride pour discriminer les classes de pixels-joueurs en présence est une solution pertinente pour la reconnaissance des équipes auxquelles appartiennent les joueurs.

Par ailleurs, la dimension de l'espace couleur hybride le plus discriminant sélectionné dans ces exemples est plus petite que la dimension des espaces classiques couleur. Dans de tels cas, la classification dans ces espaces couleur hybrides est moins coûteuse en temps de calcul que la classification opérant dans des espaces couleur classiques. En termes de complexité algorithmique des procédures de segmentation, l'utilisation de l'espace couleur hybride le plus discriminant est une solution pertinente pour la segmentation en temps réel des images couleur.

L'espace de représentation de la couleur étant sélectionné, nous nous proposons de détailler dans le chapitre suivant la phase de segmentation qui tient compte des propriétés spatio-colorimétriques des pixels.

2.5 Bibliographie

- [BB98] T. Bebie et H. Bieri. "Soccerman : reconstructing soccer games from video sequences". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 1, pages 898–902, Chicago (USA), 1998.
- [DL97] M. Dash et H. Liu. "Feature selection for classification". *Intelligent Data Analysis*, 1(3), 1997.
- [FHPZ96] C. Firmin, D. Hamad, J. G. Postaire et R. D. Zhang. "Feature extraction and selection for fault detection in production of glass bottles". *Machine Graphics and Vision*, 5(1) :77–86, 1996.
- [HATB98] J. F. Hair, R. E. Anderson, R. L. Tatham et W. Black. *Multivariate data analysis*. Prentice Hall, 1998.
- [KGP01] R. Kouassi, P. Gouton et M. Paindavoine. "Approximation of the karhunen-loève transformation and its application to colour images". *Signal Processing : Image Communication*, 16(6) :541–551, 2001.
- [KYA94] T. Kawashima, K. Yoshino et Y. Aoki. "Qualitative image analysis of group behaviour". In *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 690–693, Seattle, 1994.
- [LC99] P. Lambert et T. Carron. "Symbolic fusion of luminance-hue-chroma features for region segmentation". *Pattern Recognition*, 32(11) :1857–1872, 1999.
- [Lez00] O. Lezoray. *Segmentation d'images couleur par morphologie mathématique et classification de données par réseaux de neurones : application à la classification de cellules en cytologie des séreuses*. Thèse de doctorat, Université de Caen, janvier 2000.
- [LR97] E. Littman et E. Ritter. "Colour image segmentation : A comparison of neural and statistical methods". *IEEE Trans. on Neural Networks*, 8(1) :175–185, 1997.

- [NFM⁺00] M. Naemura, A. Fukuda, Y. Mizutani, Y. Izumi, Y. Tanaka et K. Enami. "Morphological segmentation of sport scenes using color information". *IEEE Trans. on Broadcasting*, 46(3) :181–188, 2000.
- [OKS80] Y. I. Ohta, T. Kanade et T. Sakai. "Color information for region segmentation". *Computer Graphics and Image Processing*, 13 :222–241, 1980.
- [OPR78] R. Ohlander, K. Price et D. R. Reddy. "Picture segmentation using a recursive region splitting method". *Computer Graphics and Image Processing*, 8 :313–333, 1978.
- [PO01] F. Pernkopf et P. O'Leary. "Automatic inspection system for detection and classification of flaws on turned parts". In *Proc. of the Int. Conf. on Quality Control by Artificial Vision*, pages 359–364, Le Creusot (France), 2001.
- [Rao73] C. R. Rao. *Linear statistical inference and its applications*. John Wiley and Sons, 1973.
- [Rom73] J.-M. Romeder. *Méthodes et programmes d'analyse discriminante*. Dunod, 1973. ISBN 2-04-007719-7.
- [SCKH97] Y. Seo, S. Choi, H. Kim et K.-S. Hong. "Where are the ball and players? Soccer game analysis with color-based tracking and image mosaick". In *Proc. of the IAPR Int. Conf. on Image Analysis and Processing*, pages 196–203, Florence (Italie), 1997.
- [SO99] S. Sudo et S. Ozawa. "Scene analysis of soccer game". In *Proc. of the Int. Conf. on Quality Control by Artificial Vision*, pages 119–123, Trois-Rivières (Canada), 1999.
- [SPNP99] P. Somol, P. Pudil, J. Novovicova et P. Paclik. "Adaptive floating search methods in feature selection". *Pattern Recognition Letters*, 20(8) :1157–1163, 1999.
- [Tom92] S. Tominaga. "Color classification of natural color images". *Color Research and Application*, 17(4) :230–239, 1992.
- [Van00] N. Vandenbroucke. *Segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimétriques adaptés. Application à l'analyse d'images de football*. Thèse de doctorat, Université Lille 1, décembre 2000.
- [VtP97] N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Soccer player recognition by pixels classification in an hybrid color space". In *Proc. of the SPIE Int. Conf. on Multispectral and Hyperspectral Imagery III*, volume 3071, pages 23–33, Orlando (USA), 1997.
- [VtP98a] N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color pixels classification in an hybrid color space". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 1, pages 176–180, Chicago (USA), 1998.
- [VtP98b] N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Un algorithme de classification couleur. Application à la reconnaissance de joueurs de football". Dans *Actes du Colloque Int. sur le Traitement d'Images et les Systèmes de Vision Artificielle*, volume 1, pages 121–133, Oujda (Maroc), 1998.
- [VtP00a] N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color image segmentation by supervised pixels classification in a color texture features space". In *Proc. of the IAPR Int. Conf. on Pattern Recognition*, volume 3, pages 625–628, Barcelone (Espagne), 2000.
- [VtP00b] N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color systems coding for color image processing". In *Proc. of the Int. Conf. on Color in Graphics and Image Processing*, volume 1, pages 180–185, Saint-Etienne, 2000.
- [VtP00c] N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Unsupervised color texture features extraction and selection for soccer images segmentation". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 2, pages 800–803, Vancouver (Canada), 2000.
- [VtP03] N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis.". *Computer Vision and Image Understanding*, 90(2) :190–216, 2003.
- [VtPV04] N. Vandenbroucke, **L. Macaire**, J. G. Postaire et C. Vieren. Chapitre "Applications - Analyse de séquences d'images couleur de matchs de football pour le suivi des joueurs", pages 342–351, Dans *Image couleur : de l'acquisition au traitement*, Auteurs du livre : A. Trémeau, C. Fernandez-Maloigne et P. Bonton. Editions Dunod, Paris, 2004.
- [VtV97] N. Vandenbroucke, **L. Macaire** et C. Vieren. "Suivi automatique de personnes en mouvement par analyse d'images couleurs successives. Application au suivi de joueurs de football". Dans *Actes du GRETSI'97 Colloque sur le Traitement du Signal et des Images*, volume 2, pages 917–920, Grenoble, 1997.
- [VtVP97] N. Vandenbroucke, **L. Macaire**, C. Vieren et J. G. Postaire. "Contribution of a color classification to soccer players tracking with snakes". In *Proc. of the IEEE Int. Conf. on System, Man, and Cybernetics*, volume 4, pages 3660–3665, Orlando (USA), 1997.

Chapitre 3

Segmentation par classification spatio-colorimétrique des pixels

3.1 Introduction

Généralement, les méthodes de segmentation par classification de pixels construisent les classes en privilégiant les propriétés colorimétriques des pixels. Or, l'exemple simple présenté dans la figure 3.1 montre que l'analyse seule de la distribution des points-couleur représentant les pixels dans un espace couleur ne permet pas toujours de construire les classes de pixels qui correspondent aux régions de l'image.

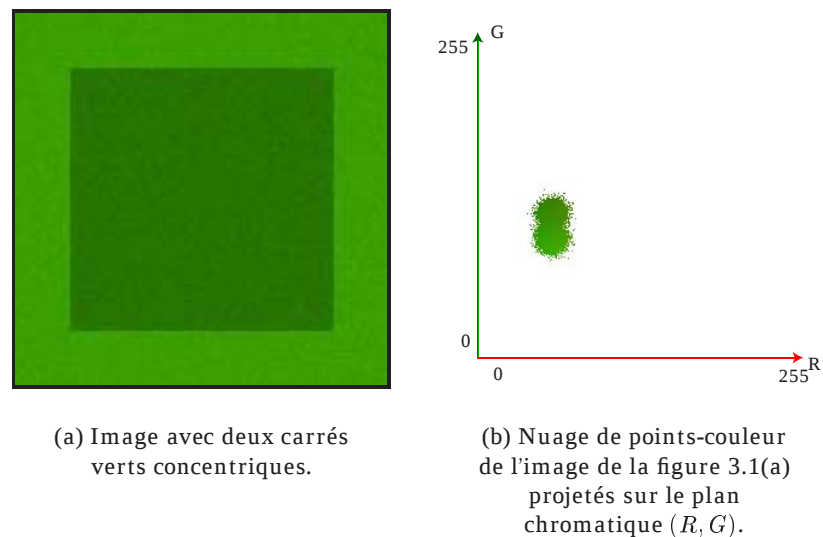


FIG. 3.1 – Deux carrés concentriques.

La prise en compte simultanée des propriétés colorimétriques et de l'arrangement spatial des pixels dans l'image constitue donc une démarche naturelle pour construire les classes de pixels de telle sorte qu'elles correspondent effectivement aux régions de l'image.

Nous avons vu au paragraphe 1.2.2 de la page 65 que nous pouvons construire un domaine couleur de l'espace qui définit un sous-ensemble-couleur de pixels dont les points-couleur appartiennent au domaine considéré (voir figure 3.2).

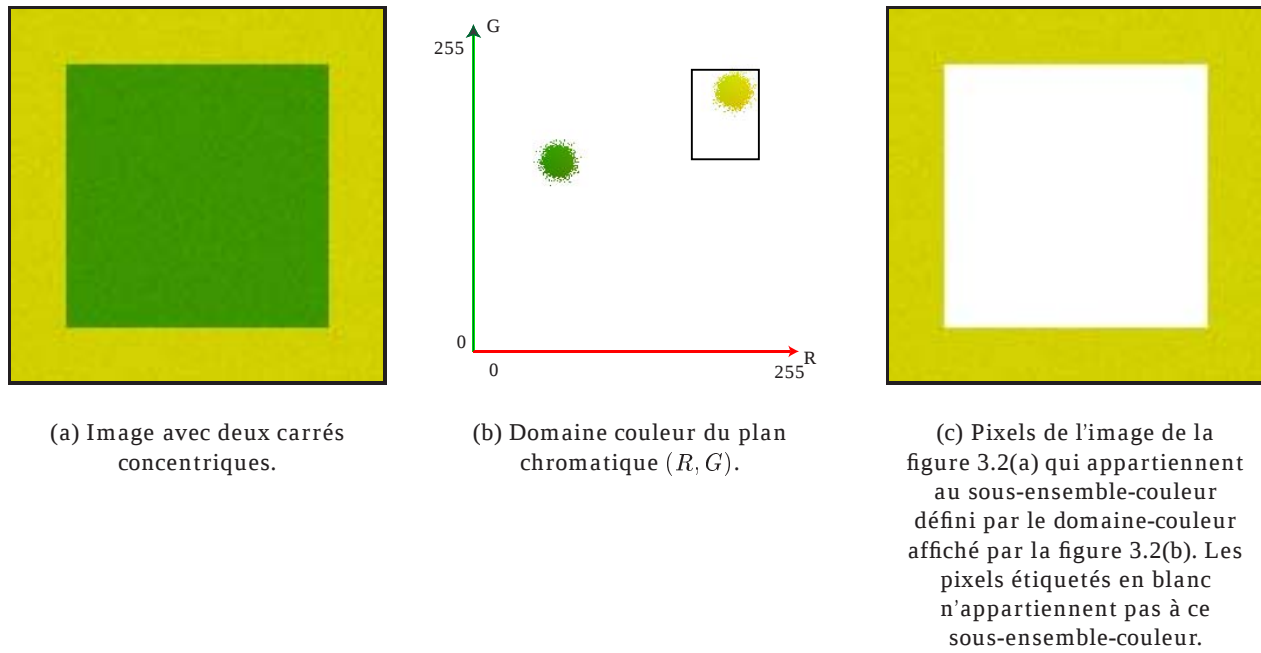


FIG. 3.2 – Un sous-ensemble-couleur de pixels défini par un domaine-couleur de l'espace couleur.

Nous supposons que chaque région de l'image peut être décrite comme un sous-ensemble-couleur de pixels connexes ayant des couleurs homogènes. La stratégie de construction de classes que nous proposons consiste à identifier des domaines-couleur de l'espace couleur qui définissent des sous-ensembles-couleur constitués de pixels connexes dans l'image avec des couleurs les plus homogènes possible.

Pour sélectionner les domaines-couleur pertinents de l'espace couleur, nous proposons d'analyser les sous-ensembles-couleur de pixels qu'il définissent, simultanément en termes de distribution de points-couleur dans l'espace couleur et de propriétés de connexité des pixels dans l'image.

3.1.1 Propriétés de connexité

Considérons un sous-ensemble-couleur S de pixels dans le plan image. Un chemin π de longueur n , partant du pixel P pour aboutir au pixel Q , est une séquence de pixels $P = P_0, \dots, P_{k-1}, P_k, \dots, P_n = Q$ tels que le pixel P_k est l'un des huit voisins de P_{k-1} , $1 \leq k \leq n$. Le pixel P est connexe au pixel Q dans le sous-ensemble-couleur S s'il existe au moins un chemin allant de P à Q , exclusivement constitué de pixels appartenant à S . Le sous-ensemble-couleur S constitue une région si chaque pixel de S est connexe à tous les autres pixels de S . Les propriétés de connexité des pixels qui appartiennent au sous-ensemble-couleur peuvent être quantifiées en supposant que plus le nombre de chemins possibles entre deux pixels P et Q du sous-ensemble-couleur S est élevé, plus les pixels de S sont connexes.

La figure 3.3 affiche trois images avec trois sous-ensembles-couleur différents constitués du même nombre de pixels étiquetés en gris. Dans la figure 3.3(a), les pixels étiquetés en gris sont dispersés dans l'image sans aucune structure ni cohésion. Il n'existe aucun chemin entre deux pixels de ce sous-ensemble-couleur. Comme les pixels du sous-ensemble-couleur de la figure 3.3(b) forment une ligne dans l'image, il existe au moins un chemin entre chaque couple de pixels de ce sous-ensemble-couleur. Dans le sous-ensemble de pixels de forme rectangulaire de la figure 3.3(c), il existe au moins deux chemins entre tout couple de pixels de ce sous-ensemble-couleur. Ainsi, les pixels du sous-ensemble-couleur de la figure 3.3(a) ne sont pas connexes, et les pixels du sous-ensemble-couleur de la fi-

gure 3.3(b) sont moins connexes que ceux du sous-ensemble-couleur de la figure 3.3(c).

Cet exemple montre que la quantification des propriétés de connexité d'un sous-ensemble-couleur s'avère intéressante pour mesurer le degré de confiance que nous pouvons accorder à la proposition : "le sous-ensemble-couleur constitue effectivement une région de l'image".

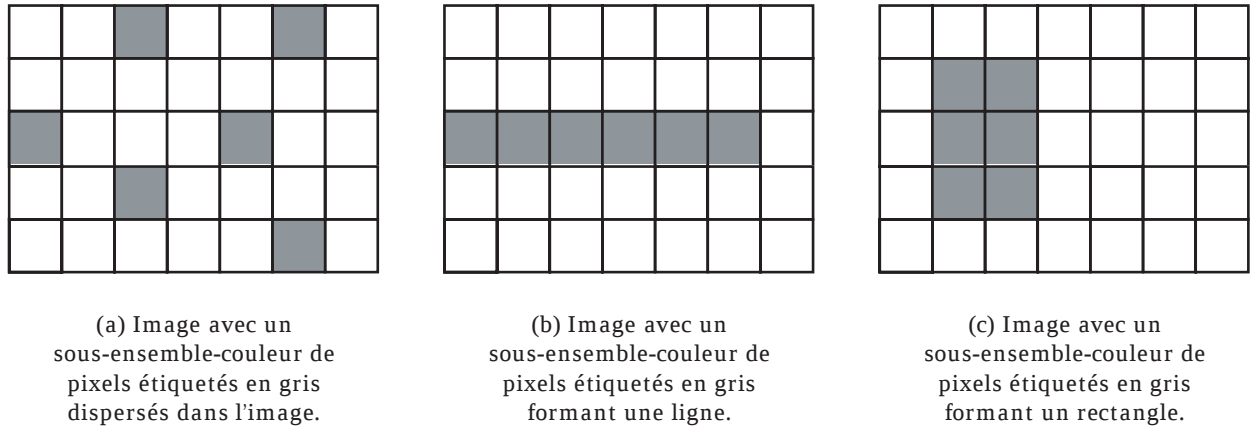


FIG. 3.3 – Trois sous-ensembles-couleur avec différentes propriétés de connexité.

3.1.2 Plan du chapitre

Le problème clé de la segmentation par classification des pixels est de construire des classes de pixels de telle sorte que chacune d'elles corresponde à une région de l'image. Nous supposons que chaque région peut être considérée comme un sous-ensemble-couleur de pixels fortement connexes ayant des couleurs homogènes. Pour construire des classes de pixels correspondant aux régions de l'image, la stratégie proposée consiste à sélectionner des domaines-couleur de l'espace couleur qui définissent des sous-ensembles-couleur constitués de pixels fortement connexes dans l'image et ayant des couleurs les plus homogènes possible.

Cette stratégie nécessite d'examiner simultanément les propriétés de connexité et d'homogénéité colorimétrique des sous-ensembles-couleur associés aux domaines-couleur considérés. Afin de mesurer ces propriétés, nous présentons le concept de *degré de connexité* et d'*homogénéité* d'un sous-ensemble-couleur dans le second paragraphe de ce chapitre. Plus les pixels d'un sous-ensemble-couleur sont connexes dans l'image et plus les points-couleur représentant ces pixels sont proches dans l'espace couleur, plus le degré de connexité et d'homogénéité de ce sous-ensemble-couleur est élevé.

La construction des classes de pixels consiste alors à sélectionner les domaines-couleur de l'espace couleur qui définissent ces classes. Nous supposons que ces domaines sont ceux qui sont associés aux sous-ensembles-couleur constitués de pixels les plus connexes possible avec des couleurs les plus homogènes possible. Dans le troisième paragraphe, nous détaillons la procédure de construction des classes basée sur cette stratégie.

Le dernier paragraphe est consacré à la présentation de résultats de segmentation afin de démontrer l'efficacité de notre procédure de segmentation basée sur une classification spatio-colorimétrique des pixels.

3.2 Degré de connexité et d'homogénéité d'un sous-ensemble-couleur

Nous supposons qu'une région peut être considérée comme un sous-ensemble-couleur de pixels fortement connexes ayant des couleurs homogènes, c'est à dire de pixels fortement connexes représentés par des points-couleur proches dans l'espace couleur. Pour analyser de telles propriétés d'un sous-ensemble-couleur, nous proposons de mesurer le *degré de connexité* et *d'homogénéité* qui quantifie simultanément l'arrangement spatial de ses pixels dans le plan image et la dispersion des points-couleur associés dans l'espace couleur. Ce degré est défini comme le produit de deux termes, le *degré de connexité* et le *degré d'homogénéité*. Le degré de connexité du sous-ensemble-couleur reflète l'arrangement spatial de ses pixels dans le plan image [FtP00c]. Le degré d'homogénéité d'un sous-ensemble-couleur dépend de la dispersion des points-couleur représentant ses pixels dans l'espace couleur.

3.2.1 Sous-ensemble couleur

Soit P un pixel d'une image couleur I dont les niveaux des composantes couleur sont les coordonnées du vecteur $I(P) = [I^1(P), I^2(P), I^3(P)]^T$ où $I^i(P)$, $i = 1, 2, 3$, est le niveau de la i^{eme} composante couleur. Dans l'espace couleur (R, G, B) , les 1^{ere} , 2^{nde} et 3^{eme} composantes couleur correspondent respectivement au rouge, au vert et au bleu. Les niveaux des composantes couleur varient entre 0 et $L - 1$, où L indique le nombre de niveaux de chaque composante couleur.

Soit $\underline{D}(c, l)$ un domaine-couleur de forme parallélépipédique rectangle, centré en un point-couleur $c = [c^1, c^2, c^3]^T$ dans l'espace couleur dont les côtés sont parallèles aux axes de cet espace. La taille du domaine-couleur $\underline{D}(c, l)$ est représentée par le *vecteur-taille* $l = [l^1, l^2, l^3]^T$. Le domaine-couleur $\underline{D}(c, l)$ centré en c contient ainsi tous les points-couleur $C = [C^1, C^2, C^3]^T$ tels que : $c^i - \frac{l^i-1}{2} \leq C^i \leq c^i + \frac{l^i-1}{2}$, $i = 1, 2, 3$. Les coordonnées l^i , $i = 1, 2, 3$, du vecteur-taille sont impaires de sorte que les valeurs $\frac{l^i-1}{2}$ soient entières. Soit $S(c, l)$ le sous-ensemble-couleur constitué de tous les pixels P associés aux points-couleur $I(P)$ appartenant au domaine-couleur $\underline{D}(c, l)$.

La figure 3.4(a) est une image de synthèse constituée de six régions de formes et tailles différentes. Cette image est différente de l'image de la figure 1.15 de la page 53, bien que les formes des régions soient identiques. Les distributions des points-couleur représentant les pixels de ces deux images dans l'espace couleur (R, G, B) sont différentes. Pour simplifier la présentation, le niveau de bleu des pixels dans cette image est fixé à 0. La figure 3.4(b) indique comment les points-couleur représentant les pixels sont projetés sur le plan couleur (R, G) . Considérons le domaine-couleur $\underline{D}(\text{orange}, [41, 41, 1]^T)$ centré au point-couleur $[210, 120, 0]^T$ qui correspond à la couleur orange. Comme le niveau de bleu des pixels est nul, la coordonnée l^3 du vecteur-taille du domaine-couleur considéré est mise à 1. Le domaine $\underline{D}(\text{orange}, [41, 41, 1]^T)$ est représenté dans la figure 3.4(b) par un rectangle situé dans la partie droite du plan couleur (R, G) . L'image de la figure 3.4(c) montre que le sous-ensemble-couleur $S(\text{orange}, [41, 41, 1]^T)$ contient tous les pixels du carré orange de l'image de la figure 3.4(a).

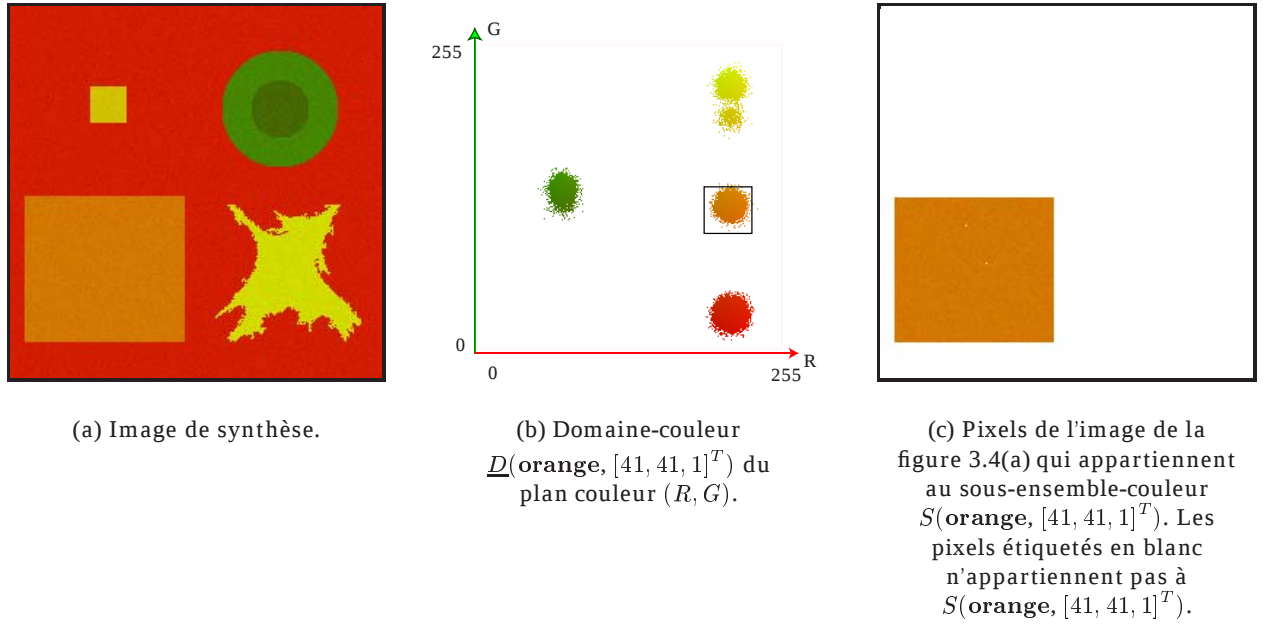


FIG. 3.4 – L'image de la figure 3.4(c) présente le sous-ensemble-couleur $S(\text{orange}, [41, 41, 1]^T)$ de l'image de la figure 3.4(a) défini par le domaine-couleur $\underline{D}(\text{orange}, [41, 41, 1]^T)$ représenté par la figure 3.4(b).

3.2.2 Degré de connexité

Définition

Soit $N_S(P)$ le sous-ensemble des pixels voisins d'un pixel P qui appartiennent à S . La connexité entre P et le sous-ensemble-couleur S , notée $\gamma_S(P)$, dépend du cardinal $N_S(P)$, à savoir du nombre de pixels voisins Q qui appartiennent à $N_S(P)$. Elle est exprimée par :

$$\gamma_S(P) = \frac{\text{Card}\{Q \in N_S(P)\}}{8}, \quad (3.1)$$

où le facteur de normalisation 8 correspond au nombre de pixels voisins considéré (voir figure 3.5).

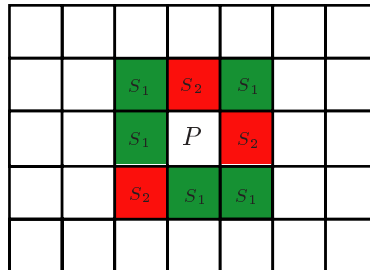


FIG. 3.5 – Cinq des 8-voisins de P appartiennent à S_1 et trois des 8-voisins de P appartiennent à S_2 . $\gamma_{S_1}(P) = \frac{5}{8}$, $\gamma_{S_2}(P) = \frac{3}{8}$.

Afin de définir une mesure de connexité du sous-ensemble-couleur S , nous introduisons le degré de

connexité, noté $DC(S)$ et défini par :

$$DC(S) = \frac{\sum_{P \in S} \gamma_S(P)}{\text{Card}\{P \in S\}}. \quad (3.2)$$

Le degré de connexité d'un sous-ensemble-couleur vide est considéré comme nul. Le degré de connexité $DC(S)$ dépend du nombre moyen de voisins des pixels de S qui appartiennent également à S . Le degré de connexité est évalué comme la moyenne des cardinaux des sous-ensembles $N_S(P)$ calculés pour tous les pixels P appartenant à S . L'une des propriétés du degré de connexité d'un sous-ensemble-couleur est son indépendance vis-à-vis du cardinal du sous-ensemble considéré.

Un degré de connexité faible, proche de 0, signifie que les pixels du sous-ensemble-couleur considéré sont dispersés dans l'image. A l'opposé, un degré de connexité élevé, proche de 1, indique que les pixels du sous-ensemble-couleur considéré sont fortement connexes dans l'image.

Degrés de connexité et régions

Afin de montrer la relation entre les degrés de connexité de sous-ensembles-couleur et les régions de l'image, considérons l'image de synthèse des figures 3.4(a) et 3.6(a). Afin de simplifier la présentation, le niveau de bleu des pixels est toujours fixé à 0. Deux domaines-couleur $\underline{D}(\text{orange}, [41, 41, 1]^T)$ et $\underline{D}(\text{rouge}, [41, 41, 1]^T)$ dont les vecteurs-taille $[41, 41, 1]^T$ sont égaux, sont représentés par des carrés dans le plan couleur (R, G) respectivement dans les figures 3.4(b) et 3.6(b). Les pixels qui appartiennent aux sous-ensembles-couleur $S(\text{orange}, [41, 41, 1]^T)$ et $S(\text{rouge}, [41, 41, 1]^T)$ des images des figures 3.4(a) et 3.6(a) sont respectivement représentés dans les images des figures 3.4(c) et 3.6(c).

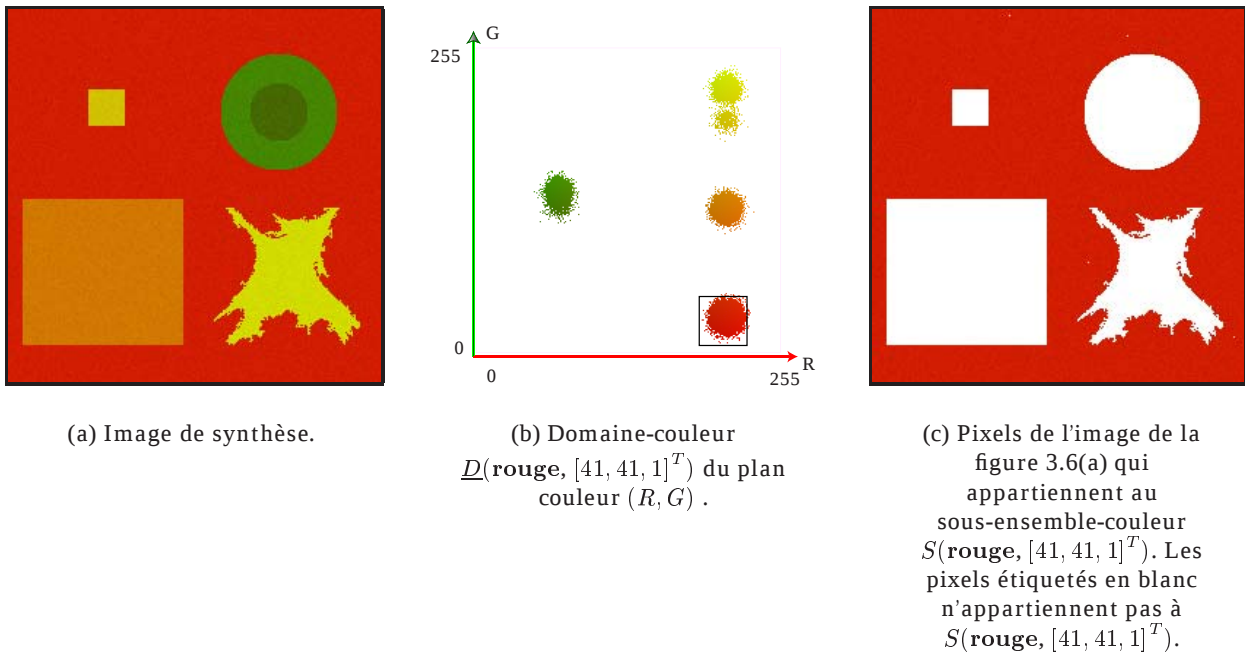


FIG. 3.6 – L'image de la figure 3.6(c) présente le sous-ensemble-couleur $S(\text{rouge}, [41, 41, 1]^T)$ de l'image de la figure 3.6(a) défini par le domaine-couleur $\underline{D}(\text{rouge}, [41, 41, 1]^T)$ représenté par la figure 3.6(b).

Le tableau 3.1 indique les degrés de connexité des sous-ensembles-couleur $S(\text{orange}, [41, 41, 1]^T)$ et

$S(\text{rouge}, [41, 41, 1]^T)$ qui correspondent respectivement au carré orange et au fond rouge dans les images des figures 3.4(c) et 3.6(c). Les degrés de connexité de ces deux sous-ensembles-couleur qui correspondent à deux régions de surfaces différentes, sont presque égaux et proches de 1. Ces exemples montrent que les pixels qui appartiennent à un sous-ensemble-couleur avec un degré de connexité élevé tendent à être fortement connexes dans l'image, de telle sorte qu'ils soient considérés comme constituant une région. Le prochain exemple montre que moins les pixels d'un sous-ensemble-couleur sont connexes, plus le degré de connexité du sous-ensemble-couleur est faible.

Mesure	$S(\text{orange}, [41, 41, 1]^T)$	$S(\text{rouge}, [41, 41, 1]^T)$
$DC(S)$	0,985	0,978

TAB. 3.1 – Degrés de connexité des sous-ensembles-couleur $S(\text{orange}, [41, 41, 1]^T)$ et $S(\text{rouge}, [41, 41, 1]^T)$ représentés respectivement dans les images des figures 3.4(c) et 3.6(c).

On considère un domaine centré sur le même point-couleur rouge, mais de taille plus petite que celles des domaines-couleur précédemment examinés (voir figure 3.7). L'image de la figure 3.7(a) est la même image de synthèse que celle des figures 3.4(a) et 3.6(a). Le domaine-couleur $\underline{D}(\text{rouge}, [7, 7, 1]^T)$ représenté par la figure 3.7(b) définit le sous-ensemble-couleur $S(\text{rouge}, [7, 7, 1]^T)$ dont les pixels sont présentés dans l'image de la figure 3.7(c). Le tableau 3.2 montre que son degré de connexité est faible. En effet, comme la couleur du fond de l'image n'est pas uniforme, les pixels qui appartiennent à ce sous-ensemble-couleur sont dispersés dans l'image (voir figure 3.7(c)).

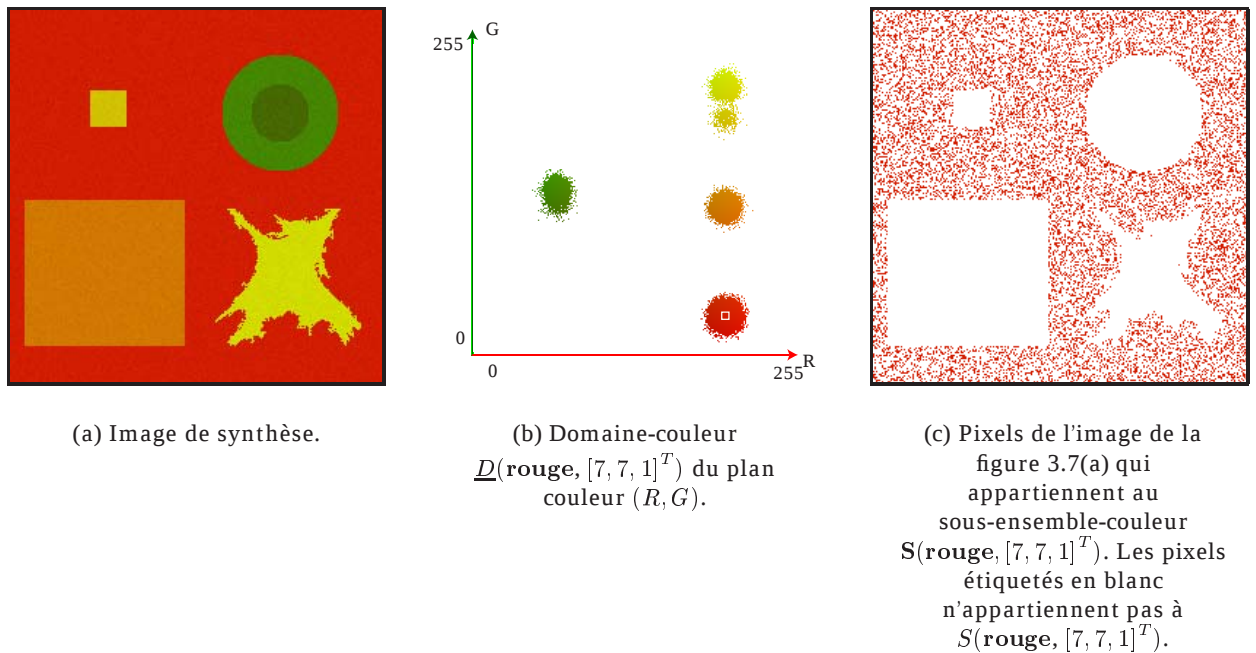


FIG. 3.7 – L'image de la figure 3.7(c) présente le sous-ensemble-couleur $S(\text{rouge}, [7, 7, 1]^T)$ de l'image de la figure 3.7(a) défini par le domaine-couleur $\underline{D}(\text{rouge}, [7, 7, 1]^T)$ représenté par la figure 3.7(b).

Considérons maintenant le sous-ensemble-couleur $S(\text{rouge}, [41, 41, 1]^T)$ présenté dans l'image de la figure 3.6(c). Comme le domaine-couleur $\underline{D}(\text{rouge}, [41, 41, 1]^T)$ est plus grand que le domaine-couleur $\underline{D}(\text{rouge}, [7, 7, 1]^T)$, le sous-ensemble-couleur $S(\text{rouge}, [41, 41, 1]^T)$ contient un nombre plus élevé de pixels que le sous-ensemble-couleur $S(\text{rouge}, [7, 7, 1]^T)$. Par ailleurs, l'image de la figure 3.6(c) montre que, comme tous les pixels de $S(\text{rouge}, [41, 41, 1]^T)$ correspondent au fond de l'image, ils sont fortement connexes. Dans ce cas, le degré de connexité de ce sous-ensemble-couleur est élevé et proche de 1.

Mesure	$S(\text{rouge}, [7, 7, 1]^T)$	$S(\text{rouge}, [41, 41, 1]^T)$
$DC(S)$	0,254	0,978

TAB. 3.2 – Degrés de connexité des sous-ensembles-couleur $S(\text{rouge}, [7, 7, 1]^T)$ et $S(\text{rouge}, [41, 41, 1]^T)$ représentés respectivement dans les images des figures 3.7(c) et 3.6(c).

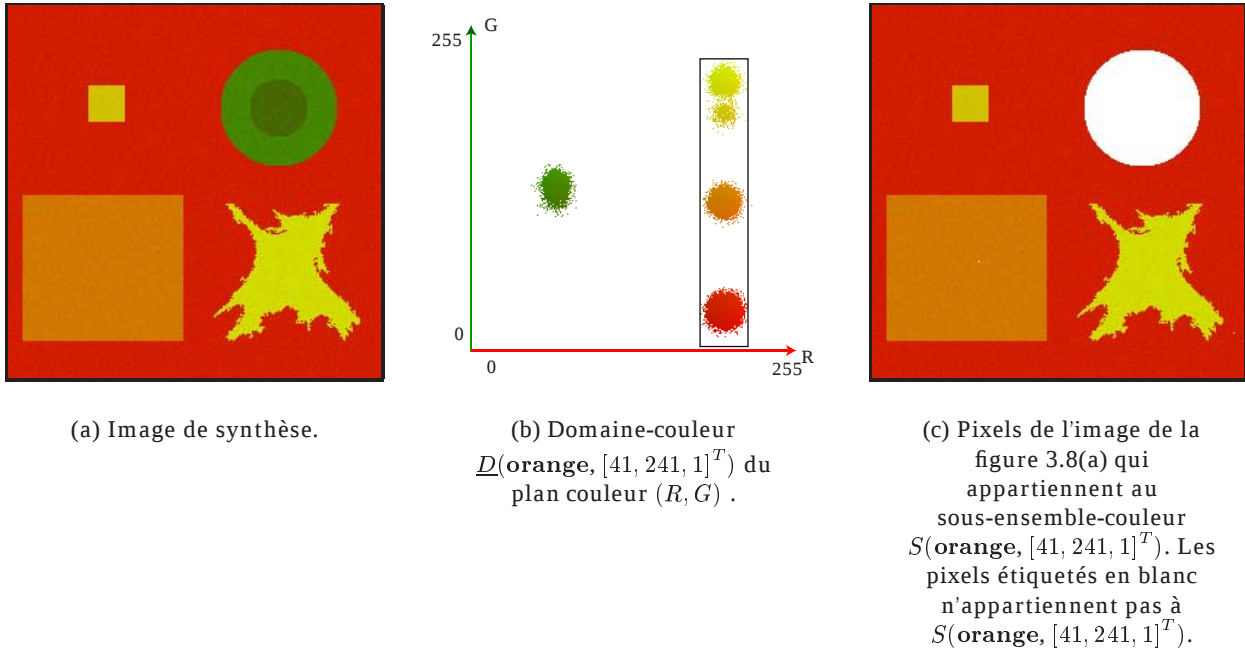


FIG. 3.8 – L'image de la figure 3.8(c) présente le sous-ensemble-couleur $S(\text{orange}, [41, 241, 1]^T)$ de l'image de la figure 3.8(a) défini par le domaine-couleur $\underline{D}(\text{orange}, [41, 241, 1]^T)$ représenté par la figure 3.8(b).

Cet exemple met en lumière l'influence du vecteur-taille du domaine-couleur sur le degré de connexité du sous-ensemble-couleur qu'il définit. Si le domaine-couleur est trop petit, le sous-ensemble-couleur qu'il définit n'est pas représentatif d'une région de l'image et son degré de connexité est faible.

Nous montrons dans l'image de la figure 3.8(c) les pixels appartenant au sous-ensemble-couleur $S(\text{orange}, [41, 241, 1]^T)$ de l'image de synthèse de la figure 3.8(a) défini par le domaine-couleur $\underline{D}(\text{orange}, [41, 241, 1]^T)$ représenté par la figure 3.8(b). Comparons les images des figures 3.4(c) et 3.8(c) qui présentent respectivement les sous-ensembles $S(\text{orange}, [41, 41, 1]^T)$ et $S(\text{orange}, [41, 241, 1]^T)$ de l'image de la figure 3.4(a). Le sous-ensemble-couleur $S(\text{orange}, [41, 41, 1]^T)$ représente le carré orange alors que $S(\text{orange}, [41, 241, 1]^T)$ correspond à plusieurs régions de différentes couleurs dans l'image. Le tableau 3.3 montre que leurs degrés de connexité atteignent tous les deux des niveaux élevés proches de 1.

Mesure	$S(\text{orange}, [41, 41, 1]^T)$	$S(\text{orange}, [41, 241, 1]^T)$
$DC(S)$	0,985	0,992

TAB. 3.3 – Degrés de connexité des sous-ensembles-couleur $S(\text{orange}, [41, 41, 1]^T)$ et $S(\text{orange}, [41, 241, 1]^T)$ représentés respectivement dans les images des figures 3.4(c) et 3.8(c).

Cet exemple démontre que, si un domaine-couleur est trop grand, il peut contenir des nuages de points-couleur séparés dans l'espace couleur qui représentent les pixels de plusieurs régions de différentes couleurs. La seule analyse des degrés de connexité des sous-ensembles-couleur ne permet pas toujours

de différencier un sous-ensemble-couleur qui correspond effectivement à une région dans l'image, d'un sous-ensemble-couleur qui constitue plusieurs régions de couleurs différentes. Par conséquent, la procédure de segmentation doit également tenir compte des propriétés d'homogénéité colorimétrique des sous-ensembles-couleur considérés.

3.2.3 Degré d'homogénéité

Définition

Nous proposons une nouvelle mesure des propriétés d'homogénéité colorimétrique d'un sous-ensemble-couleur, appelé le degré d'homogénéité. Il est basé sur la *mesure de dispersion*, notée $\sigma(S)$, des points-couleur représentant les pixels qui appartiennent au sous-ensemble-couleur S dans l'espace couleur. Cette mesure est exprimée par :

$$\sigma(S) = \frac{1}{\text{Card}\{P \in S\}} \times \sqrt{\sum_{P \in S} (\mathbf{I}(P) - \mathbf{M}(S))^T (\mathbf{I}(P) - \mathbf{M}(S))}, \quad (3.3)$$

où $\mathbf{M}(S)$ est le barycentre des points-couleur représentant les pixels qui appartiennent à S :

$$\mathbf{M}(S) = \frac{1}{\text{Card}\{P \in S\}} \times \sum_{P \in S} \mathbf{I}(P). \quad (3.4)$$

$\sigma(S)$ peut s'exprimer également comme la racine carrée de la trace de la matrice de co-variance calculée avec les points-couleur représentant les pixels du sous-ensemble-couleur S .

Afin de décider si un sous-ensemble-couleur correspond effectivement à une région, nous proposons de comparer une mesure globale de la dispersion des points-couleur représentant les pixels de ce sous-ensemble avec une mesure locale de la dispersion des points-couleur représentant les pixels voisins de tous les pixels qui appartiennent au sous-ensemble-couleur.

Pour chaque pixel P de S , nous déterminons la mesure de dispersion $\sigma(N_S(P))$ du sous-ensemble $N_S(P)$ constitué des pixels voisins de P qui appartiennent à S .

Soit $\sigma_{local}(S)$ la *mesure locale de dispersion* du sous-ensemble S , définie comme la moyenne des mesures de dispersion $\sigma(N_S(P))$ évaluées aux voisinages $N_S(P)$ de tous les pixels P appartenant à S . Cette mesure locale de dispersion du sous-ensemble-couleur S s'exprime sous la forme :

$$\sigma_{local}(S) = \frac{1}{\text{Card}\{P \in S\}} \times \sum_{P \in S} \sigma(N_S(P)). \quad (3.5)$$

Si la mesure locale de dispersion $\sigma_{local}(S)$ du sous-ensemble-couleur S est proche de $\sigma(S)$, les points-couleur représentant les pixels de ce sous-ensemble ont tendance à donner naissance à un nuage unique. Par contre, si $\sigma_{local}(S)$ est inférieure à $\sigma(S)$, les points-couleur tendent à être répartis en plusieurs nuages séparés dans l'espace couleur.

Afin de comparer $\sigma_{local}(S)$ et $\sigma(S)$, nous définissons le degré d'homogénéité, noté $DH(S)$, du sous-ensemble-couleur S par :

$$\begin{cases} DH(S) = 1 & \text{si } \sigma(S) = 0, \\ DH(S) = \frac{\sigma_{local}(S)}{\sigma(S)} & \text{si } \sigma(S) > 0. \end{cases} \quad (3.6)$$

Le degré d'homogénéité $DH(S)$ varie entre 0, quand $\sigma_{local}(S)$ est égale à 0, et 1 quand $\sigma_{local}(S)$ est

égale à $\sigma(S)$. Afin d'éviter que le degré d'homogénéité $DH(S)$ soit supérieur à 1, la valeur de $DH(S)$ est fixée à 1 si $\sigma_{local}(S)$ est supérieure à $\sigma(S)$. Nous allons examiner différents exemples pour mettre en évidence la relation qui existe entre les degrés d'homogénéité des sous-ensembles-couleur et les régions dans une image couleur.

Degrés d'homogénéité et régions

Le tableau 3.4 indique respectivement les mesures de dispersion et les degrés d'homogénéité des sous-ensembles-couleur $S(\text{rouge}, [7, 7, 1]^T)$ et $S(\text{rouge}, [41, 41, 1]^T)$ des images des figures 3.7(c) et 3.6(c). Comme les pixels de $S(\text{rouge}, [7, 7, 1]^T)$ sont dispersés dans l'image, de nombreux pixels de $S(\text{rouge}, [7, 7, 1]^T)$ comportent moins de deux voisins qui appartiennent également à $S(\text{rouge}, [7, 7, 1]^T)$. Les dispersions $\sigma(N_{S(\text{rouge}, [7, 7, 1]^T)}(P))$ mesurées aux voisinages de ces pixels P sont nulles. Ceci explique pourquoi la mesure locale de dispersion $\sigma_{local}(S(\text{rouge}, [7, 7, 1]^T))$ est significativement inférieure à la mesure de dispersion $\sigma(S(\text{rouge}, [7, 7, 1]^T))$ et pourquoi le degré d'homogénéité $DH(S(\text{rouge}, [7, 7, 1]^T))$, égal à 0,478, est relativement faible.

Mesure	$S(\text{rouge}, [7, 7, 1]^T)$	$S(\text{rouge}, [41, 41, 1]^T)$
$\sigma_{local}(S)$	1,308	6,467
$\sigma(S)$	3,92	7,067
$DH(S)$	0,478	0,915

TAB. 3.4 – Mesures de dispersion et degrés d'homogénéité des sous-ensembles-couleur $S(\text{rouge}, [7, 7, 1]^T)$ et $S(\text{rouge}, [41, 41, 1]^T)$ représentés respectivement dans les images des figures 3.7(c) et 3.6(c).

Comme le sous-ensemble-couleur $S(\text{rouge}, [41, 41, 1]^T)$ correspond effectivement au fond de l'image, les huit voisins de la plupart des pixels de $S(\text{rouge}, [41, 41, 1]^T)$ appartiennent également à $S(\text{rouge}, [41, 41, 1]^T)$ bien que leurs couleurs soient différentes. Les dispersions $\sigma(N_{S(\text{rouge}, [41, 41, 1]^T)}(P))$ mesurées aux voisinages de ces pixels P sont similaires et éloignées de 0. La mesure locale de dispersion $\sigma_{local}(S(\text{rouge}, [41, 41, 1]^T))$ est alors proche de la mesure de dispersion $\sigma(S(\text{rouge}, [41, 41, 1]^T))$, ce qui explique que le degré d'homogénéité $DH(S(\text{rouge}, [41, 41, 1]^T))$ soit élevé et égal à 0,915.

Mesure	$S(\text{orange}, [41, 41, 1]^T)$	$S(\text{orange}, [41, 241, 1]^T)$
$\sigma_{local}(S)$	6,488	9,84
$\sigma(S)$	7,080	58,40
$DH(S)$	0,916	0,168

TAB. 3.5 – Mesures de dispersion et degrés d'homogénéité des sous-ensembles-couleur $S(\text{orange}, [41, 41, 1]^T)$ et $S(\text{orange}, [41, 241, 1]^T)$ représentés respectivement dans les images des figures 3.4(c) et 3.8(c).

Le tableau 3.5 permet de comparer les mesures de dispersion et les degrés d'homogénéité des deux sous-ensembles-couleur $S(\text{orange}, [41, 41, 1]^T)$ et $S(\text{orange}, [41, 241, 1]^T)$ des images des figures 3.4(c) et 3.8(c). Comme $S(\text{orange}, [41, 41, 1]^T)$ représente le carré orange dans l'image, les points-couleur associés à ses pixels donnent naissance à un unique nuage compact dans l'espace couleur tandis que les points-couleur associés aux pixels de $S(\text{orange}, [41, 241, 1]^T)$, qui constituent plusieurs régions de différentes couleurs, forment plusieurs nuages distincts. Ceci explique pourquoi la mesure de dispersion $\sigma(S(\text{orange}, [41, 241, 1]^T))$ est supérieure à $\sigma(S(\text{orange}, [41, 41, 1]^T))$ tandis que la mesure locale de dispersion $\sigma_{local}(S(\text{orange}, [41, 241, 1]^T))$ est relativement proche de $\sigma_{local}(S(\text{orange}, [41, 41, 1]^T))$. Le degré d'homogénéité de $S(\text{orange}, [41, 41, 1]^T)$, qui représente le carré orange dans l'image est élevé, tandis

que le degré d'homogénéité de $S(\text{orange}, [41, 241, 1]^T)$, qui représente plusieurs régions de différentes couleurs, est faible.

Ces exemples montrent que le degré d'homogénéité $DH(S)$ mesure le degré de confiance que nous pouvons accorder à la proposition : "les points-couleur représentant les pixels du sous-ensemble-couleur S forment un unique nuage compact dans l'espace couleur". Par ailleurs, il donne une indication sur la probabilité que le sous-ensemble-couleur considéré corresponde effectivement à une région de l'image.

Malheureusement, cette mesure n'est pas suffisante pour déterminer dans tous les cas si un sous-ensemble-couleur correspond effectivement à une région de l'image. Considérons par exemple, le sous-ensemble-couleur $S(c, [1, 1, 1]^T)$ défini par le domaine-couleur $\underline{D}(c, [1, 1, 1]^T)$. Comme le vecteur-taille de ce domaine-couleur est égal à $[1, 1, 1]^T$, les couleurs de tous les pixels de ce sous-ensemble-couleur sont égales à c et la mesure de dispersion $\sigma(S(c, [1, 1, 1]^T))$ est nulle. Ainsi, le degré d'homogénéité $DH(S(c, [1, 1, 1]^T))$ est toujours égal à 1, même dans chacun des deux cas extrêmes suivants : quand les pixels de $S(c, [1, 1, 1]^T)$ correspondent effectivement à une région de l'image ou quand les pixels de $S(c, [1, 1, 1]^T)$ sont dispersés dans l'image et ne constituent aucune région.

L'analyse conjointe du degré de connexité et du degré d'homogénéité de ce sous-ensemble-couleur mérite alors d'être abordée pour estimer le degré de confiance à accorder à la proposition : "le sous-ensemble-couleur considéré correspond effectivement à une région de l'image".

3.2.4 Degré de connexité et d'homogénéité

Définition

Afin de tenir compte simultanément des propriétés de connexité et d'homogénéité colorimétrique d'un sous-ensemble-couleur, nous définissons le degré de connexité et d'homogénéité du sous-ensemble-couleur S , noté $DCH(S)$, comme le produit du degré de connexité avec le degré d'homogénéité :

$$DCH(S) = DC(S) \times DH(S). \quad (3.7)$$

Le degré de connexité et d'homogénéité atteint la valeur maximale 1 si son degré de connexité ainsi que son degré d'homogénéité sont tous les deux égaux à 1. Un niveau élevé du degré de connexité et d'homogénéité indique que les pixels du sous-ensemble-couleur sont fortement connexes dans l'image (degré de connexité proche de 1) et que les points-couleur représentant ses pixels forment un nuage compact dans l'espace couleur (degré d'homogénéité proche de 1). Par contre, le degré de connexité et d'homogénéité d'un sous-ensemble-couleur est proche de 0 quand soit son degré de connexité, soit son degré d'homogénéité est proche de 0. Un niveau faible du degré de connexité et d'homogénéité signifie donc que les pixels du sous-ensemble-couleur sont dispersés dans l'image ou que les points-couleur représentant les pixels du sous-ensemble ne forment pas un unique nuage compact dans l'espace couleur.

Degrés de connexité et d'homogénéité et régions

Comme pour les degrés de connexité et les degrés d'homogénéité, nous nous proposons d'examiner les degrés de connexité et d'homogénéité des sous-ensembles-couleur $S(\text{rouge}, [7, 7, 1]^T)$ et $S(\text{rouge}, [41, 41, 1]^T)$ des images des figures 3.7(c) et 3.6(c) (voir tableau 3.6). Le degré de connexité et d'homogénéité du sous-ensemble-couleur $S(\text{rouge}, [41, 41, 1]^T)$, qui constitue le fond de l'image de la figure 3.6(a), est supérieur à celui de $S(\text{rouge}, [7, 7, 1]^T)$ dont les pixels sont dispersés dans l'image. Cet exemple démontre que l'analyse du degré de connexité et d'homogénéité permet de différencier

un sous-ensemble-couleur qui ne constitue aucune région de l'image d'un sous-ensemble-couleur qui correspond effectivement à une région.

Mesure	$S(\text{rouge}, [7, 7, 1]^T)$	$S(\text{rouge}, [41, 41, 1]^T)$
$DCH(S)$	0, 123	0, 895

TAB. 3.6 – Degrés de connexité et d'homogénéité des sous-ensembles-couleur $S(\text{rouge}, [7, 7, 1]^T)$ et $S(\text{rouge}, [41, 41, 1]^T)$ représentés respectivement dans les images des figures 3.7(c) et 3.6(c).

Intéressons-nous maintenant aux degrés de connexité et d'homogénéité des deux sous-ensembles-couleur $S(\text{orange}, [41, 41, 1]^T)$ et $S(\text{orange}, [41, 241, 1]^T)$ représentés respectivement dans les images des figures 3.9(a) et 3.9(b) (voir tableau 3.7). Ce tableau montre que le degré de connexité et d'homogénéité de $S(\text{orange}, [41, 241, 1]^T)$ qui représente plusieurs régions de différentes couleurs est égal à 0, 167, tandis que $S(\text{orange}, [41, 41, 1]^T)$ qui représente le carré orange atteint 0, 903. L'analyse du degré de connexité et d'homogénéité permet de distinguer un sous-ensemble-couleur correspondant à une région d'un sous-ensemble-couleur dont les pixels constituent plusieurs régions de couleurs différentes dans l'image.

Mesure	$S(\text{orange}, [41, 41, 1]^T)$	$S(\text{orange}, [41, 241, 1]^T)$
$DCH(S)$	0, 903	0, 167

TAB. 3.7 – Degrés de connexité et d'homogénéité des sous-ensembles-couleur $S(\text{orange}, [41, 41, 1]^T)$ et $S(\text{orange}, [41, 241, 1]^T)$ représentés respectivement dans les images des figures 3.9(a) et 3.9(b).

Il est également intéressant d'examiner l'évolution des degrés de connexité et d'homogénéité des sous-ensembles-couleur définis par des domaines-couleur dont les tailles sont comprises entre celle de $\underline{D}(\text{orange}, [41, 41, 1]^T)$ et celle de $\underline{D}(\text{orange}, [41, 241, 1]^T)$. La figure 3.9(c) montre l'évolution du degré de connexité et d'homogénéité du sous-ensemble-couleur $S((\text{orange}, [41, l^2, 1]^T)$ en fonction de l^2 qui est la longueur du domaine-couleur le long de la composante verte. Quand l^2 est inférieure à 41, le degré de connexité et d'homogénéité est une fonction croissante de l^2 . Quand l^2 est égale à 41, le degré de connexité et d'homogénéité atteint sa valeur maximale car le sous-ensemble-couleur $S((\text{orange}, [41, l^2, 1]^T)$ correspond au carré orange dans l'image (voir figure 3.9(a)). Quand l^2 est comprise entre 43 et 113, aucun autre point-couleur n'est ajouté aux domaines-couleur considérés. Ainsi, la même valeur élevée du degré du sous-ensemble-couleur $S((\text{orange}, [41, l^2, 1]^T)$ est conservée. Quand l^2 dépasse 113, $DCH(S((\text{orange}, [41, l^2, 1]^T))$ diminue quand l^2 augmente, car plusieurs pixels d'autres régions extérieures au carré orange appartiennent à $S((\text{orange}, [41, l^2, 1]^T)$ (voir figure 3.9(b)). Dans ce cas, les points-couleur représentant les pixels forment plusieurs nuages séparés dans l'espace couleur.

Le degré de connexité et d'homogénéité d'un sous-ensemble-couleur est sensible au nombre de nuages distincts formés par les points-couleur représentant ses pixels dans l'espace couleur. Si les points-couleur constituent plusieurs nuages bien séparés dans cet espace, le degré de connexité et d'homogénéité du sous-ensemble-couleur considéré est proche de 0, même si les pixels sont fortement connexes dans l'image. Ces exemples démontrent que l'analyse du degré de connexité et d'homogénéité d'un sous-ensemble-couleur de pixels permet de déterminer si le sous-ensemble-couleur correspond effectivement à une région de couleurs homogènes dans l'image.

3.3 Construction des classes de pixels

Avant de décrire la procédure de construction des classes de pixels, nous nous proposons d'examiner un exemple pour illustrer la manière dont les classes de pixels peuvent être définies par des domaines-

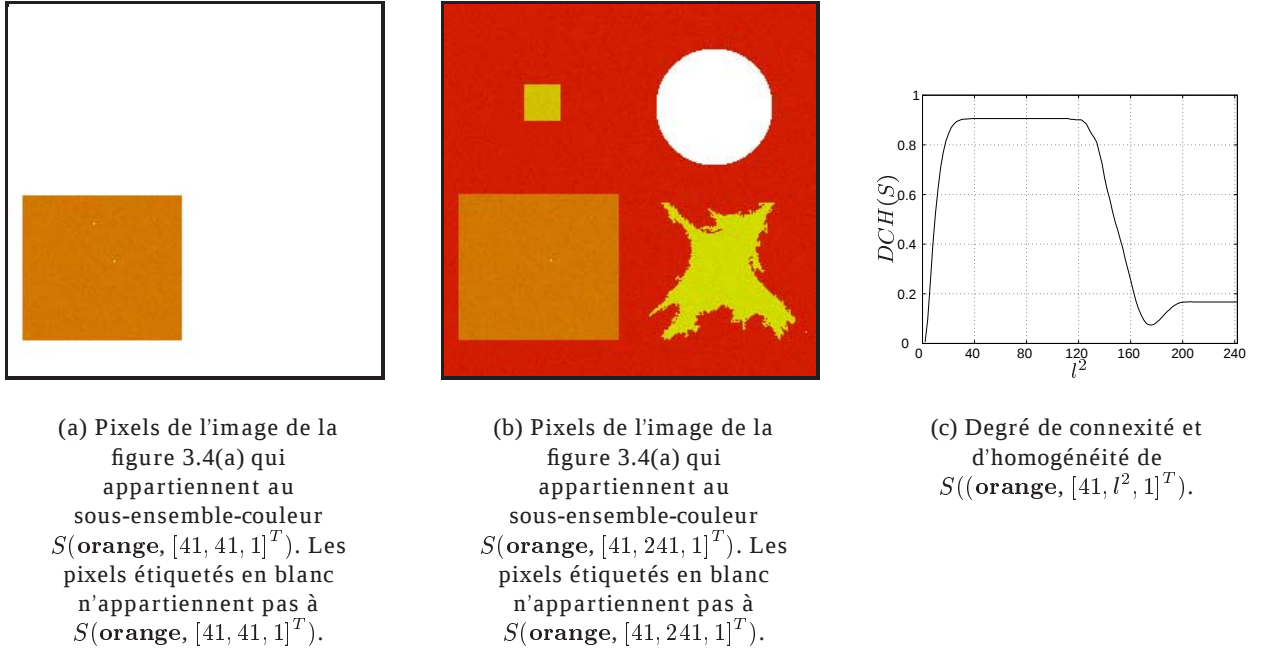


FIG. 3.9 – Évolution en fonction de l^2 du degré de connexité et d'homogénéité du sous-ensemble-couleur $S(\text{orange}, [41, l^2, 1]^T)$ dans l'image de la figure 3.4(a).

couleur afin qu'elles correspondent effectivement aux régions de l'image.

3.3.1 Classes définies par des domaines-couleur

Les six régions de l'image de synthèse de la figure 3.10(a) qui est la même que les images des figures 3.4(a), 3.6(a), 3.7(a) et 3.8(a), sont notées $\mathcal{R}_i$, $i = 1, \dots, 6$. La figure 3.10(b) présente les projections des points-couleur représentant les régions sur le plan couleur (R, G) . Les points-couleur représentatifs des régions $\mathcal{R}_1$ et $\mathcal{R}_2$ forment des nuages bien séparés tandis que les points-couleur provenant des régions $\mathcal{R}_3$ et $\mathcal{R}_4$ donnent naissance à deux nuages qui se chevauchent. Comme les points-couleur représentant les régions $\mathcal{R}_5$ et $\mathcal{R}_6$ constituent un seul nuage, ils ne peuvent pas être différenciés par la seule analyse de leur distribution dans le plan couleur (R, G) .

Cette image montre qu'il n'existe pas toujours une correspondance bijective entre les régions de l'image et les nuages de points-couleur dans l'espace couleur. Afin de résoudre ce problème, nous proposons de diviser les nuages en sous-populations de points-couleur définies par des domaines-couleur. La construction des classes passe alors par la sélection de domaines-couleur définissant des sous-ensembles-couleur qui correspondent effectivement aux régions de l'image.

Une segmentation convenable de l'image impose que les pixels qui constituent une région ne doivent appartenir qu'à un seul sous-ensemble-couleur défini par un des domaines-couleur sélectionnés. Ainsi, ces domaines-couleur sélectionnés pour définir les classes de pixels doivent être les mieux séparés possible. La procédure de construction des classes doit sélectionner des domaines-couleur ne se chevauchant pas dans l'espace couleur et définissant des sous-ensembles-couleur de sorte que chacun d'eux contienne le plus grand nombre possible de pixels constituant exclusivement la région correspondante dans l'image.

Examinons la figure 3.10(b) qui présente des domaines-couleur tels qu'ils pourraient être sélectionnés par la procédure de construction des classes car ils définissent des sous-ensembles-couleur correspon-

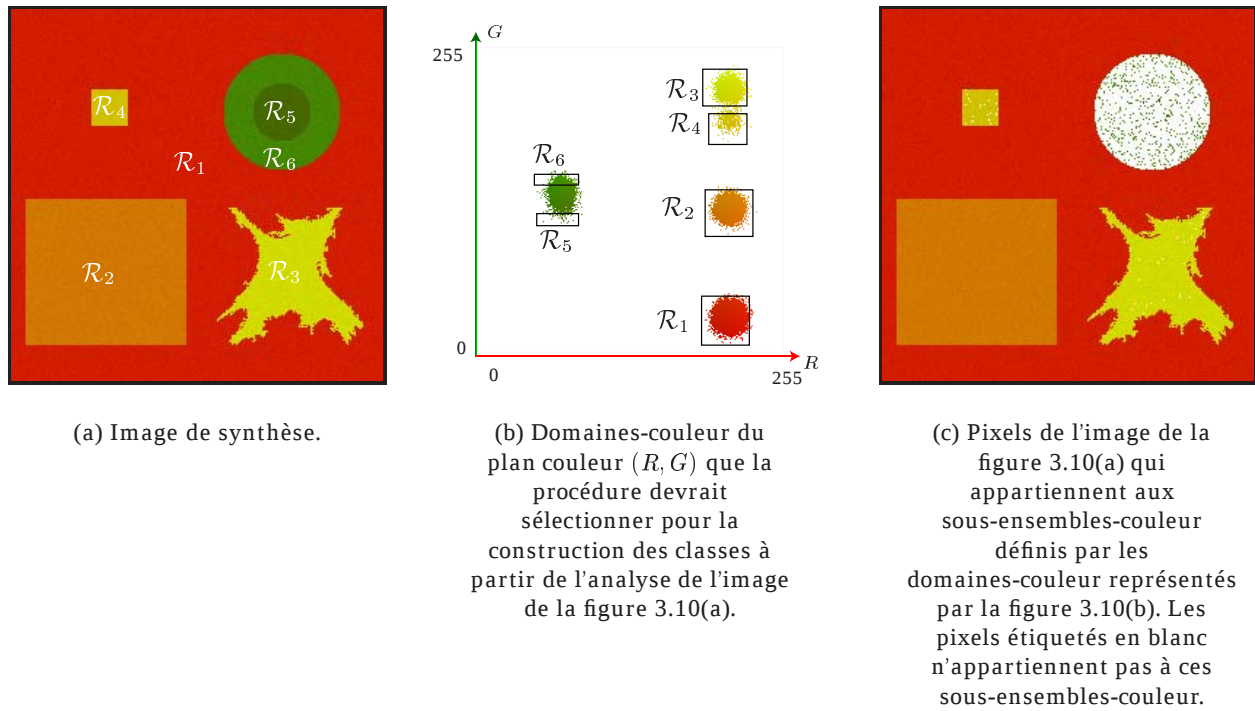


FIG. 3.10 – Exemples de domaines-couleur convenables pour la construction des classes de pixels.

dant effectivement aux régions dans l'image de la figure 3.10(a). L'image de la figure 3.10(c) représente les pixels qui appartiennent à ces sous-ensembles-couleur. La figure 3.10(b) montre que lorsque les points-couleur représentatifs des régions forment des nuages bien séparés, comme pour les régions $\mathcal{R}_1$ et $\mathcal{R}_2$, les domaines-couleur que la procédure doit sélectionner sont les plus petits domaines-couleur dans lesquels se trouvent les points-couleur représentant les pixels qui appartiennent à ces deux régions, et ceci afin d'éviter tout risque de chevauchement. Ces domaines-couleur ont été sélectionnés interactivement afin que chacun d'eux définisse un sous-ensemble-couleur exclusivement constitué de pixels qui appartiennent à l'une des deux régions $\mathcal{R}_1$ et $\mathcal{R}_2$ (voir figure 3.10(c)).

D'autre part, la figure 3.10(b) montre que les nuages formés par les points-couleur provenant de différentes régions peuvent se chevaucher, comme pour les deux couples de régions $\mathcal{R}_3$ et $\mathcal{R}_4$, $\mathcal{R}_5$ et $\mathcal{R}_6$. Afin de construire des classes qui correspondent à ces régions, la procédure doit déterminer des domaines-couleur ne se chevauchant pas de sorte que chacun d'eux contienne exclusivement les points-couleur représentatifs d'une seule région. Plus précisément, ces domaines-couleur doivent être les plus grands possible afin que chaque sous-ensemble-couleur défini par chacun d'eux contienne le plus grand nombre de pixels constituant exclusivement la région représentée (voir figure 3.10(c)). Dans ce cas, il est impossible de déterminer des domaines-couleur définissant des sous-ensembles-couleur qui contiennent tous les pixels de chacune des régions.

3.3.2 Description de la procédure

Une solution pour sélectionner les domaines-couleur qui définissent les classes de pixels correspondant effectivement aux régions de l'image couleur consisterait à analyser les degrés de connexité et d'homogénéité des sous-ensembles-couleur associés à tous les domaines-couleur pouvant être construits dans l'espace couleur.

Afin de simplifier la recherche, nous ne considérons que les domaines-couleur cubiques dont les lon-

guez des côtés sont égales à l . Même avec cette contrainte, le nombre de domaines-couleur cubiques que la procédure devrait analyser s'élèverait à $\sum_{l=1}^L (L - (l - 1))^3$, où L correspond au nombre de niveaux des composantes couleur. Quand L est fixé à 256, $1.08215 \cdot 10^9$ domaines-couleur cubiques devraient être examinés par la procédure. Même si ce nombre est inférieur à celui correspondant à l'ensemble des domaines-couleur non-cubiques pouvant être déterminés dans l'espace couleur, il n'est pas réaliste d'implanter une telle analyse exhaustive.

Ces problèmes d'implantation nous ont d'abord conduits à nous restreindre à l'analyse d'une *image bi-chromatique* où chaque pixel est caractérisé par deux composantes couleur parmi les trois disponibles. Avec Michael Fontaine, nous avons alors proposé un nouvel outil d'analyse : la *pyramide des degrés de connexité* qui recense de manière hiérarchique et organisée les degrés de connexité des sous-ensembles-couleur de pixels définis par tous les domaines-couleur pouvant être construits dans un plan couleur [FtP00b].

Le nombre total de cellules qui constituent la pyramide des degrés de connexité est alors égal à $\sum_{l=1}^L (L - (l - 1))^2$. Quand chacune des deux composantes couleur est codée sur 256 niveaux, la pyramide des degrés de connexité est constituée de 5625216 cellules. Sachant que le contenu de chaque cellule peut être codé sur 4 octets pour être mémorisée, la pyramide occupe environ 21,5 Mo d'espace mémoire. L'algorithme de construction de la pyramide des degrés de connexité a été implanté sur une architecture du type PC avec un microprocesseur Pentium II cadencé à 400 MHz. Les temps de calcul sont environ d'une dizaine d'heures pour former une pyramide complète !

Une analyse multi-échelle de cette pyramide permet de construire de manière non supervisée les différentes classes de pixels présentes dans l'image bi-chromatique [FtP00a, FtP00d]. Malgré la lourdeur de la mise en oeuvre de cette approche, elle a été exploitée avec succès pour segmenter des images du fond de l'oeil afin d'en extraire les régions présentant des lésions rétinienne [FtP99a, FtP99b] et pour la segmentation d'images couleur d'arts décoratifs arabo-andalous [ZFB⁺01].

Cette méthode de construction des classes ne nous satisfait pas pleinement car elle ne s'applique qu'à des images bi-chromatiques pour lesquelles la couleur de chaque pixel est représentée par un couple de composantes. Il est possible alors que deux classes de pixels séparables dans l'espace couleur 3D, ne le soient pas en considérant leurs projections sur un plan couleur.

Dans ce mémoire, nous présentons une procédure moins coûteuse en temps de calcul et moins encombrante en termes de place mémoire que celle que nous avons proposée avec Michael Fontaine et qui est détaillée dans [Fon01]. Cette nouvelle procédure se divise en deux étapes successives afin de construire les classes de pixels en présence dans l'image couleur par l'analyse des trois composantes couleur des pixels.

La première sélectionne les *domaines-couleur candidats* de l'espace couleur, c'est à dire des domaines-couleur qui définissent des sous-ensembles-couleur de pixels fortement connexes et ayant des couleurs les plus homogènes possible. Les centres de ces domaines-couleur candidats sont d'abord déterminés par l'analyse des degrés de connexité et d'homogénéité des sous-ensembles-couleur associés à tous les domaines-couleur cubiques de vecteur-taille initial $(l_{init}, l_{init}, l_{init})^T$ pouvant être définis dans l'espace couleur. Les coordonnées l_{init} du vecteur-taille initial fixées par un utilisateur, correspondent à la longueur initiale des côtés des domaines. Le nombre de ces domaines-couleur cubiques se limite à $(L - (l_{init} - 1))^3$. Parmi ces $(L - (l_{init} - 1))^3$ domaines-couleur cubiques, la procédure détecte ceux qui définissent les sous-ensembles-couleur dont les degrés de connexité et d'homogénéité sont les plus élevés. Leurs centres constituent les centres des domaines-couleur candidats.

Ensuite, la taille des domaines-couleur associés à ces centres est dilatée de manière itérative afin de

déterminer les vecteurs-taille des domaines-couleur candidats définissant les sous-ensembles-couleur qui correspondent le mieux possible aux régions de l'image. Le terme dilatation utilisé ici correspond à une augmentation des coordonnées des vecteurs-taille des domaines-couleur.

La seconde étape de la procédure permet de construire les classes de pixels à partir des domaines-couleur candidats. Comme plusieurs domaines-couleur candidats peuvent définir des sous-ensembles-couleur qui correspondent aux mêmes régions, le nombre de domaines-couleur candidats peut être plus élevé que le nombre de classes de pixels en présence dans l'image. Une analyse conjointe des centres des domaines-couleur candidats et des degrés de connexité et d'homogénéité des sous-ensembles-couleur associés permet d'identifier ceux qui définissent les classes correspondant effectivement aux régions de l'image.

3.3.3 Sélection des domaines-couleur candidats

Les $N_{candidat}$ domaines-couleur candidats, notés $\underline{D}(\mathbf{c}_k, \mathbf{l}_k)$, $k = 1, \dots, N_{candidat}$ sont sélectionnés en deux temps. Les centres $\mathbf{c}_k$ des domaines-couleur candidats sont identifiés, puis la procédure détermine les vecteurs-taille $\mathbf{l}_k = (l_k, l_k, l_k)^T$ des domaines-couleur candidats de sorte que leurs sous-ensembles-couleur associés correspondent le mieux possible aux régions de l'image.

Centres des domaines-couleur candidats

Pour déterminer le nombre $N_{candidat}$ de domaines-couleur candidats ainsi que leurs centres $\mathbf{c}_k$, la procédure considère tous les domaines-couleur cubiques de vecteurs-taille identiques. L'utilisateur fixe la longueur initiale l_{init} des côtés de ces domaines-couleur cubiques, de sorte que $\mathbf{l}_{init} = (l_{init}, l_{init}, l_{init})^T$ soit leur vecteur-taille initial. En déplaçant le centre $\mathbf{c}$ à travers l'espace couleur, la procédure construit $(L - (l_{init} - 1))^3$ domaines-couleur $\underline{D}(\mathbf{c}, \mathbf{l}_{init})$ qui définissent $(L - (l_{init} - 1))^3$ sous-ensembles-couleur $S(\mathbf{c}, \mathbf{l}_{init})$ dans l'image.

Les $(L - (l_{init} - 1))^3$ degrés de connexité et d'homogénéité $DCH(S(\mathbf{c}, \mathbf{l}_{init}))$ des sous-ensembles-couleur définis par ces domaines-couleur sont organisés sous la forme d'un tableau tri-dimensionnel de valeurs discrètes. Nous proposons de sélectionner comme domaines-couleur candidats ceux qui définissent des sous-ensembles-couleur dont les degrés de connexité et d'homogénéité sont des maxima locaux par rapport à $\mathbf{c}$. En effet, plus le degré de connexité et d'homogénéité d'un sous-ensemble-couleur est élevé, plus il est vraisemblable que le sous-ensemble-couleur corresponde effectivement à une région de l'image. Le nombre de maxima locaux des degrés détectés fournit le nombre $N_{candidat}$ de domaines-couleur candidats.

L'image de la figure 3.11(b) représente les degrés de connexité et d'homogénéité des sous-ensembles-couleur $S(\mathbf{c}, [3, 3, 1]^T)$ de l'image de la figure 3.11(a) associés à tous les domaines-couleur $\underline{D}(\mathbf{c}, [3, 3, 1]^T)$ qu'il est possible de définir sur le plan couleur (R, G) . Plus le degré de connexité et d'homogénéité est proche de 1, plus l'intensité correspondant à sa représentation dans l'image de la figure 3.11(b) est élevée. En examinant la figure 3.11(b) et la figure 3.10(b) qui affiche les nuages de points-couleur de l'image de la figure 3.11 projetés sur le plan (R, G) , nous constatons que seuls les degrés de connexité et d'homogénéité des sous-ensembles-couleur associés aux domaines-couleur centrés sur des points-couleur appartenant à des nuages ne sont pas nuls.

Nous représentons dans la figure 3.11(c) les centres $\mathbf{c}_k$ projetés sur le plan couleur (R, G) des $N_{candidat} = 38$ domaines-couleur candidats sélectionnés par la procédure quand la longueur initiale l_{init} des côtés des domaines-couleur cubiques considérés est fixée à 3. Cette figure montre que tous les nuages constitués de points-couleur représentatifs des régions contiennent au moins quatre centres

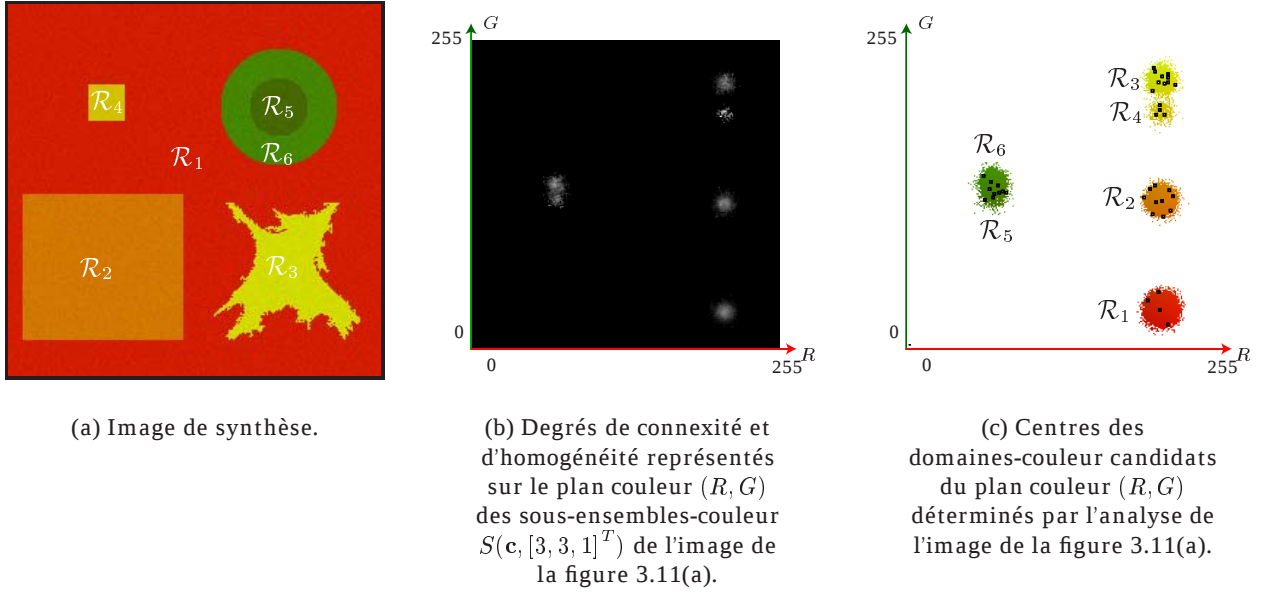


FIG. 3.11 – Centres des domaines-couleur candidats déterminés par l'analyse de l'image 3.11(a). l_{init} est fixé à 3.

déTECTÉS. Une procédure spécifique et décrite ultérieurement sélectionne les centres les plus représentatifs des régions.

Vecteurs-taille des domaines-couleur candidats

Une fois les centres identifiés, la procédure détermine les vecteurs-taille $\mathbf{l}_k$ des $N_{candidat}$ domaines-couleur candidats centrés en $\mathbf{c}_k$, $k = 1, \dots, N_{candidat}$. Ces vecteurs-taille doivent être déterminés de sorte que les domaines-couleur candidats $S(\mathbf{c}_k, \mathbf{l}_k)$ définissent des sous-ensembles-couleur correspondant le mieux possible aux régions de l'image. Par ailleurs, la détermination des vecteurs-taille doit respecter la contrainte de non-chevauchement des domaines-couleur candidats dans l'espace couleur décrite au paragraphe 3.3.1.

La procédure proposée consiste à dilater simultanément la taille des domaines-couleur cubiques centrés en $\mathbf{c}_k$, $k = 1, \dots, N_{candidat}$. Elle augmente de manière itérative, et de 2 en 2, la longueur x de leurs côtés, en partant de la valeur initiale impaire l_{init} . La longueur x des côtés des domaines examinés prend donc successivement les valeurs impaires $l_{init}, l_{init} + 2, l_{init} + 4, \dots$

A chaque itération, la procédure considère les domaines-couleur $\underline{D}(\mathbf{c}_k, \mathbf{x})$ (avec $\mathbf{x} = [x, x, x]^T$) dont les vecteurs-taille $\mathbf{l}_k$ n'ont pas été déterminés lors de l'une des itérations précédentes. La procédure examine chaque sous-ensemble-couleur $S(\mathbf{c}_k, \mathbf{x})$ associé à l'un de ces domaines-couleur. Le vecteur-taille $\mathbf{l}_k$ du domaine-couleur candidat centré en $\mathbf{c}_k$ est fixé à $\mathbf{x}$ si l'une des trois conditions suivantes est vérifiée :

- La première condition porte sur les degrés de connexité et d'homogénéité :
- $DCH(S(\mathbf{c}_k, \mathbf{x} + \mathbf{2})) \leq DCH(S(\mathbf{c}_k, \mathbf{x}))$.

Les degrés de connexité et d'homogénéité des sous-ensembles-couleur définis par les domaines-couleur cubiques centrés en $\mathbf{c}_k$ augmentent en fonction de la longueur de leurs côtés quand elle varie entre l_{init} et x . $S(\mathbf{c}_k, \mathbf{x})$ est le sous-ensemble-couleur avec le degré de connexité et d'homogénéité le plus élevé parmi tous les sous-ensembles-couleur $S(\mathbf{c}_k, \mathbf{m})$, $\mathbf{m} = [m, m, m]^T$, avec $l_{init} \leq m \leq x$. $S(\mathbf{c}_k, \mathbf{x})$ est alors considéré comme le sous-ensemble-couleur qui correspond le

mieux à l'une des régions de l'image.

– Les deux autres conditions portent sur le non-chevauchement des domaines-couleur candidats :

– $\exists j \neq k$, tel-que $\underline{D}(c_k, x+2) \cap \underline{D}(c_j, l_j) \neq \{\}$.

Cette condition signifie qu'il y a chevauchement entre le domaine-couleur $\underline{D}(c_k, x+2)$ et l'un des domaines-couleur candidats $\underline{D}(c_j, l_j)$, $j \neq k$, dont le vecteur-taille l_j a été déterminé lors de l'une des itérations précédentes.

– $\exists j \neq k$, tel-que $\underline{D}(c_k, x+2) \cap \underline{D}(c_j, x) \neq \{\}$.

Cette condition signifie qu'il y a chevauchement entre les domaines-couleur $\underline{D}(c_k, x+2)$ et $\underline{D}(c_j, x)$, avec $j \neq k$, ce dernier étant l'un des autres domaines-couleur candidats examinés à cette itération. Dans ce cas, les vecteurs-taille l_k et l_j des domaines-couleur candidats centrés en c_k et c_j sont tous les deux fixés à x .

La procédure est arrêtée quand les vecteurs-taille l_k , $k = 1, \dots, N_{candidat}$, des $N_{candidat}$ domaines-couleur candidats ont été déterminés.

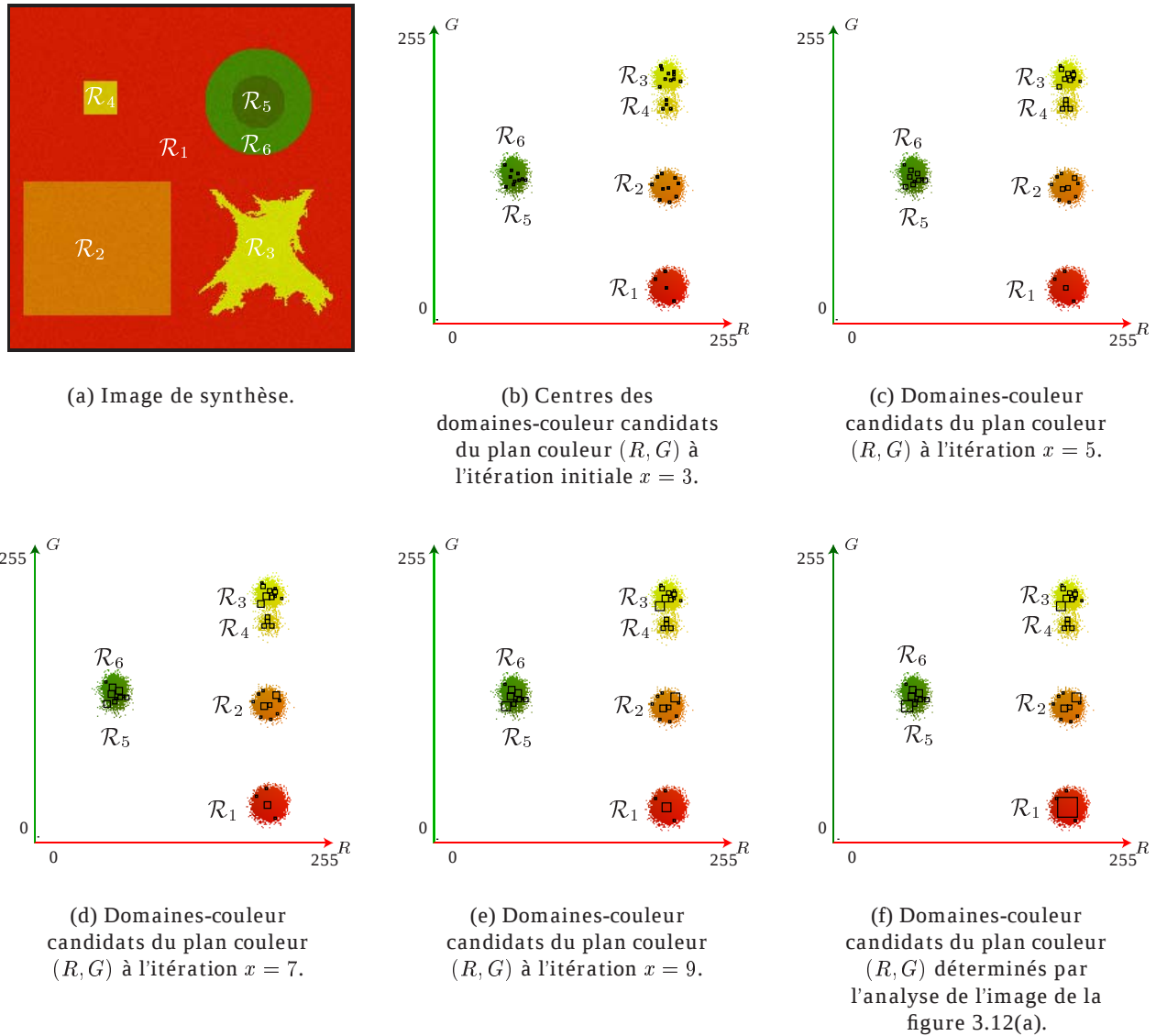


FIG. 3.12 – Détermination des domaines-couleur candidats étape par étape.

La figure 3.12(b) montre les $N_{candidat} = 38$ centres des domaines-couleur candidats qui ont été identifiés par la procédure appliquée à l'image de la figure 3.12(a). La longueur initiale l_{init} des côtés des

domaines-couleur cubiques est fixée à 3. Les figures 3.12(c), 3.12(d) and 3.12(e) montrent comment les domaines-couleur sont dilatés aux itérations correspondant respectivement aux longueurs $x = 5$, $x = 7$ et $x = 9$ de la procédure. En examinant ces figures, nous constatons que quand deux domaines-couleur se rejoignent, la procédure arrête leur dilatation. Ces figures montrent également que certains domaines-couleur ne subissent pas de dilatation, même quand aucun autre domaine-couleur ne les touche. Cette situation se produit quand les degrés de connexité et d'homogénéité des sous-ensembles-couleur associés à ces domaines-couleur n'augmentent pas en fonction de leurs dilatations.

La figure 3.12(f) représente les $N_{candidat} = 38$ domaines-couleur candidats finaux sélectionnés par la procédure. Nous remarquons que ces domaines respectent bien la contrainte de non-chevauchement.

Parmi ces domaines-couleur candidats sélectionnés, la procédure doit maintenant déterminer ceux qui définiront les classes de pixels correspondant aux six régions de l'image de la figure 3.12(a).

3.3.4 Construction des classes de pixels

Procédure itérative

Dans le contexte de la segmentation d'images couleur par classification de pixels, le problème clé réside dans l'identification des N_{class} classes de pixels, N_{class} étant le nombre de classes fixé par l'utilisateur. Le nombre $N_{candidat}$ de domaines-couleur candidats précédemment sélectionnés est probablement plus élevé que N_{class} . Par conséquent, la procédure que nous proposons doit sélectionner N_{class} domaines-couleur de définition de classes parmi les $N_{candidat}$ domaines-couleur candidats, de sorte à obtenir une correspondance bijective entre les classes de pixels et les régions de l'image.

Une solution pour sélectionner les domaines-couleur de définition de classes consiste à les trier en fonction de l'ordre décroissant des degrés de connexité et d'homogénéité des sous-ensembles-couleur associés. La figure 3.13(c) représente les 6 domaines-couleur candidats de la figure 3.13(b) associés aux sous-ensembles-couleur de l'image de la figure 3.13(a) dont les degrés de connexité et d'homogénéité sont les plus élevés. En examinant la figure 3.13(c), nous constatons qu'aucun de ces six domaines-couleur candidats ne correspond à la région $\mathcal{R}_4$ de l'image tandis que trois d'entre eux contiennent les points-couleur représentatifs des régions $\mathcal{R}_5$ et $\mathcal{R}_6$. Cet exemple montre que le tri des degrés de connexité et d'homogénéité des sous-ensembles-couleur associés aux domaines-couleur candidats ne suffit pas pour effectuer une segmentation satisfaisante de l'image.

La procédure de sélection proposée est issue de la procédure de classification "MAX-MIN" [HD86]. Elle sélectionne les N_{class} domaines-couleur de définition de classes, notés $\underline{DC}_u(c_u, \mathbf{l}_u)$, $u = 1 \dots N_{class}$, parmi les $N_{candidat}$ domaines-couleur candidats $\underline{D}(c_k, \mathbf{l}_k)$, $k = 1 \dots N_{candidat}$. Cette procédure est composée d'une étape initialisation suivie de $(N_{class} - 1)$ itérations.

- Lors de l'initialisation, la procédure identifie le premier domaine-couleur de définition de classe $\underline{DC}^1(c^1, \mathbf{l}^1)$ comme le domaine-couleur candidat associé au sous-ensemble-couleur dont le degré de connexité et d'homogénéité est le plus élevé :

$$\underline{DC}_1(c_1, \mathbf{l}_1) = \underline{D}(c, \mathbf{l}) \mid DCH(S(c, \mathbf{l})) = \max_{k=1}^{N_{candidat}} (DCH(S(c_k, \mathbf{l}_k))). \quad (3.8)$$

- A l'itération de rang u , u variant de 2 à N_{class} , la procédure retient comme u^{eme} domaine-couleur de définition de classe $\underline{DC}_u(c_u, \mathbf{l}_u)$ le domaine-couleur candidat dont le centre est l'un des plus éloignés des centres des domaines-couleur de définition de classes identifiés aux itérations précédentes et qui est associé à un sous-ensemble-couleur dont le degré de connexité et d'homogénéité est l'un des plus élevés. Pour atteindre cet objectif, le domaine-couleur candidat sélectionné maximise la fonction

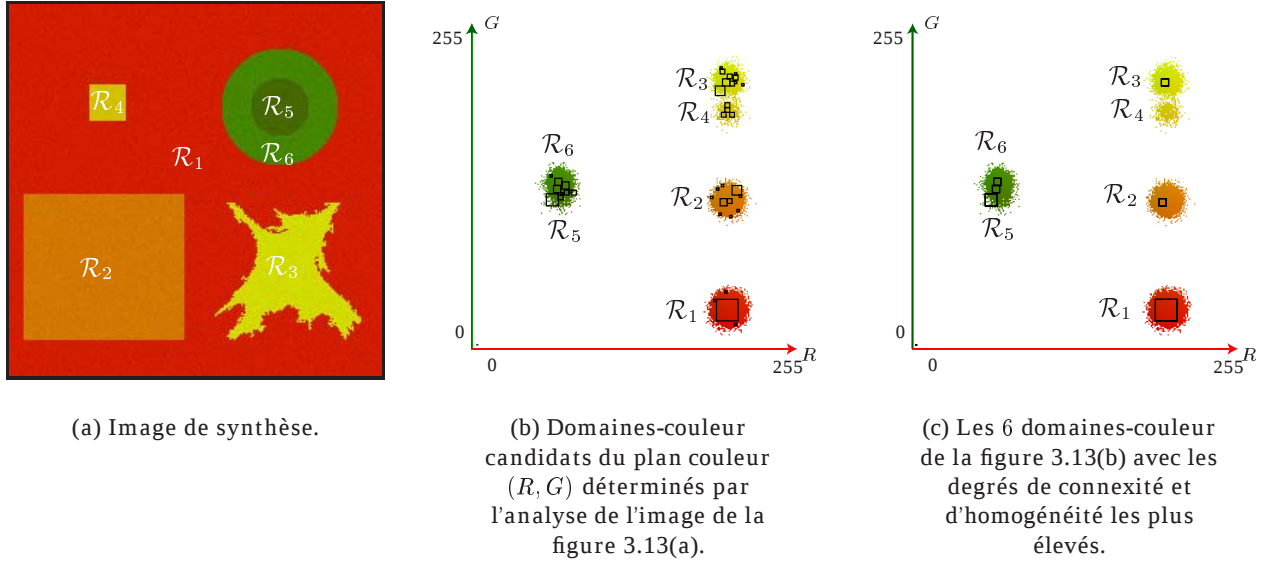


FIG. 3.13 – Sélection des domaines-couleur par le tri décroissant des degrés de connexité et d'homogénéités.

objectif J :

$$\underline{DC}_u(\mathbf{c}_u, \mathbf{l}_u) = \underline{D}(\mathbf{c}, \mathbf{l}) \mid J(u, \underline{D}(\mathbf{c}, \mathbf{l})) = \max_{k=1}^{N_{candidat}} (J(u, \underline{D}(\mathbf{c}_k, \mathbf{l}_k))), \quad (3.9)$$

avec

$$J(u, \underline{D}(\mathbf{c}_k, \mathbf{l}_k)) = \min_{v=1}^{u-1} ((\mathbf{c}_k - \mathbf{c}_v)^T \cdot (\mathbf{c}_k - \mathbf{c}_v)) \times DCH(S(\mathbf{c}_k, \mathbf{l}_k)). \quad (3.10)$$

A l'itération de rang u , la procédure évalue la fonction objectif $J(u, D(\mathbf{c}_k, \mathbf{l}_k))$ de chaque domaine-couleur candidat examiné $\underline{D}(\mathbf{c}_k, \mathbf{l}_k)$. Cette fonction est le produit de deux termes.

Le premier est le minimum des carrés des distances euclidiennes séparant le centre $\mathbf{c}_k$ du domaine-couleur candidat $\underline{D}(\mathbf{c}_k, \mathbf{l}_k)$ des centres $\mathbf{c}_v$, $v = 1, \dots, u - 1$, des $(u - 1)$ domaines-couleur de définition de classes qui ont été identifiés lors des itérations précédentes (voir équation (3.10)). Ce terme pénalise les domaines-couleur candidats dont les centres sont proches des centres des domaines-couleur de définition de classes. Nous avons vu qu'une région peut être représentée par plusieurs domaines-couleur candidats dont les centres sont proches dans l'espace couleur. Dans ce cas, quand un de ces domaines-couleur candidats est sélectionné pour définir une classe correspondant à une région, les autres domaines-couleur candidats qui représentent également cette région ont peu de chance d'être sélectionnés car ils ne maximisent pas la distance entre leurs centres et le centre de ce domaine-couleur de définition de classe.

Le second terme est un facteur de pondération constitué du degré de connexité et d'homogénéité du sous-ensemble-couleur $S(\mathbf{c}_k, \mathbf{l}_k)$ associé au domaine-couleur candidat examiné $\underline{D}(\mathbf{c}_k, \mathbf{l}_k)$. Le domaine-couleur candidat tend à maximiser ce poids s'il est associé à un sous-ensemble-couleur dont le degré de connexité et d'homogénéité est proche de 1.

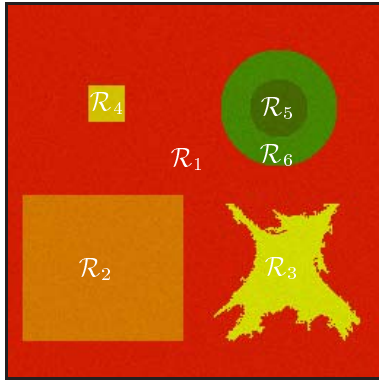
Grâce à cette fonction objectif, la procédure sélectionne à chaque rang d'itération le domaine-couleur candidat dont le centre est l'un des plus éloignés des centres des domaines-couleur de définition de classes identifiés lors des itérations précédentes et qui est associé au sous-ensemble-couleur dont le degré de connexité et d'homogénéité est l'un des plus élevés.

Résultats avec l'image de synthèse

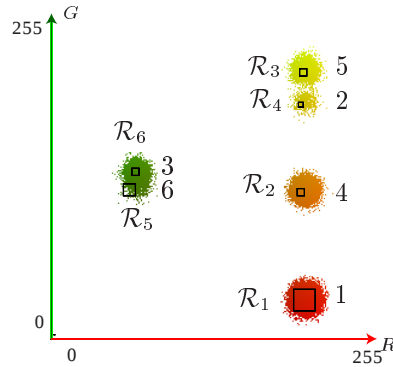
Nous montrons sur la figure 3.14(b) les 6 domaines-couleur de définition de classes qui ont été identifiées par la procédure appliquée à l'image de la figure 3.14(a), parmi les 38 domaines-couleur candidats représentés par la figure 3.13(b). Cette figure montre que chaque domaine-couleur sélectionné contient les points-couleur provenant d'au moins une région. Dans cette figure, le numéro proche du domaine indique le rang de l'itération à laquelle il a été identifié. Le premier domaine-couleur de définition de classe correspond à la région $\mathcal{R}_1$. Puis, le domaine-couleur de définition de classe correspondant à la région $\mathcal{R}_4$ est sélectionné lors de la seconde itération, bien que son centre ne soit pas le plus éloigné de celui du premier domaine-couleur de définition de classe. Ce domaine-couleur candidat est sélectionné car le degré de connexité et d'homogénéité du sous-ensemble-couleur associé est l'un des plus élevés.

Par ailleurs, nous constatons sur la figure 3.12(f) que le nuage formé par les points-couleur représentant la région $\mathcal{R}_3$ contient 10 domaines-couleur candidats avec différents vecteurs-taille. La figure 3.14(b) montre que le domaine-couleur candidat sélectionné pour représenter la région $\mathcal{R}_3$ n'est pas le plus grand. En effet, le degré de connexité et d'homogénéité du sous-ensemble-couleur associé au domaine sélectionné par la procédure est le plus élevé parmi ceux des sous-ensembles-couleur associés aux domaines-couleur candidats représentant la région $\mathcal{R}_3$.

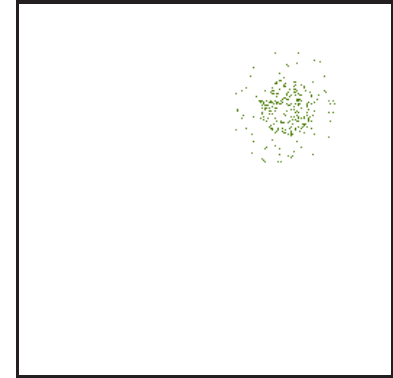
Nous constatons également sur la figure 3.14(b) que les régions $\mathcal{R}_5$ et $\mathcal{R}_6$ dont les points-couleur forment un seul nuage, sont représentées par deux domaines-couleur de définition de classes. Pour vérifier leur pertinence, nous avons affiché dans les images des figures 3.14(c) et 3.14(d) les pixels qui appartiennent aux sous-ensembles-couleur associés aux deux domaines-couleur de définition de classes correspondant respectivement aux régions $\mathcal{R}_5$ et $\mathcal{R}_6$. Quelques pixels du sous-ensemble-couleur correspondant à la région $\mathcal{R}_5$ ($\mathcal{R}_6$, respectivement) appartiennent à l'autre région $\mathcal{R}_6$ ($\mathcal{R}_5$, respectivement). Ces erreurs de classification sont provoquées par le fort chevauchement des distributions des points-couleur représentatifs de ces deux régions.



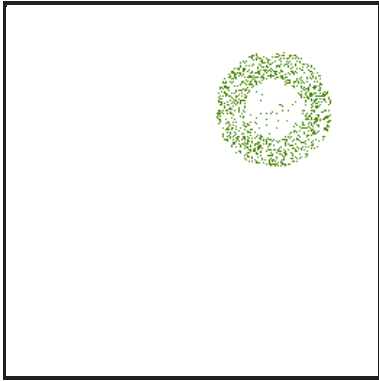
(a) Image de synthèse.



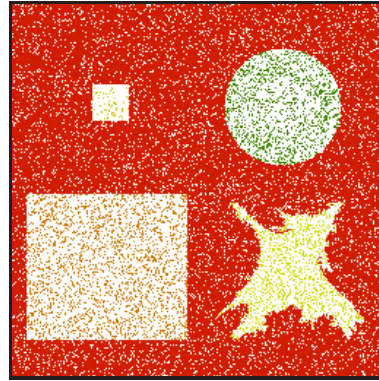
(b) Domaines-couleur de définition de classes de l'image de la figure 3.14(a) identifiés par la procédure. Le numéro proche de chaque domaine correspond au rang d'itération à laquelle il a été identifié.



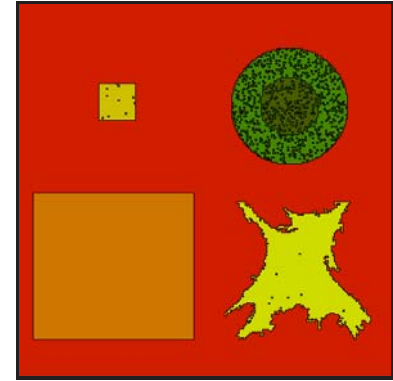
(c) Sous-ensemble couleur de l'image de la figure 3.14(a) associé au domaine de définition de classe correspondant à la région $\mathcal{R}_5$. Les pixels étiquetés en blanc n'appartiennent pas à ce sous-ensemble-couleur.



(d) Sous-ensemble couleur de l'image de la figure 3.14(a) associé au domaine de définition de classe correspondant à la région $\mathcal{R}_6$. Les pixels étiquetés en blanc n'appartiennent pas à ce sous-ensemble-couleur.



(e) Sous-ensembles couleur de l'image de la figure 3.14(a) associés aux domaines-couleur de définition de classes de la figure 3.14(b). Les pixels étiquetés en blanc n'appartiennent pas à ces sous-ensembles-couleur.



(f) Image couleur segmentée à partir de l'analyse de l'image de la figure 3.14(a). Les frontières des régions reconstruites sont affichées en noir. Les pixels n'appartenant à aucun sous-ensemble-couleur de la figure 3.14(e) sont assignés aux classes selon une règle de décision spécifique.

FIG. 3.14 – Construction des classes de pixels et segmentation de l'image.

Nous représentons dans l'image de la figure 3.14(e) les sous-ensembles-couleur associés aux 6 domaines-couleur de définition de classes de l'image de la figure 3.14(a). Pour la segmentation, les pixels qui appartiennent à chacun de ces sous-ensembles-couleur sont considérés comme constituant les noyaux des classes correspondantes. Chaque pixel qui n'appartient à aucun de ces noyaux est assigné à la classe dont la frontière du domaine-couleur de définition est la plus proche du point-couleur le représentant dans l'espace couleur. Nous pouvons alors reconstruire les régions qui sont composées de pixels connexes assignés aux mêmes classes.

L'image d'étiquettes de la figure 3.14(f) montre comment sont reconstruites les régions de l'image de la figure 3.14(a) à partir des domaines-couleur de définition de classes représentés par la figure 3.14(b). Dans l'image de la figure 3.14(f), les frontières entre les régions reconstruites sont affichées en noir. Cette image montre que les 6 régions sont bien reconstruites. Le chevauchement des distributions des points-couleur provenant des régions $\mathcal{R}_3$ et $\mathcal{R}_4$ explique pourquoi quelques pixels épars de la région $\mathcal{R}_3$ sont mal classés. Malgré le fort chevauchement des distributions des points-couleur représentatifs des deux régions $\mathcal{R}_5$ et $\mathcal{R}_6$, la plupart des pixels de ces deux régions sont correctement classés. Ce résultat montre que notre méthode est capable de discriminer des classes de pixels non équiprobables dont les distributions des points-couleur se chevauchent dans l'espace couleur.

Discussion

Nous montrons dans la figure 3.15(a) les domaines-couleur de définition de classes identifiés interactivement par un utilisateur de sorte que les sous-ensembles-couleur associés contiennent exclusivement un grand nombre de pixels appartenant aux régions correspondantes dans l'image de la figure 3.14(a) (cette figure est identique à la figure 3.10(b)). En comparant cette figure à la figure 3.15(b) qui représente les domaines-couleur de définition de classes identifiés par la procédure, nous constatons que ces domaines-couleur sont plus petits que ceux interactivement sélectionnés par l'utilisateur. Ceci est dû au choix de la longueur initiale l_{init} des côtés des domaines-couleur. En effet, quand l_{init} est fixée à 3, la procédure sélectionne de nombreux domaines-couleur candidats dont les centres sont proches les uns des autres dans l'espace couleur (voir figure 3.12(b)). Par conséquent, la procédure qui détermine les vecteurs-taille l_k ne peut pas dilater les domaines-couleur candidats afin d'éviter d'éventuels chevauchements.

La figure 3.15(c) représente les domaines-couleur de définition de classes identifiés par la procédure quand la longueur initiale l_{init} des côtés est fixée à 5. La plupart des domaines-couleur identifiés par la procédure sont presque identiques à ceux interactivement sélectionnés par l'utilisateur.

Le tableau 3.3.4 indique les taux moyens d'erreur de classification des pixels obtenus par notre méthode appliquée à l'image de la figure 3.14(a) dont l'unique paramètre ajustable l_{min} varie de 3 à 9. Ce tableau indique que la meilleure segmentation est obtenue quand l_{init} est fixée à 5. L'augmentation du taux moyen d'erreur de classification quand l_{init} est fixée à 7 est provoquée par une plus mauvaise classification des pixels des régions $\mathcal{R}_5$ et $\mathcal{R}_6$ que lorsque l_{init} est fixée à 5. Enfin, les résultats se détériorent encore plus quand l_{init} est fixée à 9, car les pixels de la région $\mathcal{R}_3$ et $\mathcal{R}_4$ sont assignés à une seule classe construite. Nous déduisons de cet exemple que la valeur du paramètre l_{init} fixée à 3 ou à 5, permet de sélectionner des petits domaines-couleur candidats qui définissent des sous-populations de points-couleur parmi l'ensemble de tous ceux qui donnent naissance à un nuage unique identifié dans l'espace couleur.

Notre méthode de segmentation a été implantée sur une architecture du type PC avec un microprocesseur Pentium II cadencé à 400 MHz. Les temps de calcul dépendent essentiellement de trois paramètres, à savoir :

l_{min}	taux moyen d'erreur de classification	temps de calcul
3	1,49	5'13
5	1,27	5'45
7	2,19	6'05
9	2,85	6'36

TAB. 3.8 – Taux moyens d'erreur de classification et temps de calcul nécessaires à la segmentation de l'image de la figure 3.14(a).

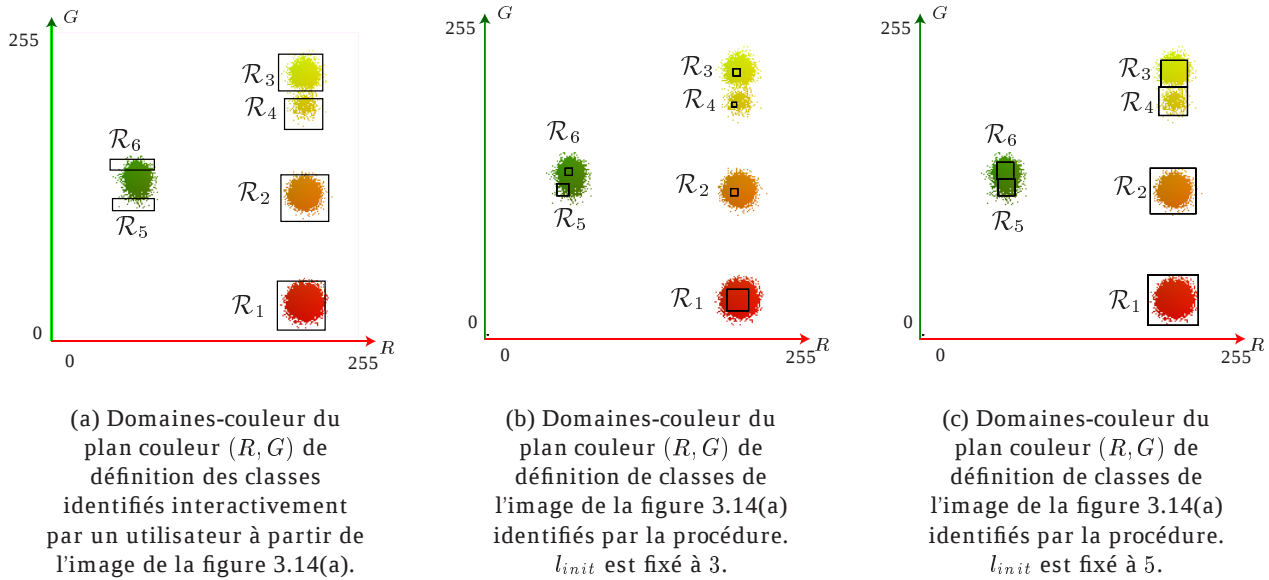


FIG. 3.15 – Comparaison entre les domaines-couleur de définition de classes identifiés interactivement par un utilisateur et ceux identifiés par la procédure.

- la taille de l'image à analyser,
- le nombre de couleurs présentes dans l'image qui fournit le nombre de domaines-couleur à examiner,
- la longueur l_{min} initiale des côtés des domaines-couleur examinés.

Nous présentons dans le tableau 3.3.4 les temps de calcul nécessaires à la segmentation de l'image de la figure 3.14(a) de taille 256×256 pixels en fonction de l_{min} .

Ce tableau montre que la procédure de construction des classes prend d'autant plus de temps que la longueur initiale des côtés des domaines-couleur est élevée. Il montre également que les temps de calcul nécessaires à la procédure présentée sont nettement inférieurs à ceux requis par la procédure d'analyse des pyramides des degrés de connexité que nous avons mise en place avec Michael Fontaine [Fon01].

3.4 Résultats

La procédure de construction des classes présentée dans ce chapitre est basée sur l'analyse conjointe des propriétés de connexité et d'homogénéité colorimétrique de sous-ensembles-couleur de pixels. Afin de démontrer l'intérêt de cette nouvelle approche, nous proposons de comparer les résultats de segmentation d'images obtenus par notre procédure avec ceux obtenus par l'algorithme de classification des pixels basé sur un apprentissage compétitif et proposé par Uchiyama [UA94]. Cette méthode, que nous dénommons apprentissage compétitif, est reconnue comme l'une des méthodes de classification

de pixels fournissant des résultats de segmentation satisfaisants par l'analyse privilégiée de la distribution des points-couleur représentant les pixels dans l'espace couleur [CJSW01].

Ces deux algorithmes requièrent la connaissance a priori du nombre de classes N_{class} . Les valeurs des autres paramètres de la méthode basée sur un apprentissage compétitif ont été proposées par l'auteur pour segmenter la plupart des images couleur [UA94]. Suite aux conclusions déduites des résultats sur l'image de synthèse, la valeur de l'unique paramètre l_{init} requis par notre approche est fixée à 5.

3.4.1 Les images

Nous proposons de tester les deux procédures en segmentant trois images de référence, l'image 'jelly beans' (voir figure 3.16(a)), l'image 'woman' (voir figure 3.17(a)) et l'image 'hand' (voir figure 3.18(a)) par les deux procédures à tester. Comme nous n'abordons pas, dans ce chapitre, l'influence des espaces de représentation de la couleur sur les résultats de segmentation, les couleurs des pixels sont simplement représentées dans l'espace couleur (R, G, B) .

L'image 'jelly beans' de la figure 3.16(a) contient principalement cinq classes de pixels qui correspondent au fond clair de l'image et aux bonbons de couleurs rouge, jaune, noire et verte. Comme l'image contient autant de bonbons de chaque couleur, les quatre classes des pixels correspondant aux bonbons sont sensiblement équiprobables. La reconstruction des régions est difficile car elle doit tenir compte des effets provoqués par des ombres et des reflets présents dans l'image. La figure 3.16(b) montre que les points-couleur représentatifs des pixels de l'image de la figure 3.16(a) ne forment pas des nuages de points bien séparés dans l'espace couleur.

Les pixels de l'image 'woman' peuvent être divisés en quatre classes, à savoir le fond de l'image, la peau, le vêtement bleu et les cheveux noirs (voir figure 3.17(a)). Une segmentation correcte de cette image est difficile du fait que les points-couleur représentant les pixels de l'image constituent de nombreux nuages dispersés dans l'espace couleur (voir figure 3.17(b)).

Dans l'image 'hand', nous pouvons discerner quatre classes de pixels qui représentent le fond de l'image, la main, les ongles et l'anneau bleu (voir figure 3.18(a)). Par ailleurs, il n'est pas facile de segmenter correctement cette image car les points-couleur représentant les pixels donnent naissance à des nuages se chevauchant dans l'espace couleur (voir figure 3.18(b)).

Une comparaison des résultats de segmentation va fournir un aperçu des comportements des deux procédures que nous proposons d'exécuter avec un nombre de classes N_{class} variant de quatre à sept. Les images des figures 3.19, 3.20 et 3.21 montrent respectivement les résultats obtenus par l'apprentissage compétitif et l'analyse des degrés de connexité et d'homogénéité appliqués aux images 'jelly beans', 'woman', et 'hand'.

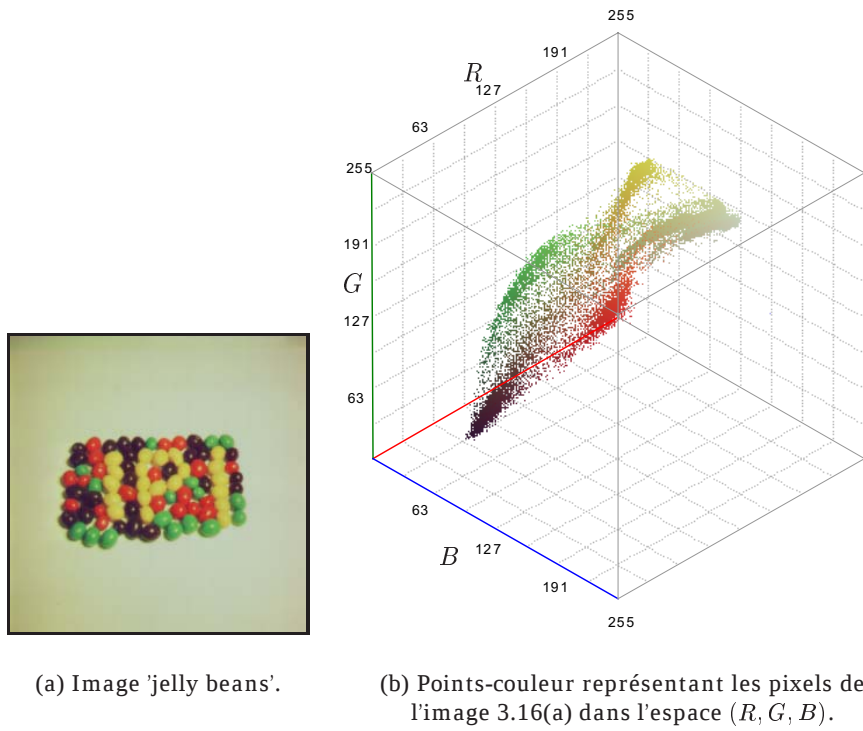


FIG. 3.16 – Image 'jelly beans'.

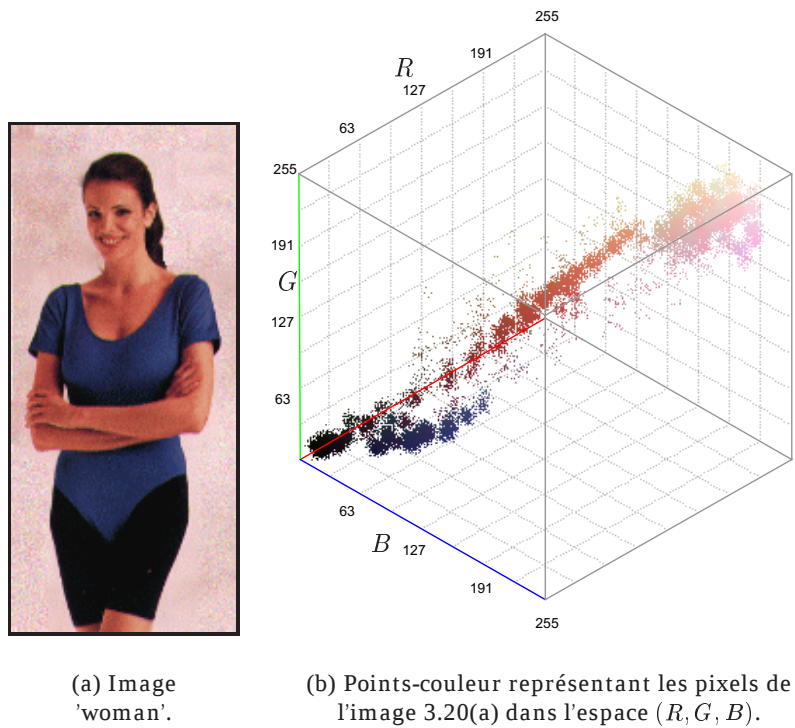


FIG. 3.17 – Image 'woman'.



(a) Image 'hand'.

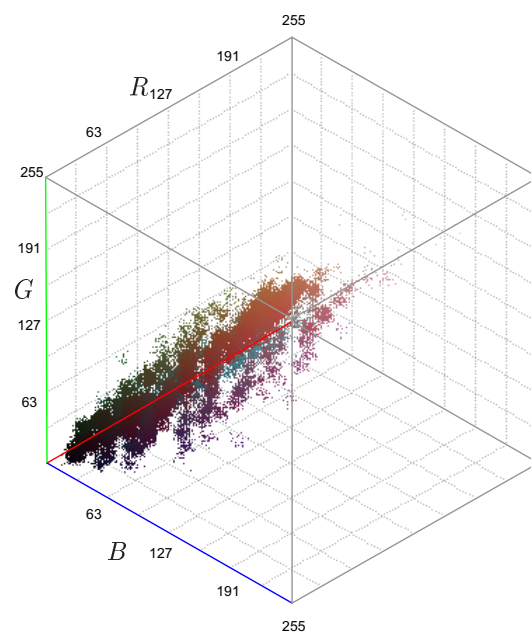
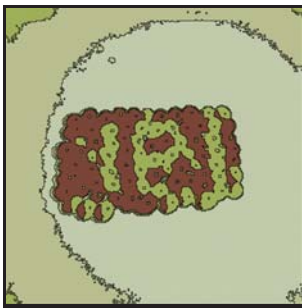
(b) Points-couleur représentant les pixels de l'image 3.21(a) dans l'espace (R, G, B) .

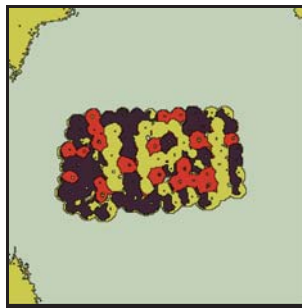
FIG. 3.18 – Image 'hand'.



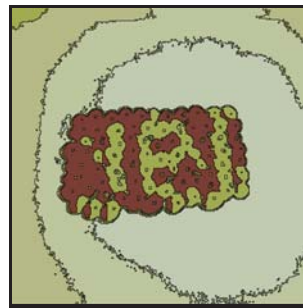
(a) Image 'jelly beans'.



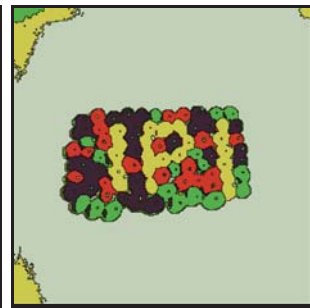
(b) Apprentissage
compétitif.
 $N_{Class} = 4.$



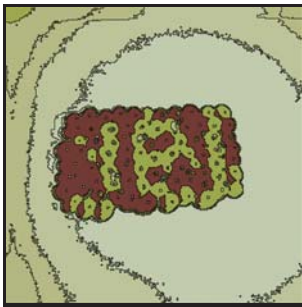
(c) Analyse des degrés
de connexité et
d'homogénéité.
 $N_{Class} = 4.$



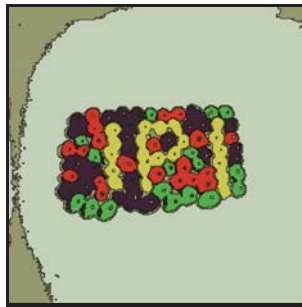
(d) Apprentissage
compétitif.
 $N_{Class} = 5.$



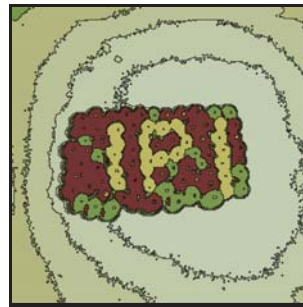
(e) Analyse des degrés
de connexité et
d'homogénéité.
 $N_{Class} = 5.$



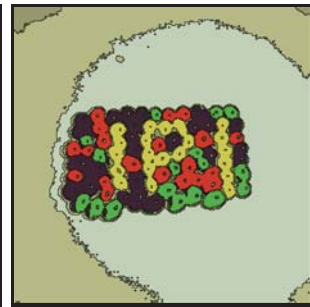
(f) Apprentissage
compétitif.
 $N_{Class} = 6.$



(g) Analyse des degrés
de connexité et
d'homogénéité.
 $N_{Class} = 6.$



(h) Apprentissage
compétitif.
 $N_{Class} = 7.$



(i) Analyse des degrés
de connexité et
d'homogénéité.
 $N_{Class} = 7.$

FIG. 3.19 – Segmentation de l'image 'jelly beans'.

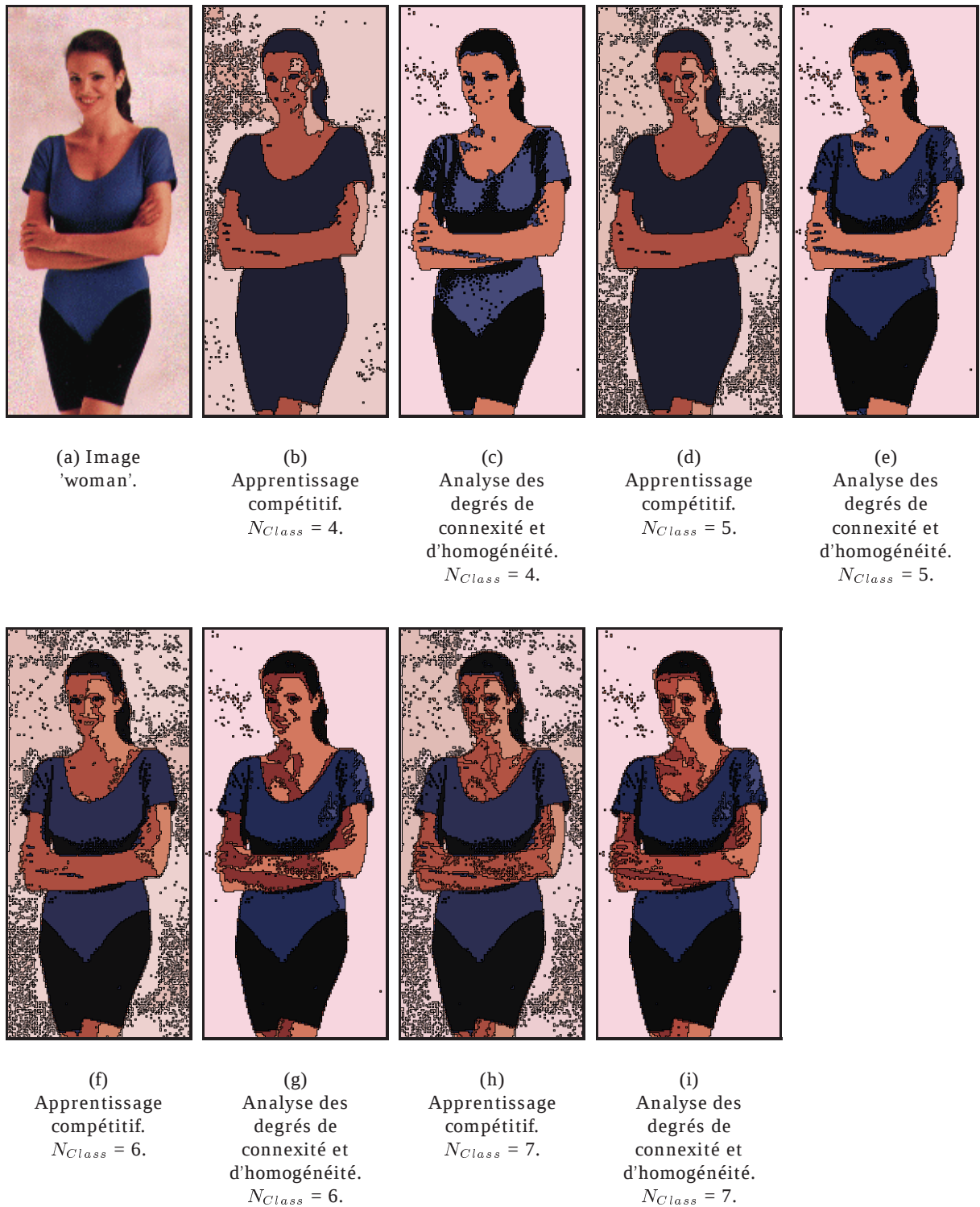
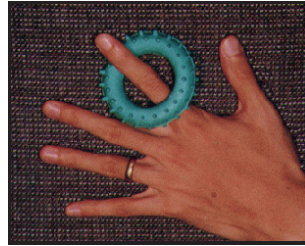


FIG. 3.20 – Segmentation de l'image 'woman'.



(a) Image 'hand'.



(b)
Apprentissage
compétitif.
 $N_{Class} = 4$.



(c) Analyse des degrés
de connexité et
d'homogénéité.
 $N_{Class} = 4$.



(d)
Apprentissage
compétitif.
 $N_{Class} = 5$.



(e) Analyse des degrés
de connexité et
d'homogénéité.
 $N_{Class} = 5$.



(f)
Apprentissage
compétitif.
 $N_{Class} = 6$.



(g) Analyse des degrés
de connexité et
d'homogénéité.
 $N_{Class} = 6$.



(h)
Apprentissage
compétitif.
 $N_{Class} = 7$.



(i) Analyse des degrés
de connexité et
d'homogénéité.
 $N_{Class} = 7$.

FIG. 3.21 – Segmentation de l'image 'hand'.

3.4.2 Évaluation qualitative de la qualité de segmentation

En examinant visuellement la segmentation de l'image 'jelly beans', nous constatons que les deux procédures mettent trois classes de bonbons en évidence quand N_{class} est fixé à quatre (voir figures 3.19(b) et 3.19(c)). Lorsque N_{class} est fixé à cinq, l'apprentissage compétitif ne parvient pas à discriminer les cinq classes de pixels en présence dans l'image tandis que l'analyse des degrés de connexité et d'homogénéité réussit à les distinguer (voir figures 3.19(d) et 3.19(e)). L'intensité des pixels du fond situés dans la partie gauche de l'image est supérieure à celle des pixels situés dans la partie droite. Cette non uniformité de l'intensité et la taille importante de la population des pixels formant le fond de l'image expliquent que l'apprentissage compétitif divise ces pixels en deux classes distinctes. Par contre, comme les degrés de connexité et d'homogénéité sont indépendants des tailles des populations des sous-ensembles-couleur analysés, la procédure proposée parvient à regrouper la plupart des pixels du fond de l'image en une seule classe.

Quand N_{class} est supérieur à cinq, l'apprentissage compétitif sépare les pixels du fond de l'image en plusieurs régions tandis que l'analyse des degrés de connexité et d'homogénéité met en valeur certains détails des bonbons (voir figures 3.19(f)... 3.19(i)), ce qui semble plus acceptable en termes de pertinence de segmentation.

Examinons visuellement les résultats de la segmentation de l'image 'woman' (voir figure 3.20). Quand N_{class} varie de quatre à sept, l'apprentissage compétitif reconstruit le fond de l'image en plusieurs petites régions tandis que l'analyse des degrés de connexité et d'homogénéité parvient à représenter la quasi-totalité de ce fond par une seule région principale. Ces bons résultats sont dus à la fonction objectif J utilisée pour la construction des classes de pixels (voir équation (3.10)), qui tient compte du degré de connexité et d'homogénéité des sous-ensembles-couleur examinés. Le degré de connexité et d'homogénéité du sous-ensemble-couleur représentant le fond de l'image est si élevé que le domaine-couleur candidat associé est sélectionné sans ambiguïté par la procédure pour définir une classe de pixels. En examinant les autres régions reconstruites, il apparaît que l'analyse des degrés de connexité et d'homogénéité fournit une partition plus pertinente en régions représentant la femme que l'apprentissage compétitif.

Nous débouchons sur les mêmes conclusions concernant les résultats de segmentation de l'image 'hand' (voir figure 3.21). Tandis que l'apprentissage compétitif tend à diviser les pixels représentant la main et le fond de l'image en plusieurs classes quand le nombre de classes augmente, l'analyse des degrés de connexité et d'homogénéité construit des domaines-couleur de définition de classes de sorte que la main soit toujours représentée par une région principale. Seul l'anneau bleu est décomposé en plusieurs petites régions. Par ailleurs, les ongles de la main sont correctement reconstruits par l'analyse des degrés de connexité et d'homogénéité, tandis qu'ils sont ignorés par l'apprentissage compétitif. Quand le nombre désiré de classes est plus élevé que le nombre de classes de pixels en présence dans l'image, l'analyse des degrés de connexité et d'homogénéité n'a pas tendance à décomposer les grands objets en plusieurs petites régions.

3.4.3 Évaluation quantitative de la qualité de segmentation

Comme il est difficile d'estimer visuellement la qualité de résultats d'une procédure de segmentation, nous proposons d'utiliser des critères quantitatifs [Zha96]. En l'absence d'une "vérité-terrain" (segmentation de référence), il existe assez peu de critères quantitatifs permettant de comparer des résultats de segmentation [PF01].

Le premier critère utilisé, noté R , est le nombre de régions reconstruites constituant l'image segmen-

tée. Le second critère, noté E^2 , est la moyenne des carrés des distances euclidiennes séparant les points-couleur représentant les pixels des centres des classes auxquelles ils ont été assignés. Comme ce critère, nommé écart quadratique, reflète la compacité des classes, il constitue l'un des critères les plus utilisés pour évaluer la qualité des résultats des procédures de classification. Cependant, ce critère souffre d'un gros inconvénient pour l'évaluation de la qualité des résultats de segmentation car il ne tient pas compte des propriétés spatiales des régions reconstruites.

Aussi, nous proposons d'utiliser également le critère proposé par Borsotti, noté Q , qui est spécifiquement conçu pour l'évaluation de la qualité des résultats de segmentation d'images couleur et qui s'exprime de la manière suivante [BCS98] :

$$Q = \frac{\sqrt{R}}{X \times Y} \times \sum_{i=1}^R \left(\frac{\Delta_i^2}{1 + \log A_i} + \frac{R(A_i)^2}{A_i^2} \right) \quad (3.11)$$

où

- $X \times Y$ indique la taille de l'image,
- R correspond au nombre de régions reconstruites de l'image segmentée,
- A_i est la surface de la i^{eme} région reconstruite $\mathcal{R}_i$,
- $R(A_i)$ est le nombre de régions reconstruites de surface A_i ,
- $Max(A_i)$ est la surface de la plus grande région reconstruite de l'image segmentée,
- Δ_i^2 est la somme des carrés des distances euclidiennes séparant le point-couleur représentant chaque pixel de $\mathcal{R}_i$ du centre des points-couleur représentatifs des pixels de $\mathcal{R}_i$. Cette valeur est également d'autant plus élevée que les régions sont petites.

Plus la valeur du critère Q est faible et plus le résultat de la segmentation est considéré comme étant satisfaisant.

Ce critère est le produit de trois termes :

- $\frac{1}{X \times Y}$ est un coefficient de normalisation nécessaire pour s'affranchir de la taille de l'image analysée,
- $\sqrt{R}$ permet de pénaliser les images sur-segmentées,
- la somme se décompose en deux termes. Le premier prend une valeur élevée soit lorsque les régions reconstruites sont constituées de pixels dont les couleurs ne sont pas homogènes, soit lorsque le nombre de petites régions reconstruites est élevé. Le deuxième prend une valeur élevée lorsqu'il y a beaucoup de régions reconstruites de mêmes surfaces. Il est d'autre part d'autant plus élevé que les régions reconstruites sont petites, comme c'est le cas dans une image sur-segmentée.

Les tableaux 3.9, 3.10 et 3.11 indiquent respectivement les valeurs de ces trois critères obtenues par les deux méthodes à comparer, quand le nombre de classes N_{class} varie de quatre à sept. Ces tableaux montrent que lorsque N_{class} augmente, le nombre R de régions reconstruites a tendance à augmenter également, tandis que l'écart quadratique E^2 diminue.

Pour tous les résultats de segmentation d'images, l'analyse des degrés de connexité et d'homogénéité reconstruit un nombre R de régions plus faible que celui fourni par l'apprentissage compétitif. Ceci signifierait que l'analyse des degrés de connexité et d'homogénéité produit une segmentation plus grossière que l'apprentissage compétitif. Par ailleurs, la valeur de l'écart quadratique E^2 obtenue avec l'apprentissage compétitif est généralement plus faible que celle évaluée avec l'analyse des degrés de connexité et d'homogénéité. En effet, l'apprentissage compétitif est développé pour minimiser la valeur de l'écart quadratique sans tenir compte des propriétés spatiales des pixels. L'examen des valeurs prises par ces deux premiers critères indique que l'apprentissage compétitif fournit de meilleurs résultats de segmentation que l'analyse des degrés de connexité et d'homogénéité.

Par contre, les tableaux 3.9, 3.10 et 3.11 montrent que les valeurs du critère Q obtenues par l'analyse

Critère/méthode	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité
N_{Class}	4	4	5	5
Image	Fig. 3.19(b)	Fig. 3.19(c)	Fig. 3.19(d)	Fig. 3.19(e)
$R(I)$	324	279	531	401
E^2	818	1449	761	1041
Q	1727	1597	6162	1682
Critère/méthode	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité
N_{Class}	6	6	7	7
Image	Fig. 3.19(f)	Fig. 3.19(g)	Fig. 3.19(h)	Fig. 3.19(i)
R	752	348	844	470
E^2	752	868	587	443
Q	25415	1342	30160	3180

TAB. 3.9 – Evaluation de la qualité de segmentation de l'image 'jelly beans' de la figure 3.19(a).

des degrés de connexité et d'homogénéité sont toujours plus faibles que celles évaluées avec l'apprentissage compétitif. En se basant sur les valeurs du critère Q , spécifiquement proposé pour l'évaluation de la qualité des résultats de segmentation, nous concluons que l'analyse des degrés de connexité et d'homogénéité fournit de meilleurs résultats de segmentation pour les trois images, et ce pour tous les nombres de classes considérés.

Ces exemples montrent qu'en l'absence de vérité-terrain, les conclusions sur l'efficacité d'une méthode de segmentation varient selon les critères d'évaluation retenus. Si nous désirons minimiser l'écart quadratique, alors notre choix porte sur l'apprentissage compétitif. Par contre, si nous préférons obtenir un nombre restreint de régions de grande surface, alors notre choix privilégie l'analyse des degrés de connexité et d'homogénéité.

3.5 Conclusion

Dans ce chapitre, nous avons présenté une nouvelle méthode de segmentation d'images couleur qui reconstruit les régions par classification non supervisée des pixels. Cette méthode a été développée pour tenter de répondre au cas qui n'a rien d'exceptionnel, où il n'y a pas de correspondance bijective entre les nuages de points-couleur dans l'espace couleur et les régions de l'image. Quand les points-couleur provenant de différentes régions de l'image constituent un seul nuage dans l'espace couleur, la méthode proposée divise ce nuage en sous-populations de points-couleur associées à des domaines-couleur. Toute la difficulté consiste à retrouver les domaines-couleur qui définissent des classes de pixels correspondant effectivement aux régions de l'image.

Pour ce faire, la procédure analyse conjointement les propriétés de connexité et d'homogénéité colorimétrique des sous-ensembles-couleur de pixels associés aux domaines-couleur considérés.

La sélection des domaines-couleur pertinents s'appuie sur une nouvelle mesure, le degré de connexité et d'homogénéité, qui indique le degré de confiance que nous pouvons accorder à la proposition : "le sous-ensemble-couleur associé au domaine-couleur examiné correspond à une région de l'image".

Une procédure spécifique est développée afin de sélectionner les domaines-couleur de définition de

Critère/méthode	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité
N_{Class}	4	4	5	5
Image	Fig. 3.20(b)	Fig. 3.20(c)	Fig. 3.20(d)	Fig. 3.20(e)
R	448	386	808	551
E^2	1135	2077	1005	1682
Q	17388	9967	117667	23066
Critère/méthode	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité
N_{Class}	6	6	7	7
Image	Fig. 3.20(f)	Fig. 3.20(g)	Fig. 3.20(h)	Fig. 3.20(i)
R	944	528	1081	630
E^2	508	1290	484	1114
Q	182697	29337	285208	44478

TAB. 3.10 – Evaluation de la qualité de segmentation de l'image 'woman' de la figure 3.20(a).

Critère/méthode	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité
N_{Class}	4	4	5	5
Image	Fig. 3.21(b)	Fig. 3.21(c)	Fig. 3.21(d)	Fig. 3.21(e)
R	4755	3849	4365	3948
E^2	857	2164	564	2147
Q	17935255	3660819	8114611	4426336
Critère/méthode	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité	Apprentissage compétitif	Analyse des degrés de connexité et d'homogénéité
N_{Class}	6	6	7	7
Image	Fig. 3.21(f)	Fig. 3.21(g)	Fig. 3.21(h)	Fig. 3.21(i)
R	5030	3986	8186	2355
E^2	484	2088	455	1210
Q	12883146	4676083	54310340	1958540

TAB. 3.11 – Evaluation de la qualité de segmentation de l'image 'hand' de la figure 3.21(a).

classes, de sorte que les classes correspondent effectivement aux régions de l'image. La procédure proposée nécessite l'ajustement d'un seul paramètre, à savoir la longueur initiale l_{init} des côtés des domaines-couleur examinés. La finesse de la segmentation des images dépend de ce paramètre. Les domaines-couleur de définition de classes dans l'espace couleur sont d'autant plus grands que la longueur initiale des côtés des domaines examinés est élevée. Plus les tailles des domaines-couleur de définition de classes sont élevées dans l'espace couleur, plus le nombre de pixels constituant les noyaux des classes est élevé, et par conséquent, plus les régions reconstruites tendent à être étendues.

Nous avons présenté les détails de la méthode proposée en utilisant une image de synthèse. Cet exemple a permis de montrer que notre méthode est capable de discriminer des classes de pixels non équiprobables et qu'elle reste performante même lorsque les distributions des points-couleur présentent de forts chevauchements dans l'espace couleur.

Par ailleurs, nous avons testé notre approche sur des images réelles et comparé les résultats obtenus par notre méthode avec ceux obtenus par un apprentissage compétitif qui privilégie l'analyse de la distribution des points-couleur représentant les pixels dans l'espace couleur. Ces expériences ont mis en évidence l'intérêt de l'analyse des degrés de connexité et d'homogénéité pour segmenter des images couleur par classification de pixels.

Il serait intéressant de poursuivre ce travail en tentant d'améliorer la méthode de segmentation proposée sur les deux points suivants : la construction des classes de pixels et la reconstruction des régions à partir des classes construites.

Pour des raisons de complexité algorithmique, la procédure proposée de construction des classes de pixels analyse seulement des domaines-couleur cubiques. Elle suppose que les points couleur représentatifs des régions peuvent être rassemblés au sein de domaines-couleur cubiques ne se chevauchant pas dans l'espace couleur. Il serait intéressant de relâcher cette contrainte en considérant des domaines couleur non cubiques pour sélectionner ceux qui représentent les classes de pixels en présence dans l'image. Cependant, le relâchement de cette contrainte s'accompagnerait irrémédiablement d'une augmentation de la complexité calculatoire de la procédure.

Par ailleurs, la méthode proposée sélectionne généralement un seul domaine-couleur de définition de classe parmi plusieurs domaines-couleur candidats qui représentent la même région de l'image. Les domaines candidats qui sont rejetés ne sont pas pris en compte pour construire les domaines de définition de classes, et donc pour classer les pixels. Les nombreux pixels qui n'appartiennent à aucun sous-ensemble couleur associé à un des domaines-couleur de définition de classes, sont alors assignés par une règle de décision qui ne tient pas compte des domaines-couleur candidats. La procédure de construction des classes pourrait analyser les domaines-couleur candidats afin de fusionner ceux qui représentent les mêmes régions, de sorte que les domaines-couleur de définition de classes représenteraient encore mieux les régions de l'image.

La reconstruction des régions est basée sur l'assignation des pixels aux classes construites. Il est dommage que cette assignation ne tienne pas compte de la disposition spatiale des pixels analysés. Comme notre objectif est de regrouper les pixels qui appartiennent aux mêmes régions de l'image à partir des classes de pixels construites, il serait intéressant de modéliser l'image par un graphe d'adjacence de régions. Une analyse floue du graphe telle que celle proposée par Fontaine [Fon01] permettrait de reconstruire les régions à partir des classes construites en tenant compte simultanément des propriétés spatiales et colorimétriques des pixels.

Les régions étant reconstruites, elles peuvent être analysées afin d'identifier les objets qu'elles représentent. Le prochain chapitre est consacré aux problèmes soulevés par les modifications des images couleur consécutives à des changements d'illuminants pour la reconnaissance d'objets.

3.6 Bibliographie

- [BCS98] M. Borsotti, P. Campadelli et R. Schettini. "Quantitative evaluation of color image segmentation results". *Pattern Recognition Letters*, 19(8) :741–747, 1998.
- [CJSW01] H. D. Cheng, X. H. Jiang, Y. Sun et J. Wang. "Color image segmentation : advances and prospects". *Pattern Recognition*, 34(12) :2259–2281, 2001.
- [Fon01] M. Fontaine. *Segmentation non supervisée d'images couleur par analyse de la connexité des pixels*. Thèse de doctorat, Université Lille 1, decembre 2001.
- [FtP99a] M. Fontaine, **L. Macaire** et J. G. Postaire. "Fundus images segmentation by unsupervised classification". In *Proc. of the 12th Conf. on Vision Interface*, pages 296–302, Trois-Rivières (Canada), 1999.
- [FtP99b] M. Fontaine, **L. Macaire** et J. G. Postaire. "Segmentation non supervisée d'images par analyse de la connexité, Application à la détection de lésions virales sur des images de fond d'oeil". Dans *Actes du GRETSI'99 Colloque sur le Traitement du Signal et des Images*, pages 35–38, Vannes, 1999.
- [FtP00a] M. Fontaine, **L. Macaire** et J. G. Postaire. "Color image segmentation based on a multiscale analysis of pixel connectivity properties". In *Proc. of the Int. Conf. on Color in Graphics and Image Processing*, volume 1, pages 78–83, Saint-Etienne, 2000.
- [FtP00b] M. Fontaine, **L. Macaire** et J. G. Postaire. "Image segmentation based on an original multiscale analysis of the pixel connectivity properties". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 1, pages 804–807, Vancouver (Canada), 2000.
- [FtP00c] M. Fontaine, **L. Macaire** et J. G. Postaire. "Pixel classification scheme for color image segmentation". In *Proc. of the Int. Symp. on Image/Video Communications over Fixed and Mobile Networks*, pages 1–8, Rabat (Maroc), 2000.
- [FtP00d] M. Fontaine, **L. Macaire** et J. G. Postaire. "Unsupervised segmentation based on connectivity analysis". In *Proc. of the IAPR Int. Conf. on Pattern Recognition*, volume 1, pages 660–663, Barcelone (Espagne), 2000.
- [HD86] G. Houle et E. Dubois. "Quantization of color images for display on graphics terminals". In *Proc. of the IEEE Global Telecommunications Conf. GLOBECOM86*, pages 284–297, Houston (USA), december 1986.
- [PF01] S. Philipp-Foliguet. "Evaluation de méthodes de segmentation d'images couleur". Dans *Actes de l'Ecole de printemps du GDR ISIS Operation Imagerie Couleur*, Pau, avril 2001.
- [UA94] T. Uchiyama et M. A. Arbib. "Color image segmentation using competitive learning". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(12) :1197–1206, 1994.
- [ZFB⁺01] A. Zarghili, M. Fontaine, R. Benslimane, **L. Macaire** et J. G. Postaire. "Indexation et recherche des images par similarité de formes. application à la recherche documentaire pour les arts décoratifs arabo-andalous". *Archéologie Islamique*, 11 :1–12, 2001.
- [Zha96] Y. Zhang. "A survey of evaluation methods for image segmentation". *Pattern Recognition*, 29(8) :1335–1346, 1996.

Chapitre 4

Reconnaissance d'objets sous éclairage non contrôlé

4.1 Introduction

4.1.1 Reconnaissance d'objets éclairés par différents illuminants

Dans ce chapitre, nous abordons la reconnaissance d'objets par la recherche de toutes les images candidates d'une base qui contiennent le même objet que celui représenté par une image requête. Ce problème est un cas particulier du domaine de la recherche d'images par l'exemple [TFMB04].

La reconnaissance d'objets est obtenue par la comparaison de la signature de l'image requête avec celles des images candidates. Les images candidates sont triées en fonction de mesures de similarité entre leurs signatures et celle de l'image requête. Sont retrouvées ainsi les images candidates qui contiennent des objets semblables à celui représenté dans l'image requête.

Les images considérées dans ce chapitre contiennent un seul objet placé sur un fond uniforme et éclairé avec un illuminant qui diffère d'une image à l'autre. Elles sont acquises dans les conditions suivantes (voir figure 4.1) :

- les paramètres de réglage de la caméra ne sont pas modifiés entre les acquisitions,
- les modifications d'éclairage résultent de changements d'illuminant, c'est à dire d'une modification de la température et/ou de l'intensité de l'illuminant utilisé.



FIG. 4.1 – Les images (a) et (b) contiennent le même objet éclairé avec un illuminant différent lors de l'acquisition de chaque image. Elles forment un couple d'images similaires. L'image (c) contient un objet différent de celui représenté par les images (a) et (b). Les couples d'images ((a), (c)) et ((b), (c)) forment des couples d'images différentes. Ces images extraites de la base de l'Université Simon Fraser sont disponibles sur l'internet à l'adresse : http://www.cs.sfu.ca/colour/image_db.

Nous nous proposons de nous intéresser aux problèmes soulevés par les modifications des images

couleur consécutives à des changements d'illuminant pour la reconnaissance d'objets. La procédure de recherche des images candidates contenant un objet similaire à celui représenté par l'image requête consiste à analyser chaque couple (I_{req}, I_{cand}) constitué de l'image requête I_{req} et de l'une des images candidates I_{cand} . On rencontre deux cas de figure (voir figure 4.1) :

- l'image requête I_{req} et l'image candidate I_{cand} contiennent le même objet observé dans des conditions identiques mais éclairé avec des illuminants différents. Elles forment alors un *couple d'images similaires*.
- les images I_{req} et I_{cand} contiennent des objets différents. Elles constituent alors un *couple d'images différentes*.

La solution classique pour comparer les images de chaque couple consiste à calculer l'histogramme dit "invariant" de chacune des images, à savoir l'histogramme qui est aussi peu sensible que possible aux changements d'illuminant [tM03]. Les images sont ensuite comparées à l'aide de l'intersection décrite par l'équation (1.5) de la page 56 entre leurs histogrammes invariants.

Nous avons montré au paragraphe 1.2.3 de la page 69, que la valeur de l'intersection entre les histogrammes couleur invariants d'un couple d'images similaires peut être inférieure à celle d'un couple d'images différentes. Les résultats expérimentaux obtenus en termes de reconnaissance d'objets par l'intersection entre les différents histogrammes couleur invariants ne sont alors pas très satisfaisants [FBM98, FS01].

En effet, les différents histogrammes couleur invariants sont calculés à partir de modèles de changement d'illuminant qui s'appuient sur des hypothèses sur les conditions d'éclairage, la nature des éléments de surface observés et les fonctions de sensibilité des capteurs de la caméra [GS96]. Les conditions d'acquisition des images avec lesquelles nous travaillons sont très strictes. Ceci permet d'émettre des hypothèses qui ne portent pas sur les conditions d'acquisition, mais sur les conséquences de changements d'illuminant sur les couleurs des pixels.

Ces conséquences ne modifient pas directement les niveaux des composantes couleur des pixels, mais des mesures de rang des pixels dans les images de composante. Dans chacune des images de composante I^i , $i = R, G, B$, pouvant être extraites de l'image couleur I , les pixels P sont triés dans l'ordre croissant de leurs niveaux $I^i(P)$. Nous associons des mesures de rang aux pixels, de telle sorte que les premiers classés (niveaux les plus faibles) obtiennent une mesure de rang proche de 0 et les derniers (niveaux les plus élevés), une mesure de rang égale à 1.

4.1.2 Plan du chapitre

Deux hypothèses sont présentées dans le second paragraphe de ce chapitre. La première hypothèse, proposée par Finlayson [FHST03], stipule qu'un changement d'illuminant ne provoque aucune modification des mesures de rang des pixels dans chaque image de composante. Dans ce second paragraphe, nous montrons à partir d'exemples concrets que cette hypothèse, pourtant très intéressante, *de conservation des mesures de rang* n'est pas toujours vérifiée.

Cette première hypothèse a servi de base pour proposer la seconde hypothèse, appelée hypothèse *des plus proches mesures de rang*, qui stipule qu'un changement d'illuminant provoque de légères modifications des mesures de rang des pixels dans chaque image de composante. Le deuxième paragraphe se termine par la vérification de cette hypothèse à partir de couples d'images similaires.

Ces deux hypothèses permettent de définir les fonctions qui transforment les histogrammes couleur afin de pouvoir comparer les images des objets éclairés par différents illuminants.

Le troisième paragraphe est consacré à la présentation de l'histogramme couleur invariant calculé pour

chaque image à partir de la transformation qui s'appuie sur l'hypothèse de conservation des mesures de rang. Cet histogramme est appelé *histogramme égalisé* car la transformation est décomposée en trois égalisations des histogrammes mono-dimensionnels issus de l'histogramme couleur de l'image examinée. L'utilisation de l'intersection entre les histogrammes égalisés présente cependant certaines limites pour la reconnaissance d'objets éclairés par différents illuminants.

La stratégie classique pour la reconnaissance d'objets éclairés par différents illuminants qui consiste à calculer l'histogramme invariant de chaque image, et ce de manière indépendante des autres images, ne permet pas d'atteindre toujours des résultats satisfaisants en termes de reconnaissance d'objets.

Ce constat nous amène à ne pas mettre uniquement en cause les hypothèses sur lesquelles s'appuient les transformations pour calculer les histogrammes invariants, mais à mettre également en question la stratégie elle-même de comparaison des images. Nous pensons que la reconnaissance d'objets sous éclairage non contrôlé nécessite la mise en place d'une procédure originale de comparaison d'images. Nous proposons d'analyser chaque couple constitué de l'image requête et d'une image candidate pris en considération durant la recherche des images candidates similaires à l'image requête. La comparaison des deux images de chaque couple s'effectue à partir d'un couple d'histogrammes couleur spécifiques à ce couple d'images, déterminés par la procédure décrite dans le quatrième paragraphe.

Ce couple d'histogrammes couleur spécifiques est calculé de telle sorte que l'intersection entre ces histogrammes soit plus élevée lorsque les images sont similaires que lorsque les images sont différentes. Le couple d'histogrammes couleur spécifiques est le résultat d'une transformation du couple des histogrammes couleur des images. L'hypothèse des plus proches mesures de rang peut être utilisée pour transformer le couple d'histogrammes couleur en un couple d'histogrammes couleur spécifiques car les mesures de rang des pixels dans les couples d'images de composante sont directement déduites des histogrammes couleur.

Dans le cinquième paragraphe, notre procédure de reconnaissance d'objets est testée sur deux bases d'images publiques représentant un seul objet placé sur un fond uniforme et éclairé avec des illuminants qui diffèrent d'une image à l'autre. Afin de démontrer l'intérêt de la méthode proposée, nous comparons les résultats de reconnaissance d'objets obtenus par l'intersection entre différents types d'histogrammes couleur invariants avec ceux obtenus par l'intersection entre les histogrammes couleur spécifiques.

4.2 Hypothèses basées sur les mesures de rang

Les modifications des points-couleur associés aux pixels consécutives à des changements d'illuminant sont généralement modélisées par des transformations linéaires de leurs coordonnées [MGD98]. Dans ce paragraphe, nous présentons des modèles de changement d'illuminant qui décrivent ces modifications par des transformations non-linéaires.

A partir de ces transformations non-linéaires sont émises des hypothèses à propos des conséquences d'un changement d'illuminant sur les mesures de rang des pixels dans les images de composante.

Dans un premier temps, nous décrivons l'hypothèse de conservation des mesures de rang proposée par Finlayson qui stipule qu'un changement d'illuminant ne provoque aucune modification des mesures de rang des pixels dans une image de composante.

Nous mettons ensuite en lumière les limites de cette hypothèse et présentons une hypothèse originale, l'hypothèse des plus proches mesures de rang. Cette hypothèse prend en compte les éventuelles modifications des mesures de rang qui sont les conséquences d'un changement d'illuminant.

4.2.1 Hypothèse de conservation des mesures de rang

Modèle de changement d'illuminant de Finlayson

Le modèle proposé par Finlayson [FHST03] examine de manière indépendante chacun des trois couples d'images de composante requête et candidate (I_{req}^i, I_{cand}^i) , $i = R, G, B$, pouvant être extraits d'un couple d'images similaires (I_{req}, I_{cand}) .

Soit un pixel P_{req} de l'image I_{req} et un pixel P_{cand} de l'image I_{cand} sur lesquels se projette un même élément de surface de l'objet. Nous appelons le couple (P_{req}, P_{cand}) un *couple de pixels correspondants*. Le niveau $I_{cand}^i(P_{cand})$, $i = R, G, B$, du pixel P_{cand} est déduit du niveau $I_{req}^i(P_{req})$ du pixel correspondant P_{req} à l'aide d'une fonction monotone croissante $f_{req,cand}^i$ qui dépend du couple d'images de composante (I_{req}^i, I_{cand}^i) :

$$I_{cand}^i(P_{cand}) = f_{req,cand}^i(I_{req}^i(P_{req})), i = R, G, B. \quad (4.1)$$

Ces trois fonctions monotones croissantes $f_{req,cand}^i$ ne sont pas a priori des fonctions linéaires.

Considérons maintenant deux couples de pixels correspondants (P_{req}, P_{cand}) et (Q_{req}, Q_{cand}) . Si le niveau $I_{req}^i(P_{req})$ est supérieur à $I_{req}^i(Q_{req})$, $i = R, G, B$, nous déduisons de l'équation (4.1) que le niveau $I_{cand}^i(P_{cand})$ est supérieur à $I_{cand}^i(Q_{cand})$:

$$I_{req}^i(P_{req}) > I_{req}^i(Q_{req}) \Rightarrow f_{req,cand}^i(I_{req}^i(P_{req})) > f_{req,cand}^i(I_{req}^i(Q_{req})) \Rightarrow I_{cand}^i(P_{cand}) > I_{cand}^i(Q_{cand}). \quad (4.2)$$

Cette relation nous amène à examiner les mesures de rang des pixels dans les images de composante requête et candidate.

Mesures de rang

Dans chaque image de composante I^i , les pixels sont triés dans l'ordre croissant de leurs niveaux. Nous associons une mesure de rang à chaque pixel, de telle sorte qu'elle soit proche de 0 pour les premiers pixels et égale à 1 pour les derniers pixels. Pour chaque pixel P , la mesure de rang $M^i[\mathbf{I}](P)$ dans l'image de composante I^i est exprimée par :

$$M^i[\mathbf{I}](P) = \frac{\sum_{y=0}^{I^i(P)} H^i[\mathbf{I}](y)}{\sum_{y=0}^{L-1} H^i[\mathbf{I}](y)}, i = R, G, B, \quad (4.3)$$

où L indique le nombre de niveaux de quantification des composantes couleur (L est généralement égal à 256), et $H^i[\mathbf{I}](y)$ est le nombre de pixels dont le niveau est égal à y .

Hypothèse de conservation des mesures de rang

A partir de l'équation (4.2), Finlayson [FHST03] pose l'hypothèse de conservation des mesures de rang qui stipule que les mesures de rang des couples de pixels correspondants (P_{req}, P_{cand}) dans chaque couple d'images de composante (I_{req}^i, I_{cand}^i) sont égales :

$$M^i[\mathbf{I}_{req}](P_{req}) = M^i[\mathbf{I}_{cand}](P_{cand}), i = R, G, B. \quad (4.4)$$

Les mesures de rang des pixels dans chaque image de composante sont supposées être invariantes à

n'importe quel changement d'illuminant.

Finlayson vérifie la validité de cette hypothèse avec un échantillon de 462 fonctions de réflectance spectrale caractérisant 462 plaquettes-échantillons extraites de l'atlas de Munsell et les fonctions de distribution spectrale de 16 illuminants différents [FHST03]. A l'aide de l'équation (1.4) décrite à la page 45, il calcule les niveaux des composantes couleur des pixels qui représentent les 462 plaquettes-échantillons éclairées par l'un des 16 illuminants. Il compare ensuite les niveaux calculés de ces pixels et montre ainsi qu'un changement d'illuminant provoque très peu de modifications des mesures de rang des pixels dans une image de composante.

Vérification avec des images couleur similaires

Comme la validité de cette hypothèse n'a été vérifiée qu'avec des niveaux de pixels calculés, nous proposons de la vérifier avec les couples d'images vertes issues des couples d'images couleur similaires de la figure 4.2.



FIG. 4.2 – Les images (a),..., (e) sont des images similaires extraites de la base de l'Université Simon Fraser, disponibles sur l'internet à l'adresse : http://www.cs.sfu.ca/colour/image_db.

Comme dans ce cas particulier l'objet n'a pas bougé entre les acquisitions de ces images, les pixels de mêmes coordonnées spatiales dans ces cinq images forment des couples de pixels correspondants. Ces images ont été choisies car elles permettent de vérifier facilement la validité de l'hypothèse de conservation des mesures de rang. En effet, l'équation (4.4) est vraie uniquement si les mesures de rang des pixels de mêmes coordonnées spatiales dans les images vertes sont égales. Le tableau 4.1 indique les pourcentages de pixels de mêmes coordonnées spatiales dont les mesures de rang dans les différentes images vertes issues des images couleur de la figure 4.2 sont égales. Pour cela, nous faisons abstraction du fond et ne tenons compte que des pixels qui représentent l'objet dans chaque image. La moyenne de ces pourcentages est égale à 6,9%, ce qui montre que l'hypothèse de conservation des mesures de rang ne peut pas être considérée comme vérifiée avec ces images couleur. En effet, sous cette hypothèse, on ignore totalement les éventuelles modifications de mesures de rang des pixels provoquées par un changement d'illuminant.

Image requête \ Image candidate	Image requête				
	4.2(a)	4.2(b)	4.2(c)	4.2(d)	4.2(e)
4.2(a)	X	4	2	5	13
4.2(b)	4	X	7	9	8
4.2(c)	2	7	X	3	3
4.2(d)	5	9	3	X	15
4.2(e)	13	8	3	15	X

TAB. 4.1 – Pourcentages de pixels de mêmes coordonnées spatiales dont les mesures de rang dans les images requête et candidate vertes issues des images couleur de la figure 4.2 sont égales (Les composantes sont quantifiées avec 256 niveaux).

4.2.2 Hypothèse des plus proches mesures de rang

Extension du modèle de Finlayson

Comme la validité de l'hypothèse de conservation des mesures de rang n'est pas vérifiée, nous proposons un nouveau modèle de changement d'illuminant. Il décrit les relations entre les niveaux de composante de deux pixels correspondants (P_{req}, P_{cand}) dans un couple d'images similaires $(\mathbf{I}_{req}, \mathbf{I}_{cand})$. Le niveau $I_{cand}^i(P_{cand})$ du pixel P_{cand} est exprimé en fonction du niveau $I_{req}^i(P_{req})$ du pixel correspondant P_{req} par :

$$I_{cand}^i(P_{cand}) = f_{req,cand}^i(I_{req}^i(P_{req})) + \rho^i(P_{cand}), \quad i = R, G, B, \quad (4.5)$$

où la fonction $f_{req,cand}^i$ est monotone croissante et dépend du couple d'images de composante requête et candidate (I_{req}^i, I_{cand}^i) considérées. La fonction ρ^i est une fonction de biais qui dépend du pixel P_{cand} dans l'image de composante candidate I_{cand}^i . Cette fonction représente les éventuelles modifications de mesures de rang des pixels correspondants dans les deux images de composante.

Ainsi, en considérant deux couples de pixels correspondants (P_{req}, P_{cand}) et (Q_{req}, Q_{cand}) pour lesquels le niveau $I_{req}^i(P_{req})$ est supérieur au niveau $I_{req}^i(Q_{req})$, l'équation (4.5) ne permet pas de conclure que le niveau $I_{cand}^i(P_{cand})$ est supérieur au niveau $I_{cand}^i(Q_{cand})$.

Hypothèse des plus proches mesures de rang

Ce nouveau modèle permet de proposer l'hypothèse des plus proches mesures de rang qui porte sur chacun des trois couples d'images de composante (I_{req}^i, I_{cand}^i) pouvant être extraits d'un couple d'images couleur similaires $(\mathbf{I}_{req}, \mathbf{I}_{cand})$. Elle stipule que chaque couple de pixels correspondants (P_{req}, P_{cand}) respecte deux conditions :

- la mesure de rang de P_{req} dans l'image de composante I_{req}^i est, parmi toutes les mesures de rang des pixels dans I_{req}^i , celle qui est la plus proche de la mesure de rang de P_{cand} dans I_{cand}^i :

$$\|M^i[\mathbf{I}_{cand}](P_{cand}) - M^i[\mathbf{I}_{req}](P_{req})\| = \min_{P \in I_{req}^i} \|M^i[\mathbf{I}_{cand}](P_{cand}) - M^i[\mathbf{I}_{req}](P)\|, \quad i = R, G, B, \quad (4.6)$$

- la mesure de rang de P_{cand} dans l'image de composante I_{cand}^i est, parmi toutes les mesures de rang des pixels dans I_{cand}^i , celle qui est la plus proche de la mesure de rang de P_{req} dans I_{req}^i :

$$\|M^i[\mathbf{I}_{cand}](P_{cand}) - M^i[\mathbf{I}_{req}](P_{req})\| = \min_{P \in I_{cand}^i} \|M^i[\mathbf{I}_{req}](P_{req}) - M^i[\mathbf{I}_{cand}](P)\|, \quad i = R, G, B. \quad (4.7)$$

Pour simplifier la suite de la présentation, nous indiquons que les mesures de rang d'un couple de pixels sont *les plus proches dans le couple d'images de composante* (I_{req}^i, I_{cand}^i) si ce couple de pixels respecte les équations (4.6) et (4.7).

L'hypothèse des plus proches mesures de rang stipule donc que lorsque les images $\mathbf{I}_{req}$ et $\mathbf{I}_{cand}$ sont similaires, les couples de pixels correspondants sont ceux ayant les mesures de rang les plus proches dans chaque couple d'images de composante (I_{req}^i, I_{cand}^i) .

Par le biais des équations (4.6) et (4.7), l'hypothèse des plus proches mesures de rang prend en considération les mesures de rang de tous les pixels dans chaque couple d'images de composante. De ce fait, elle se démarque de l'hypothèse de conservation des mesures de rang.

Vérification avec des images couleur similaires

De la même manière que pour l'hypothèse de conservation des mesures de rang, nous proposons de vérifier la validité de l'hypothèse des plus proches mesures de rang avec les images vertes issues des images couleur de la figure 4.2, où les pixels de mêmes coordonnées spatiales forment des couples de pixels correspondants.

Pour chaque couple d'images vertes, nous recherchons les couples de pixels avec les mesures de rang les plus proches. Le tableau 4.2 indique les pourcentages de cas où les pixels de ces couples ont les mêmes coordonnées spatiales. Comme pour le tableau 4.1, nous ne prenons en considération que les pixels représentant l'objet dans chaque image en faisant abstraction du fond. La moyenne des pourcentages de ces cas s'élève à 49%.

En comparant les tableaux 4.1 et 4.2, nous constatons que les pourcentages de cas où les couples de pixels avec les mêmes coordonnées spatiales forment des couples de pixels avec les mesures de rang les plus proches sont bien plus élevés que les pourcentages de couples de pixels de mêmes coordonnées spatiales avec des mesures de rang identiques. A partir de ces exemples, nous concluons que l'hypothèse des plus proches mesures de rang reflète nettement mieux les conséquences d'un changement d'illuminant sur les mesures de rang des pixels que l'hypothèse de conservation des mesures de rang.

Dans ces conditions, nous considérons que la validité de l'hypothèse des plus proches mesures de rang est partiellement vérifiée avec les images de la figure 4.2.

Image requête \ Image candidate	4.2(a)	4.2(b)	4.2(c)	4.2(d)	4.2(e)
4.2(a)	X	36	31	44	46
4.2(b)	36	X	70	53	62
4.2(c)	31	70	X	33	40
4.2(d)	44	53	33	X	75
4.2(e)	46	62	40	75	X

TAB. 4.2 – Pourcentages de cas où les pixels des couples avec les mesures de rang les plus proches ont les mêmes coordonnées spatiales dans les images vertes issues des images couleur de la figure 4.2 (Les composantes sont quantifiées avec 256 niveaux).

4.3 Histogrammes couleur égalisés

Avant de décrire les histogrammes couleur spécifiques calculés sous l'hypothèse des plus proches mesures de rang, nous présentons les histogrammes couleur invariants calculés sous l'hypothèse de conservation des mesures de rang. Finlayson [FHST03] propose de transformer l'histogramme couleur $\mathbf{H}[\mathbf{I}]$ d'une image $\mathbf{I}$ en un histogramme couleur invariant, appelé *histogramme couleur égalisé* et noté $\mathbf{H}_{egal}[\mathbf{I}]$, à l'aide d'une fonction $\mathbf{T}_{egal}$ qui prend en compte les mesures de rang des pixels dans les images de composante I^i , $i = R, G, B$.

La fonction $\mathbf{T}_{egal}$ est décomposée en trois fonctions $T_{egal}^R, T_{egal}^G, T_{egal}^B$, chacune transformant respectivement les histogrammes mono-dimensionnels $H^R[\mathbf{I}]$, $H^G[\mathbf{I}]$, $H^B[\mathbf{I}]$. Chaque histogramme mono-dimensionnel $H^i[\mathbf{I}]$, $i = R, G, B$, est le résultat de la projection ϕ^i de l'histogramme couleur $\mathbf{H}[\mathbf{I}]$ sur la i^{eme} composante couleur :

$$H^i[\mathbf{I}] = \phi^i(\mathbf{H}[\mathbf{I}]), \quad i = R, G, B. \quad (4.8)$$

Ainsi, pour déterminer l'histogramme couleur égalisé $\mathbf{H}_{egal}[\mathbf{I}]$, on fait appel à trois fonctions T_{egal}^i , $i = R, G, B$ qui transforment de manière indépendante l'histogramme couleur $\mathbf{H}[\mathbf{I}]$ selon le schéma :

$$\mathbf{H}_{egal}[\mathbf{I}] = T_{egal}^R \circ T_{egal}^G \circ T_{egal}^B(\mathbf{H}[\mathbf{I}]). \quad (4.9)$$

Les trois composantes étant supposées orthogonales, l'histogramme mono-dimensionnel égalisé $H_{egal}^i[\mathbf{I}]$ est le résultat de la projection ϕ^i de $T_{egal}^i(\mathbf{H}[\mathbf{I}])$ sur la i^{eme} composante couleur :

$$H_{egal}^i[\mathbf{I}] = \phi^i(T_{egal}^i(\mathbf{H}[\mathbf{I}])) , i = R, G, B. \quad (4.10)$$

Comme les fonctions T_{egal}^i transforment de manière indépendante les trois histogrammes mono-dimensionnels $H^i[\mathbf{I}]$, la détermination de l'histogramme couleur égalisé $\mathbf{H}_{egal}[\mathbf{I}]$ passe par le calcul des trois histogrammes mono-dimensionnels égalisés $H_{egal}^i[\mathbf{I}]$. Nous allons donc examiner le calcul de chacun de ces histogrammes mono-dimensionnels égalisés.

Précédemment, nous avons vu que Finlayson propose de caractériser les pixels de l'image couleur $\mathbf{I}$ par leurs mesures de rang dans chacune des trois images de composante I^i de manière à obtenir une nouvelle représentation qui serait insensible à l'éclairage. Comme les mesures de rang de tous les pixels de même niveau sont égales, par souci de simplification de notation, la mesure de rang d'un pixel correspond dans la suite de ce chapitre à la mesure de rang, notée $M^i[\mathbf{I}](l)$, du niveau l de ce pixel.

La mesure de rang $M^i[\mathbf{I}](l)$ du niveau l présent dans l'image de composante I^i est obtenue par l'égalisation de l'histogramme mono-dimensionnel $H^i[\mathbf{I}]$:

$$M^i[\mathbf{I}](l) = \frac{\sum_{y=0}^l H^i[\mathbf{I}](y)}{\sum_{y=0}^{L-1} H^i[\mathbf{I}](y)}, l = 0, \dots, (L-1). \quad (4.11)$$

L'histogramme mono-dimensionnel égalisé $H_{egal}^i[\mathbf{I}]$ contient les mesures de rang des niveaux présents dans I^i . Le nombre N de cellules de l'histogramme mono-dimensionnel égalisé est un paramètre ajusté par l'utilisateur à une valeur généralement inférieure ou égale au nombre L de niveaux utilisés pour quantifier chaque composante couleur (N est généralement fixé à 16).

L'histogramme mono-dimensionnel égalisé $H_{egal}^i[\mathbf{I}]$ s'exprime à partir de l'histogramme mono-dimensionnel $H^i[\mathbf{I}]$ de la manière suivante :

$$H_{egal}^i[\mathbf{I}](n) = \sum_{LUT_{egal}^i[\mathbf{I}](l)=n} H^i[\mathbf{I}](l), n = 0, \dots, (N-1), \quad (4.12)$$

avec :

$$LUT_{egal}^i[\mathbf{I}](l) = int \left(N \cdot \frac{\sum_{y=0}^l H^i[\mathbf{I}](y)}{\sum_{y=0}^{L-1} H^i[\mathbf{I}](y)} \right), l = 0, \dots, (L-1). \quad (4.13)$$

La cellule de l'histogramme mono-dimensionnel égalisé $H_{egal}^i[\mathbf{I}]$ associée au niveau n accumule les valeurs des cellules de l'histogramme mono-dimensionnel $H^i[\mathbf{I}]$ associées aux niveaux l , pour lesquels $LUT_{egal}^i[\mathbf{I}](l)$ est égale à n .

Comme cette approche transforme l'histogramme couleur $\mathbf{H}[\mathbf{I}]$ en un histogramme couleur égalisé

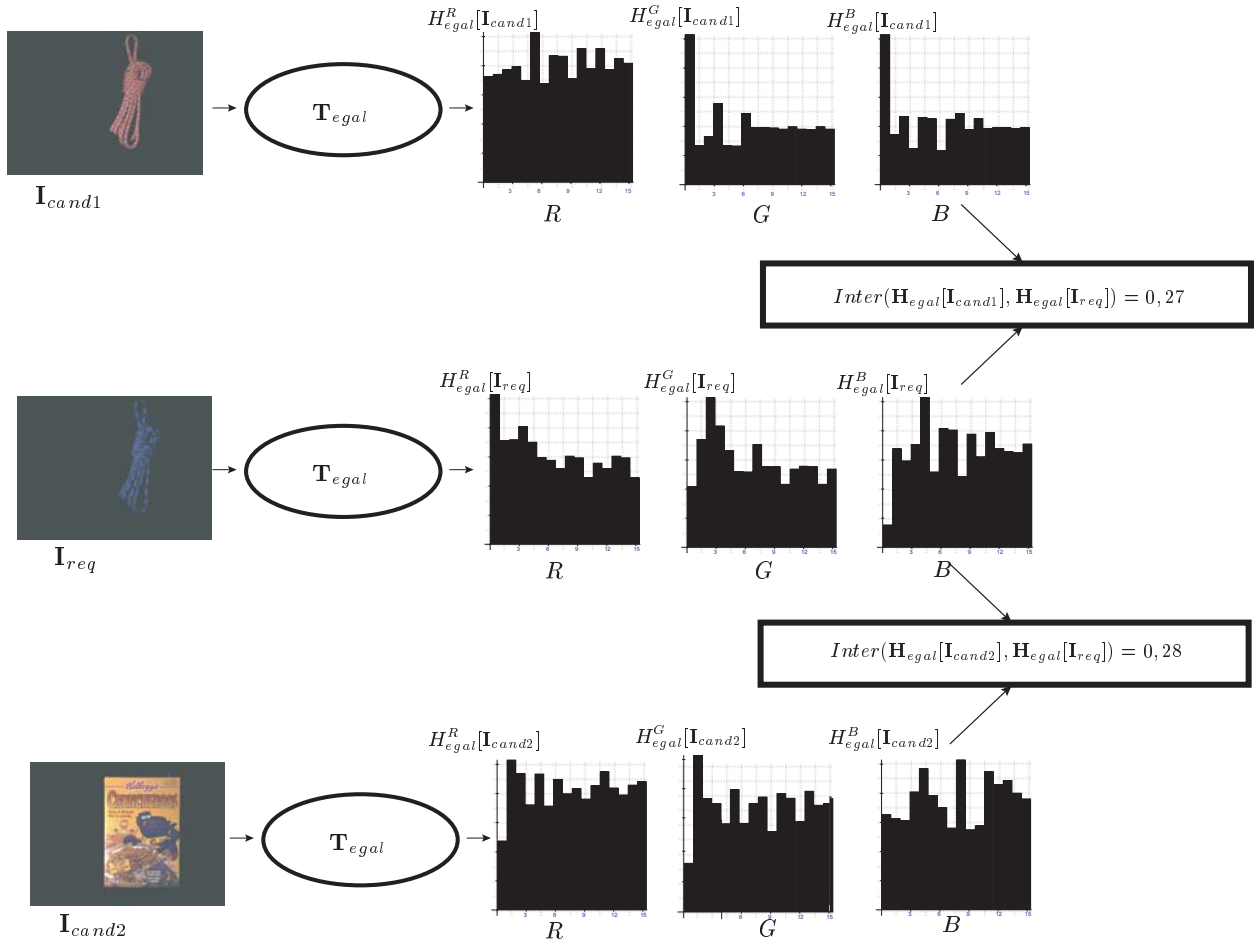


FIG. 4.3 – Comparaison entre une image requête I_{req} et deux images candidates I_{cand1} et I_{cand2} par les intersections entre leurs histogrammes couleur égalisés. N est fixé à 16.

$H_{egal}[\mathbf{I}]$ en utilisant successivement trois tables mono-dimensionnelles de correspondance $LUT_{egal}^i[\mathbf{I}]$, le temps de calcul nécessaire à l'exécution de cette transformation est très faible.

La figure 4.3 présente la comparaison entre une image requête I_{req} et deux images candidates I_{cand1} et I_{cand2} par l'intersection entre les histogrammes couleur égalisés. Bien que les images I_{req} et I_{cand1} soient similaires, leurs histogrammes mono-dimensionnels égalisés ne semblent pas être identiques sur la figure. La valeur de l'intersection entre leurs histogrammes couleur égalisés $H_{egal}[I_{req}]$ et $H_{egal}[I_{cand1}]$ n'atteint que 0,27.

Les images I_{req} et I_{cand2} sont différentes et l'intersection entre leurs histogrammes couleur égalisés $H_{egal}[I_{req}]$ et $H_{egal}[I_{cand2}]$ s'élève à 0,28, valeur supérieure à celle de l'intersection entre les histogrammes couleur égalisés de I_{req} et I_{cand1} .

Cet exemple montre que l'intersection entre ces histogrammes couleur égalisés ne permet pas toujours de discerner le cas où les images requête et candidate sont similaires de celui où elles sont différentes.

Plus généralement, les procédures basées sur le calcul d'histogrammes couleur invariants analysent toutes les images de la base indépendamment les unes des autres. Les résultats obtenus par ces approches ne sont pas satisfaisants en termes de reconnaissance d'objets éclairés avec différents illuminants [FBM98, FS01]. C'est pourquoi nous proposons, dans le paragraphe suivant, une nouvelle approche qui détermine un couple d'histogrammes couleur spécifiques à chaque couple d'images considérées.

4.4 Histogrammes couleur spécifiques aux couples d'images

Pour comparer l'image requête $\mathbf{I}_{req}$ et l'une des images candidates $\mathbf{I}_{cand}$, nous proposons non pas de calculer les histogrammes couleur invariants des deux images, et ce de manière indépendante pour chacune d'elles, mais plutôt d'adapter leurs histogrammes couleur. L'intersection entre ces histogrammes couleur spécifiques doit être plus élevée lorsque les images sont similaires que lorsqu'elles sont différentes [MtBP03].

Dans la première partie de ce paragraphe, nous décrivons comment une fonction $\mathbf{T}_{req,cand}$ transforme le couple d'histogrammes couleur $(\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}])$ en un couple d'histogrammes couleur spécifiques $(\mathbf{H}_{req,cand}[\mathbf{I}_{req}], \mathbf{H}_{req,cand}[\mathbf{I}_{cand}])$. Cette transformation s'appuie sur des couples de tables de correspondance présentées dans la seconde partie. Dans la dernière partie de ce paragraphe, nous détaillons la procédure qui permet de déterminer ces couples de tables de correspondance.

4.4.1 Couple d'histogrammes couleur spécifiques

La détermination du couple d'histogrammes couleur spécifiques passe par l'examen des projections du couple d'histogrammes couleur sur les composantes couleur.

Notons $(H^i[\mathbf{I}_{req}], H^i[\mathbf{I}_{cand}])$, $i = R, G, B$, le couple d'histogrammes mono-dimensionnels qui résulte de la projection ψ^i du couple d'histogrammes couleur $(\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}])$ sur la i^{eme} composante couleur :

$$(H^i[\mathbf{I}_{req}], H^i[\mathbf{I}_{cand}]) = \psi^i(\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}]), \quad i = R, G, B. \quad (4.14)$$

La fonction $\mathbf{T}_{req,cand}$ est décomposée en trois fonctions $T_{req,cand}^i$, $i = R, G, B$, qui transforment uniquement les projections des couples d'histogrammes couleur sur les composantes couleur. Pour être plus précis, pour déterminer le couple d'histogrammes couleur spécifiques $(\mathbf{H}_{req,cand}[\mathbf{I}_{req}], \mathbf{H}_{req,cand}[\mathbf{I}_{cand}])$, nous appliquons successivement trois fonctions $T_{req,cand}^i$ au couple d'histogrammes couleur $(\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}])$:

$$(\mathbf{H}_{req,cand}[\mathbf{I}_{req}], \mathbf{H}_{req,cand}[\mathbf{I}_{cand}]) = T_{req,cand}^R \circ T_{req,cand}^G \circ T_{req,cand}^B(\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}]). \quad (4.15)$$

Notons $(H_{req,cand}^i[\mathbf{I}_{req}], H_{req,cand}^i[\mathbf{I}_{cand}])$ le couple d'histogrammes mono-dimensionnels spécifiques qui est le résultat de la projection ψ^i sur la i^{eme} composante couleur du couple d'histogrammes couleur spécifiques $(\mathbf{H}_{req,cand}[\mathbf{I}_{req}], \mathbf{H}_{req,cand}[\mathbf{I}_{cand}])$.

Puisque les fonctions $T_{req,cand}^i$ transforment de manière indépendante les trois couples d'histogrammes mono-dimensionnels $(H^i[\mathbf{I}_{req}], H^i[\mathbf{I}_{cand}])$, la détermination du couple d'histogrammes couleur spécifiques $(\mathbf{H}_{req,cand}[\mathbf{I}_{req}], \mathbf{H}_{req,cand}[\mathbf{I}_{cand}])$ nécessite le calcul des trois couples d'histogrammes mono-dimensionnels spécifiques $(H_{req,cand}^i[\mathbf{I}_{req}], H_{req,cand}^i[\mathbf{I}_{cand}])$. Nous allons donc examiner le calcul de chacun de ces couples d'histogrammes mono-dimensionnels spécifiques.

Le couple d'histogrammes mono-dimensionnels spécifiques $(H_{req,cand}^i[\mathbf{I}_{req}], H_{req,cand}^i[\mathbf{I}_{cand}])$ est déterminé à partir du couple d'histogrammes mono-dimensionnels $(H^i[\mathbf{I}_{req}], H^i[\mathbf{I}_{cand}])$ à l'aide du couple de tables mono-dimensionnelles de correspondance $(LUT_{req,cand}^i[\mathbf{I}_{req}], LUT_{req,cand}^i[\mathbf{I}_{cand}])$:

$$H_{req,cand}^i[\mathbf{I}_{req}](n) = \sum_{LUT_{req,cand}^i[\mathbf{I}_{req}](l)=n} H^i[\mathbf{I}_{req}](l), \quad n = 0, \dots, (N-1), \quad (4.16)$$

et :

$$H_{req,cand}^i[\mathbf{I}_{cand}](n) = \sum_{LUT_{req,cand}^i[\mathbf{I}_{cand}](l)=n} H^i[\mathbf{I}_{cand}](l), \quad n = 0, \dots, (N-1). \quad (4.17)$$

La cellule de l'histogramme mono-dimensionnel spécifique $H_{req,cand}^i[\mathbf{I}_{req}]$ associée au niveau n , avec $n = 0, \dots, (N-1)$, accumule les valeurs des cellules de l'histogramme mono-dimensionnel $H^i[\mathbf{I}_{req}]$ associées aux niveaux l , avec $l = 0, \dots, (L-1)$, pour lesquels $LUT_{req,cand}^i[\mathbf{I}_{req}](l)$ est égale à n .

De la même manière, la cellule de l'histogramme mono-dimensionnel spécifique $H_{req,cand}^i[\mathbf{I}_{cand}]$ associée au niveau n accumule les valeurs des cellules de l'histogramme mono-dimensionnel $H^i[\mathbf{I}_{cand}]$ associées aux niveaux l , pour lesquels $LUT_{req,cand}^i[\mathbf{I}_{cand}](l)$ est égale à n .

La procédure suivante détermine le couple de tables mono-dimensionnelles de correspondance $(LUT_{req,cand}^i[\mathbf{I}_{req}], LUT_{req,cand}^i[\mathbf{I}_{cand}])$ spécifiques à chaque couple d'images de composante (I_{req}^i, I_{cand}^i) .

4.4.2 Couple de tables mono-dimensionnelles de correspondance

Les histogrammes mono-dimensionnels requête et candidat contiennent au maximum N cellules non vides. Le nombre N est fixé par l'utilisateur à une valeur inférieure ou égale à L , le nombre de niveaux de chaque composante couleur. La procédure analyse les valeurs des L cellules de chaque histogramme mono-dimensionnel pour déterminer les N cellules de l'histogramme mono-dimensionnel spécifique qui en découlent. Nous considérons uniquement des accumulations des valeurs de cellules associées à des niveaux adjacents pour simplifier la complexité de la procédure de détermination des couples de tables mono-dimensionnelles de correspondance.

Pour construire l'histogramme mono-dimensionnel requête (respectivement candidat) spécifique, la procédure sélectionne $(N+1)$ niveaux notés $X_{req}^i(n)$, avec $n = 0, \dots, N$ (respectivement notés $X_{cand}^i(n)$) parmi les L possibles. Ces $(N+1)$ niveaux délimitent N intervalles le long de la composante couleur considérée. Les valeurs des cellules de l'histogramme mono-dimensionnel requête (respectivement candidat) associées aux niveaux appartenant à l'un des N intervalles sont accumulées dans l'une des N cellules de l'histogramme mono-dimensionnel requête (respectivement candidat) spécifique.

Le couple de tables mono-dimensionnelles de correspondance $(LUT_{req,cand}^i[\mathbf{I}_{req}], LUT_{req,cand}^i[\mathbf{I}_{cand}])$ s'exprime à partir des couples de niveaux retenus $(X_{req}^i(n), X_{cand}^i(n))$ comme :

$$LUT_{req,cand}^i[\mathbf{I}_{req}](l) = n \text{ pour } l \in [X_{req}^i(n), X_{req}^i(n+1)[, \quad n = 0, \dots, (N-1), \quad (4.18)$$

et :

$$LUT_{req,cand}^i[\mathbf{I}_{cand}](l) = n \text{ pour } l \in [X_{cand}^i(n), X_{cand}^i(n+1)[, \quad n = 0, \dots, (N-1). \quad (4.19)$$

Les couples de niveaux retenus $(X_{req}^i(0), X_{cand}^i(0))$ et $(X_{req}^i(N), X_{cand}^i(N))$ sont respectivement fixés à $(0,0)$ et $(L-1, L-1)$, de telle sorte que les tables mono-dimensionnelles de correspondance soient définies pour tous les niveaux compris entre 0 et $(L-1)$. Il reste alors $(N-1)$ couples de niveaux $(X_{req}^i(n), X_{cand}^i(n))$ à sélectionner.

La table mono-dimensionnelle de correspondance $LUT_{req,cand}^i[\mathbf{I}_{req}]$ accumule les valeurs des cellules de $H^i[\mathbf{I}_{req}]$ associées aux niveaux compris entre $X_{req}^i(n)$ et $X_{req}^i(n+1) - 1$, avec $n = 0, \dots, (N-1)$, dans la cellule $H_{req,cand}^i[\mathbf{I}_{req}](n)$ de l'histogramme mono-dimensionnel spécifique requête $H_{req,cand}^i[\mathbf{I}_{req}]$.

De la même manière, la table mono-dimensionnelle de correspondance $LUT_{req,cand}^i[\mathbf{I}_{cand}]$ accumule les valeurs des cellules de $H^i[\mathbf{I}_{cand}]$ associées aux niveaux compris entre $X_{cand}^i(n)$ et $X_{cand}^i(n+1)-1$ dans la cellule $H_{req,cand}^i[\mathbf{I}_{cand}](n)$ de l'histogramme mono-dimensionnel spécifique candidat $H_{req,cand}^i[\mathbf{I}_{cand}]$.

Déterminer les couples de tables mono-dimensionnelles de correspondance $(LUT_{req,cand}^i[\mathbf{I}_{req}], LUT_{req,cand}^i[\mathbf{I}_{cand}])$, $i = R, G, B$, consiste à sélectionner les $(N-1)$ couples de niveaux $(X_{req}^i(n), X_{cand}^i(n))$, avec $n = 1, \dots, (N-1)$.

Nous proposons que ces couples soient sélectionnés parmi ceux dont les mesures de rang sont les plus proches dans le couple d'images de composante. Ainsi, la procédure respecte l'hypothèse des plus proches mesures de rang quand les images sont similaires.

D'autre part, la sélection des couples de niveaux influence directement l'intersection entre les histogrammes spécifiques. En effet, les équations (4.18) et (4.19) montrent que les sous-ensembles de pixels de I_{req}^i dont les niveaux appartiennent à $[X_{req}^i(n), X_{req}^i(n+1)[$ et les sous-ensembles de pixels de I_{cand}^i dont les niveaux appartiennent à $[X_{cand}^i(n), X_{cand}^i(n+1)[$, sont tous les deux représentés par des cellules de mêmes coordonnées dans les histogrammes mono-dimensionnels spécifiques. L'intersection entre les histogrammes mono-dimensionnels spécifiques est basé sur le calcul du minimum des cardinaux de ces deux sous-ensembles.

Nous proposons que les couples de niveaux soient sélectionnés de manière à "réduire" les effets du changement d'illuminant se produisant entre les acquisitions des images requête et candidate similaires sur les histogrammes couleur spécifiques requête et candidat. Nous considérons que ces effets sont effectivement "réduits" si l'intersection entre les histogrammes couleur spécifiques atteint une valeur plus élevée lorsque les images sont similaires que lorsque les images sont différentes.

Cependant, nous ne savons pas a priori si les images requête et candidate sont similaires. Dans ces conditions, les $3 \times (N-1)$ couples de niveaux $(X_{req}^i(n), X_{cand}^i(n))$ sont sélectionnés de telle sorte que les trois intersections entre les histogrammes mono-dimensionnels spécifiques $H_{req,cand}^i[\mathbf{I}_{req}]$ et $H_{req,cand}^i[\mathbf{I}_{cand}]$ soient élevées, et ce que les images requête et candidate soient similaires ou différentes.

Après la sélection des $3 \times (N-1)$ couples de niveaux $(X_{req}^i(n), X_{cand}^i(n))$, nous pouvons construire les 3 couples de tables mono-dimensionnelles de correspondance $(LUT_{req,cand}^i[\mathbf{I}_{req}], LUT_{req,cand}^i[\mathbf{I}_{cand}])$ grâce aux équations (4.18) et (4.19). Ces couples de tables mono-dimensionnelles de correspondance sont ensuite utilisés pour transformer le couple d'histogrammes couleur $(\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}])$ en un couple d'histogrammes couleur spécifiques $(\mathbf{H}_{req,cand}[\mathbf{I}_{req}], \mathbf{H}_{req,cand}[\mathbf{I}_{cand}])$.

4.4.3 Sélection des couples de niveaux

La sélection des couples de niveaux se déroule en deux étapes successives. La première étape recherche les couples de niveaux dont les mesures de rang sont les plus proches dans le couple d'images de composante requête et candidate. Cette première étape permet de respecter l'hypothèse émise sur les modifications des mesures de rang des pixels consécutives à un changement d'illuminant quand les images sont similaires. La seconde sélectionne parmi ces couples de niveaux ceux qui permettent d'obtenir une valeur élevée d'intersection entre les histogrammes mono-dimensionnels spécifiques, afin de "réduire" les effets du changement d'illuminant sur le couple d'histogrammes couleur spécifiques quand les images sont similaires.

Recherche des couples de niveaux dont les mesures de rang sont les plus proches

La première étape recherche les $nb_couples^i$ couples de niveaux dont les mesures de rang dans le couple d'images de composante (I_{req}^i, I_{cand}^i) sont les plus proches.

Ces $nb_couples^i$ couples de niveaux, notés $(x_{req}^i(a), x_{cand}^i(a))$, avec $a = 1, \dots, nb_couples^i$, sont ceux qui respectent les deux conditions suivantes :

- la mesure de rang de $x_{req}^i(a)$ dans l'image de composante I_{req}^i est, parmi toutes les mesures de rang des niveaux présents dans I_{req}^i , celle qui est la plus proche de la mesure de rang de $x_{cand}^i(a)$ dans I_{cand}^i :

$$x_{req}^i(a) = y \text{ tel que } \|M^i[\mathbf{I}_{cand}](x_{cand}^i(a)) - M^i[\mathbf{I}_{req}](y)\| = \min_{w=0, \dots, L-1} \|M^i[\mathbf{I}_{cand}](x_{cand}^i(a)) - M^i[\mathbf{I}_{req}](w)\|. \quad (4.20)$$

- la mesure de rang de $x_{cand}^i(a)$ dans l'image de composante I_{cand}^i est, parmi toutes les mesures de rang des niveaux présents dans I_{cand}^i , celle qui est la plus proche de la mesure de rang de $x_{req}^i(a)$ dans I_{req}^i :

$$x_{cand}^i(a) = y \text{ tel que } \|M^i[\mathbf{I}_{req}](x_{req}^i(a)) - M^i[\mathbf{I}_{cand}](y)\| = \min_{w=0, \dots, L-1} \|M^i[\mathbf{I}_{req}](x_{req}^i(a)) - M^i[\mathbf{I}_{cand}](w)\|. \quad (4.21)$$

L'hypothèse des plus proches mesures de rang suppose que les pixels caractérisés par le niveau $x_{req}^i(a)$ dans I_{req}^i et les pixels caractérisés par le niveau $x_{cand}^i(a)$ dans I_{cand}^i constituent des couples de pixels correspondants lorsque les images sont similaires. Ainsi, les couples de niveaux $(x_{req}^i(a), x_{cand}^i(a))$ caractérisent des couples de pixels correspondants lorsque les images sont similaires. Il faut noter que tous les niveaux présents dans chacune des images de composante ne forment pas nécessairement des couples avec des mesures de rang les plus proches.

Comme le nombre $nb_couples^i$ de couples de niveaux ainsi déterminés est probablement supérieur au nombre $(N - 1)$, nous proposons de ne retenir que $(N - 1)$ couples, notés $(X_{req}^i(n), X_{cand}^i(n))$ avec $n = 1, \dots, (N - 1)$. Dans le cas contraire, tous les couples sont sélectionnés et les histogrammes mono-dimensionnels spécifiques ne contiennent que $nb_couples^i$ cellules.

Sélection des $(N - 1)$ couples $(X_{req}^i(n), X_{cand}^i(n))$

Les $(N - 1)$ couples de niveaux $(X_{req}^i(n), X_{cand}^i(n))$ sont sélectionnés pour obtenir une valeur élevée de l'intersection entre les histogrammes spécifiques mono-dimensionnels.

Une solution consisterait à sélectionner les $(N - 1)$ couples avec les écarts de mesures de rang les plus faibles parmi les $nb_couples^i$ couples de niveaux ainsi déterminés. Cette solution ne serait pas satisfaisante en termes de discrimination des images car elle aurait tendance à sélectionner des couples de niveaux qui représentent de faibles populations de pixels dans le couple d'images de composante.

Pour pallier ce problème, la procédure proposée détermine le nombre de pixels dans I_{req}^i caractérisés par les niveaux compris entre $x_{req}^i(a)$ et $(x_{req}^i(a + 1) - 1)$. En effet, l'intersection des histogrammes mono-dimensionnels spécifiques tient compte du minimum entre le nombre de pixels dont les niveaux sont compris entre $x_{req}^i(a)$ et $(x_{req}^i(a + 1) - 1)$ dans I_{req}^i et le nombre de pixels dont les niveaux sont compris entre $x_{cand}^i(a)$ et $(x_{cand}^i(a + 1) - 1)$ dans I_{cand}^i . La procédure retient ensuite les $(N - 1)$ niveaux $x_{req}^i(a)$ avec ces nombres de pixels les plus élevés. A chaque niveau $x_{req}^i(a)$ sélectionné, nous associons le niveau $x_{cand}^i(a)$ pour constituer l'un des $(N - 1)$ couples de niveaux à sélectionner $(X_{req}^i(n), X_{cand}^i(n))$.

Ces $(N - 1)$ couples de niveaux sélectionnés permettent de construire le couple d'histogrammes mono-dimensionnels spécifiques dont l'intersection atteint une valeur élevée, que les images soient similaires ou différentes.

4.4.4 Exemple

La figure 4.4 présente la comparaison entre une image requête $\mathbf{I}_{req}$ et deux images candidates $\mathbf{I}_{cand1}$ et $\mathbf{I}_{cand2}$ par les intersections entre les histogrammes couleur spécifiques. Les images $\mathbf{I}_{req}$ et $\mathbf{I}_{cand1}$ sont similaires alors que l'image $\mathbf{I}_{cand2}$ contient un objet différent.

Cette figure montre que les histogrammes mono-dimensionnels spécifiques $H_{req,cand1}^i[\mathbf{I}_{req}]$ de l'image requête calculés pour comparer $\mathbf{I}_{req}$ et $\mathbf{I}_{cand1}$ sont différents des histogrammes mono-dimensionnels spécifiques $H_{req,cand2}^i[\mathbf{I}_{req}]$ de l'image requête calculés pour comparer $\mathbf{I}_{req}$ et $\mathbf{I}_{cand2}$. Dans cette figure, nous avons encerclé les valeurs des intersections entre les histogrammes mono-dimensionnels spécifiques $H_{req,cand}^i[\mathbf{I}_{req}]$ et $H_{req,cand}^i[\mathbf{I}_{cand}]$. Ces valeurs d'intersection sont comprises entre 0,93 et 0,98, ce qui confirme que les couples d'histogrammes mono-dimensionnels spécifiques sont construits de sorte que leurs intersections atteignent des valeurs élevées, et ce que les images soient similaires ou différentes.

La comparaison des images est basée sur l'intersection entre les histogrammes couleur spécifiques. L'intersection entre les histogrammes couleur spécifiques $\mathbf{H}_{req,cand1}[\mathbf{I}_{req}]$ et $\mathbf{H}_{req,cand1}[\mathbf{I}_{cand1}]$ des images similaires atteint 0,41. L'intersection entre les histogrammes couleur spécifiques $\mathbf{H}_{req,cand2}[\mathbf{I}_{req}]$ et $\mathbf{H}_{req,cand2}[\mathbf{I}_{cand2}]$ des images différentes est égale à 0,36 valeur inférieure à l'intersection entre les histogrammes couleur spécifiques des images similaires.

Cet exemple illustre que la détermination des couples de niveaux $(X_{req}^i(n), X_{cand}^i(n))$ respecte le compromis entre l'obtention de valeurs élevées des intersections entre les histogrammes mono-dimensionnels spécifiques pour "réduire" les effets d'un changement d'illuminant sur le couple d'histogrammes couleur spécifiques quand les images sont similaires, et la discrimination grâce à l'intersection des histogrammes couleur spécifiques entre le cas où les images sont similaires et celui où elles sont différentes.

La procédure présentée détermine le couple d'histogrammes couleur spécifiques à partir du couple d'histogrammes couleur requête et candidat construit à chaque comparaison effectuée pendant la recherche d'images candidates similaires à l'image requête. Le temps de calcul des couples d'histogrammes spécifiques doit être par conséquent faible pour une implantation réaliste de cette procédure. Le couple d'histogrammes couleur est transformé en un couple d'histogrammes couleur spécifiques en appliquant successivement trois transformations s'appuyant sur des couples de tables mono-dimensionnelles de correspondance. La complexité de cette transformation est de l'ordre de L , le nombre de niveaux utilisés pour quantifier les composantes couleur.

Chaque couple de tables mono-dimensionnelles de correspondance est construit en déterminant $(N - 1)$ couples de niveaux $(X_{req}^i(n), X_{cand}^i(n))$ qui sont sélectionnés parmi les L^2 possibles. Comme cette sélection consiste à rechercher les couples de niveaux dont les mesures de rang sont les plus proches, la complexité de cette construction est de l'ordre de L^2 , ce qui permet d'envisager l'intégration de notre méthode à un système de reconnaissance automatique d'objets éclairés par différents illuminants.

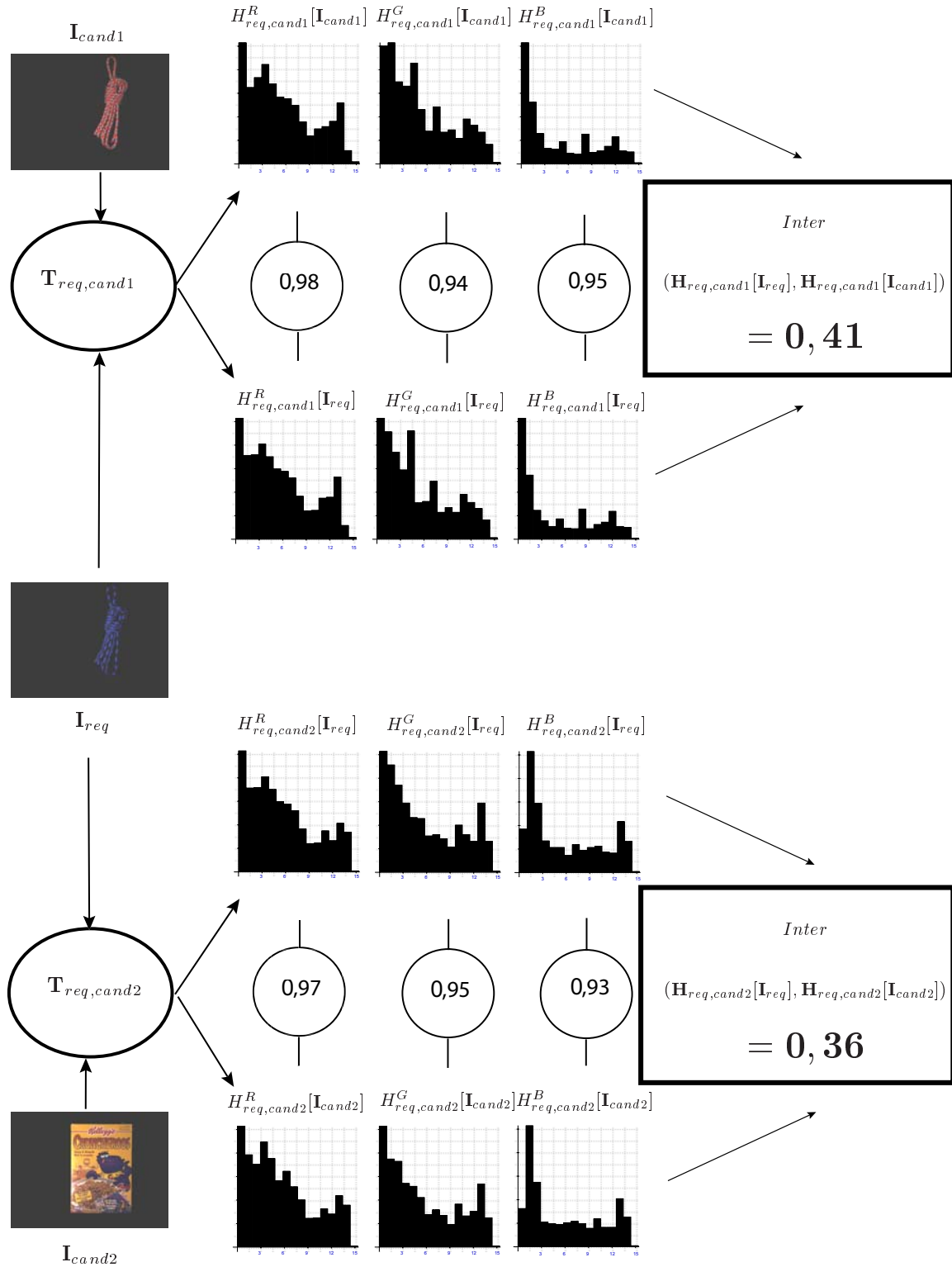


FIG. 4.4 – Comparaison entre une image requête I_{req} et deux images candidates I_{cand1} et I_{cand2} par les intersections entre les histogrammes couleur spécifiques. N est fixé à 16.

4.5 Résultats expérimentaux

La présentation de résultats expérimentaux va permettre de mettre en évidence l'intérêt de notre méthode pour la reconnaissance d'objets éclairés par différents illuminants, en comparant une image requête à différentes images candidates grâce à l'intersection entre leurs histogrammes couleur spécifiques. Pour ceci, nous utilisons deux bases d'images publiques spécifiquement constituées pour mesurer l'influence des changements d'illuminant sur les résultats obtenus par des procédures de reconnaissance des objets. Comme ces bases ont déjà été exploitées pour effectuer de nombreux tests, nous pouvons comparer les résultats de reconnaissance d'objets obtenus par notre approche avec ceux obtenus par l'intersection entre différents types d'histogrammes couleur invariants.

4.5.1 Reconnaissance d'objets de la base de l'Université Simon Fraser

La base de l'Université Simon Fraser, disponible à l'adresse internet http://www.cs.sfu.ca/~colour/image_db, se compose de 55 images couleur contenant chacune un objet choisi parmi 11 et placé sur un fond uniforme (voir figure 4.5). Chaque image représente un objet éclairé par l'un des 5 illuminants utilisés, un même objet conservant la même position dans les 5 images qui le représentent (voir figure 4.6).

Pour chaque recherche d'images, les images acquises avec un illuminant, appelé illuminant candidat, sont considérées comme étant les images candidates et l'une des images acquises avec l'un des 4 autres illuminants, appelé illuminant requête, est considérée comme étant l'image requête. Ainsi, il y a 5×4 couples d'illuminants requête-candidat différents. La recherche d'images est répétée pour chacun des 11 objets. Finalement, 220 recherches sont réalisées ($11 \text{ objets} \times 5 \times 4 \text{ couples d'illuminants}$).

Pour chaque recherche d'images, les 11 images candidates sont triées en fonction des intersections entre leurs histogrammes couleur invariants (respectivement spécifiques) et l'histogramme couleur invariant (respectivement spécifique) de l'image requête considérée. Lorsque l'image candidate similaire à l'image requête est classée première, le résultat de la recherche est considéré comme parfait.

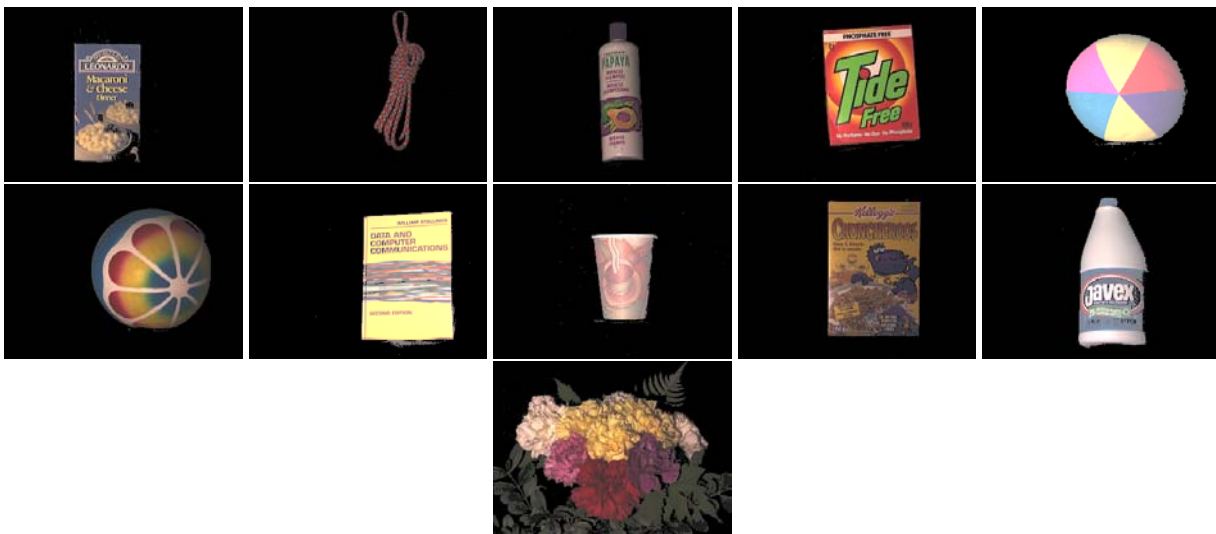


FIG. 4.5 – Les 11 objets représentés par les images de la base de l'Université Simon Fraser. Le même illuminant a été utilisé lors de l'acquisition de ces images.

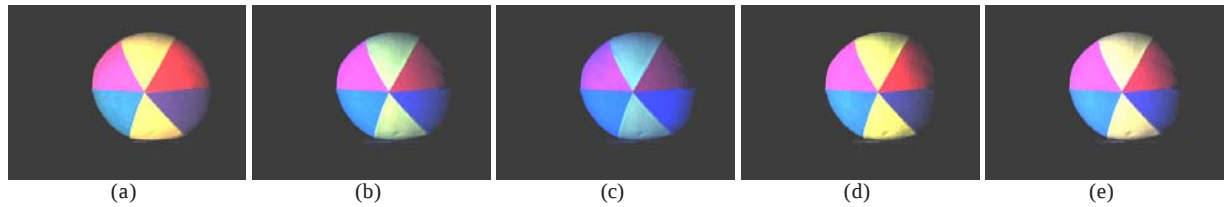


FIG. 4.6 – Les images (a)...(e) contiennent le même objet éclairé avec un illuminant différent lors de l'acquisition de chaque image.

Funt conclut à partir d'expériences menées sur cette base que les meilleurs résultats de reconnaissance d'objets sont obtenus par l'intersection entre les histogrammes couleur invariants basés sur la normalisation "greyworld" [FBM98]. Finlayson montre que l'intersection entre les histogrammes couleur égalisés (voir équation (4.12)) apporte des résultats supérieurs à ceux obtenus par la normalisation "greyworld" [FHST03]. Par conséquent, nous proposons de comparer les résultats obtenus par ces deux procédures avec ceux obtenus par l'intersection entre les histogrammes couleur spécifiques.

Le tableau 4.3 indique les résultats de reconnaissance d'objets obtenus par les intersections entre les histogrammes couleur invariants testés et par l'intersection entre les histogrammes couleur spécifiques, lorsque le nombre N est fixé à 16. Chaque case de la colonne correspondant au rang i indique le pourcentage cumulé de recherches d'images où l'image candidate similaire à l'image requête obtient la j^{eme} place parmi les images candidates, avec $j \leq i$. Ainsi, la procédure conduit à des résultats de reconnaissance d'objets parfaits si la valeur correspondant au rang 1 atteint 100%.

Intersection entre les	Rang = 1	Rang ≤ 2	Rang ≤ 3
histogrammes couleur invariants (normalisation "greyworld")	98,18	99,09	99,55
histogrammes couleur égalisés	98,18	99,55	100
histogrammes couleur spécifiques	100	100	100

TAB. 4.3 – Résultats de reconnaissance d'objets obtenus par les intersections entre les différents histogrammes avec la base de l'Université Simon Fraser. N est fixé à 16.

La comparaison des résultats présentés dans le tableau 4.3 montre que l'intersection entre les histogrammes couleur invariants calculés par la normalisation "greyworld" fournit les plus mauvais résultats. L'intersection entre les histogrammes couleur spécifiques fournit des résultats parfaits, meilleurs que ceux obtenus par l'intersection entre les histogrammes couleur égalisés.

Comme ces résultats dépendent du nombre N de cellules des histogrammes monodimensionnels invariants et spécifiques, il est intéressant de mesurer l'influence de son ajustement [MtBP04]. La figure 4.7 représente les pourcentages de cas où l'image candidate similaire à l'image requête a été classée première, en fonction de N .

Sur la figure 4.7, nous constatons que la qualité des résultats de reconnaissance d'objets obtenus par l'intersection entre les histogrammes invariants est très sensible au nombre de niveaux avec lesquels sont quantifiées les coordonnées des cellules. La qualité des résultats de reconnaissance d'objets obtenus par l'intersection entre les histogrammes invariants diminue lorsque N augmente, tandis que les résultats obtenus par l'intersection entre les histogrammes couleur spécifiques restent stables.

La base d'images de l'Université Simon Fraser est trop réduite pour permettre de juger complètement de l'efficacité d'une méthode pour la reconnaissance d'objets éclairés par différents illuminants. Elle permet néanmoins de mettre en évidence le fait que les histogrammes couleur invariants sont très

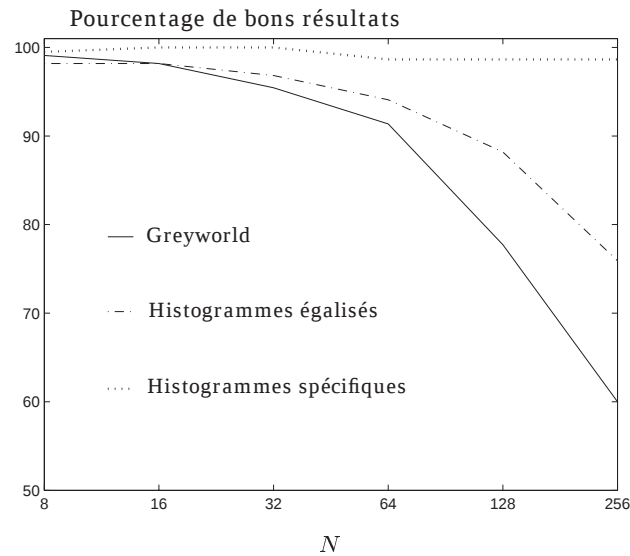


FIG. 4.7 – Évolution des résultats de reconnaissance d'objets sur la base de l'Université Simon Fraser en fonction de N .

sensibles au nombre de niveaux avec lesquels sont quantifiées les coordonnées de leurs cellules. Par ailleurs, les résultats obtenus par notre méthode illustrent sa robustesse par rapport à l'ajustement de ce nombre N .

Afin de démontrer l'intérêt de notre méthode dans des situations non triviales, nous proposons de l'appliquer sur une autre base qui contient une large gamme d'objets de même type éclairés par différents illuminants.

4.5.2 Reconnaissance d'objets de la base de l'Université East Anglia

La seconde base d'images utilisée est maintenue disponible par l'Université East Anglia à l'adresse internet <http://vision.doc.ntu.ac.uk/research/CATSI/database.html>. Ses 336 images contiennent 28 tapisseries multicolores (voir figure 4.8) éclairées par l'un des 3 illuminants utilisés et acquises avec des conditions d'observation identiques par l'une des 4 caméras disponibles (voir figure 4.9). La discrimination entre ces 28 tapisseries multicolores n'est pas triviale puisque les distributions des points-couleur associés aux pixels dans les images (a), (b), (c) et (d), ainsi que dans les images (e), (f) et (g) de la figure 4.8 sont presque identiques [FS01].

Ceci constitue l'une des raisons pour lesquelles aucun histogramme couleur invariant ne fournit de résultats satisfaisants de reconnaissance d'objets avec cette base [FS01].

Afin de comparer les résultats de reconnaissance d'objets obtenus par l'intersection entre les histogrammes couleur invariants et ceux obtenus par l'intersection entre les histogrammes couleur spécifiques, nous utilisons le même protocole expérimental que celui décrit dans [FS01, FHST03]. Cependant, comme notre but n'est pas de tester l'insensibilité de notre procédure aux changements de caméra, les sous-ensembles d'images acquises par chacune des 4 caméras sont analysés séparément.

Pour chacun de ces 4 sous-ensembles, chaque recherche d'images considère une image requête parmi les 28 images acquises par une caméra sous un illuminant requête. La recherche consiste à retrouver parmi les 28 images candidates acquises avec un illuminant candidat qui est différent de l'illumi-

nant requête, l'image candidate similaire à l'image requête. Il y a 3×2 couples d'illuminants requête-candidat différents. Cette procédure est répétée pour chacune des 28 tapisseries. Finalement, 672 recherches sont réalisées ($28 \text{ tapisseries} \times 3 \times 2 \text{ couples d'illuminants requête-candidat} \times 4 \text{ caméras}$).

Pour chaque recherche, les 28 images sont triées dans l'ordre décroissant des valeurs d'intersection entre leurs histogrammes couleur invariants (respectivement spécifiques) et l'histogramme couleur invariant (respectivement spécifique) de l'image requête considérée. Quand l'image candidate similaire à l'image requête est classée première, le résultat de reconnaissance est considéré comme parfait.

Pour pouvoir comparer les résultats présentés par Finlayson, nous utilisons le critère qu'il a retenu, à savoir le pourcentage de reconnaissance moyen, noté AMP et calculé de la manière suivante :

$$AMP = \frac{1}{N_{recherche}} \times \sum_{i=1}^{N_{recherche}} \frac{N_{cand} - Rang(i)}{N_{cand} - 1}, \quad (4.22)$$

où $N_{recherche}$ est le nombre total de recherches ($N_{recherche} = 672$), N_{cand} est le nombre total d'images candidates analysées pour chaque recherche ($N_{cand} = 28$) et $Rang(i)$ est le rang obtenu par l'image candidate similaire à l'image requête lors de la i^{eme} recherche. Les résultats sont parfaits lorsque la valeur de l' AMP est égale à 100%.

Le tableau 4.4 montre les pourcentages de reconnaissance moyens (AMP) obtenus par l'intersection entre les histogrammes couleur invariants basés sur la normalisation "greyworld" [FS01], par l'intersection entre les histogrammes couleur égalisés [FHST03], et ceux obtenus par l'intersection entre les histogrammes couleur spécifiques. Comme le préconisent les auteurs, le nombre N est fixé à 16.

Intersection entre les	AMP
histogrammes couleur invariants (normalisation "greyworld") [FS01]	93,96
histogrammes couleur égalisés [FHST03]	96,72
histogrammes couleur spécifiques	98,83

TAB. 4.4 – Pourcentages de reconnaissance moyens obtenus par les intersections entre les différents histogrammes sur la base de l'Université East Anglia. N est fixé à 16.

Le tableau 4.4 confirme que l'intersection entre les histogrammes couleur spécifiques fournit de meilleurs résultats que ceux obtenus par les intersections entre les histogrammes couleur invariants. Il est difficile de dégager d'autres conclusions à partir de la seule valeur de l' AMP . Examinons en détail les résultats de reconnaissance d'objets obtenus par l'intersection entre les histogrammes couleur spécifiques.

Intersection entre les	Rang = 1	Rang $\leq$ 2	Rang $\leq$ 3
histogrammes couleur spécifiques	86,61	94,35	97,32

TAB. 4.5 – Résultats de reconnaissance d'objets de la base de l'Université East Anglia obtenus par l'intersection entre les histogrammes couleur spécifiques. N est fixé à 16.

Le tableau 4.5 montre que l'image candidate similaire à l'image requête est classée première par l'intersection entre les histogrammes couleur spécifiques dans 86,61% des cas. Les distributions des points-couleur associés aux pixels dans plusieurs images de la base sont presque identiques. Dans ces conditions, le score obtenu par l'intersection entre les histogrammes couleur spécifiques peut être considéré comme satisfaisant.

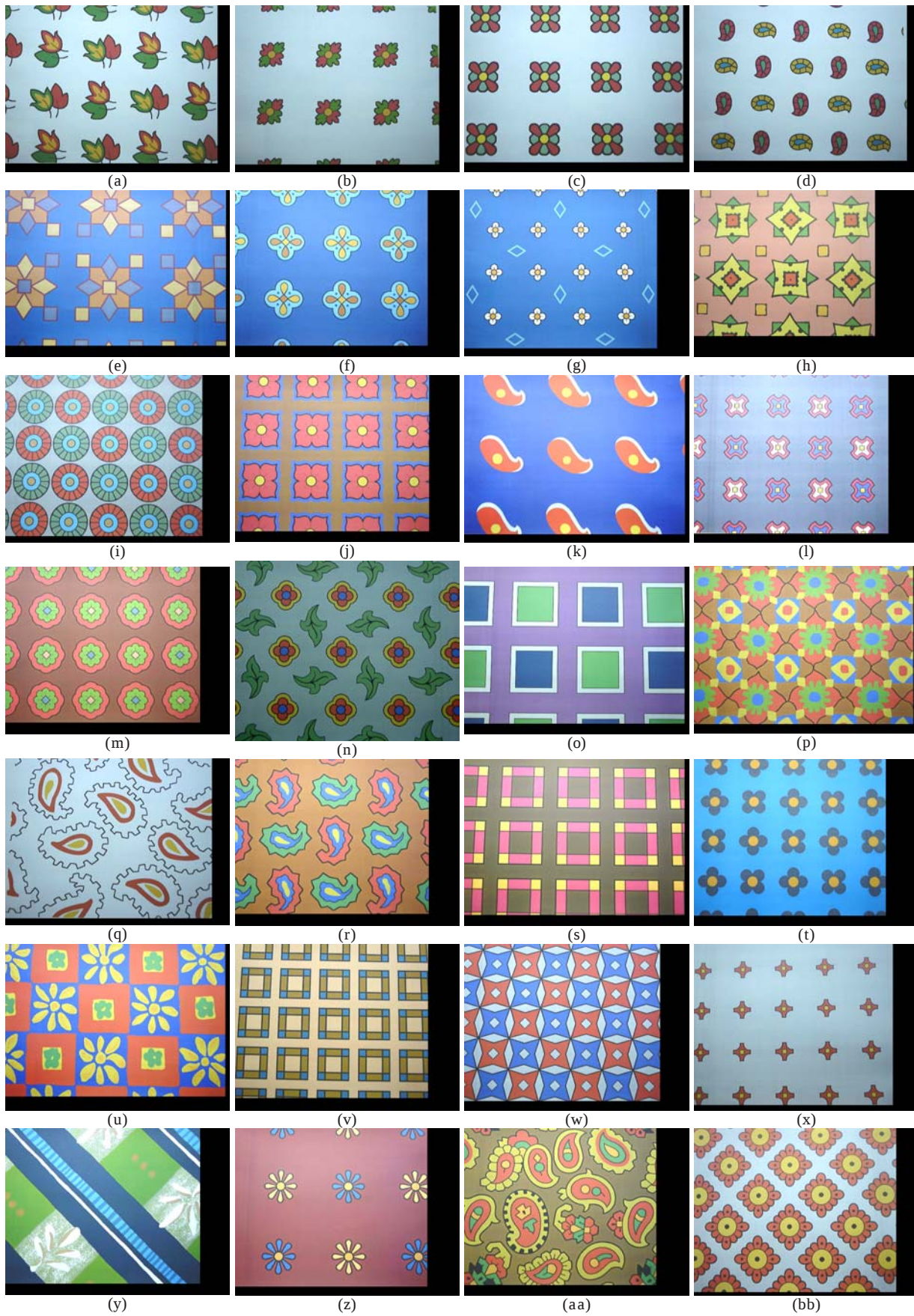


FIG. 4.8 – Les 28 tapisseries représentées par les images de la base de l'Université East Anglia. Le même illuminant a été utilisé lors de l'acquisition de ces images.

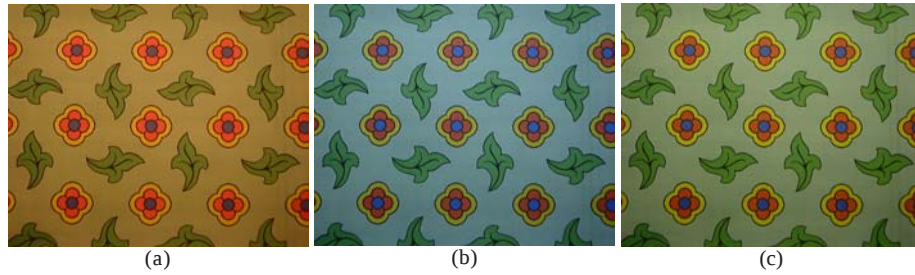


FIG. 4.9 – Les images (a)...(c) sont acquises par la même caméra et contiennent la même tapisserie éclairée avec un illuminant différent lors de l'acquisition de chaque image.

4.5.3 Discussion

La supériorité des résultats obtenus par notre procédure s'explique par les trois points principaux suivants.

Tout d'abord, le calcul des couples d'histogrammes couleur spécifiques est basé sur une hypothèse originale, l'hypothèse des plus proches mesures de rang qui est moins restrictive que l'hypothèse de conservation des mesures de rang. Ces différents résultats expérimentaux montrent le rôle essentiel de l'hypothèse à propos des conséquences d'un changement d'illuminant sur les mesures de rang des pixels retenue pour définir la transformation des histogrammes couleur des images considérées.

Par ailleurs, les résultats de reconnaissance d'objets dépendent du nombre N de cellules des histogrammes mono-dimensionnels invariants ou spécifiques. Pour calculer les histogrammes couleur invariants, la réduction du nombre de cellules de $(L)^3$ à $(N)^3$ est réalisée par une quantification uniforme des niveaux des composantes couleur. Cette quantification regroupe des cellules tout en négligeant la distribution des points-couleur représentant les pixels dans l'espace couleur (R, G, B) . La procédure qui calcule les couples d'histogrammes couleur spécifiques prend en compte la distribution des points-couleur, notamment lors de la sélection des couples de niveaux $(X_{req}^i(n), X_{cand}^i(n))$. Par conséquent, les couples d'histogrammes couleur spécifiques ne subissent pas une étape de quantification uniforme qui dégrade la qualité de la représentation des distributions des points-couleur présents dans le couple d'images considérées.

La stratégie de comparaison des images est le dernier élément d'explication de cette supériorité des résultats. Les histogrammes couleur invariants sont déterminés en considérant les images requête et candidate indépendamment les unes des autres. Les résultats expérimentaux montrent que cette stratégie ne permet pas toujours de discriminer le cas où les images sont similaires du cas où elles sont différentes. La démarche que nous proposons est de considérer chaque couple constitué de l'image requête et de l'une des images candidates, puis de calculer le couple d'histogrammes couleur spécifiques à ce couple d'images. Cette stratégie est plus coûteuse en temps de calcul que celle adoptée par le calcul des histogrammes invariants, car elle nécessite le calcul du couple d'histogrammes couleur spécifiques pour chaque comparaison entre deux images. Les résultats prometteurs obtenus en suivant cette stratégie justifient le sur-coût en temps de calcul induit par l'analyse de chaque couple d'images.

4.6 Conclusion

Dans ce chapitre, nous avons proposé une approche originale pour répondre au problème de la reconnaissance d'objets éclairés par différents illuminants. Plutôt que de calculer l'histogramme couleur

invariant de chaque image de la base, nous considérons chaque couple constitué de l'image requête et de l'une des images candidates. Pour chaque couple, la procédure proposée détermine le couple d'histogrammes couleur spécifiques aux deux images considérées.

Cette procédure est basée sur une nouvelle hypothèse à propos des conséquences d'un changement d'illuminant sur les mesures de rang des pixels, l'hypothèse des plus proches mesures de rang. Elle suppose que lorsque les images sont similaires, les couples de pixels correspondants forment des couples avec les plus proches mesures de rang. Cette hypothèse est originale car elle ne considère pas indépendamment chaque image de composante, mais tient compte des mesures de rang des pixels dans chacun des trois couples d'images de composante.

Les résultats expérimentaux montrent l'intérêt de notre démarche qui consiste à calculer le couple d'histogrammes spécifiques à chaque couple d'images, de telle sorte que dans le cas d'images similaires, leur intersection soit plus élevée que dans le cas où elles sont différentes. La détermination du couple d'histogrammes couleur spécifiques passe par le calcul des trois couples d'histogrammes mono-dimensionnels spécifiques. Chacun de ces couples est calculé de sorte que leur intersection atteigne une valeur élevée, et ce que les images soient similaires ou différentes. Quand les images sont similaires, ces couples d'histogrammes mono-dimensionnels spécifiques permettent de "réduire" les effets d'un changement d'illuminant sur le couple d'histogrammes couleur spécifiques.

Nous avons suivi cette même démarche en utilisant des fonctions de calcul des histogrammes couleur spécifiques qui s'appuient sur un modèle de changement d'illuminant très simple à savoir le modèle diagonal [MtPK02]. Les résultats obtenus avec une base restreinte d'images montre que les transformations des histogrammes basées sur ce modèle peuvent s'avérer adaptées à des cas simples de reconnaissance d'objets sous éclairage non contrôlé.

Par ailleurs, les résultats de reconnaissance d'objets pourraient être améliorés en exploitant des signatures qui prennent en compte simultanément les distributions des points-couleur et les interactions spatiales entre les pixels. Ces signatures spatio-colorimétriques, telles que les matrices de co-occurrences couleur [StP00], ont été appliquées pour la reconnaissance de personnes [MtP02] et l'identification de véhicules se déplaçant sur autoroute [MMtP03].

La procédure proposée s'applique à la reconnaissance d'objets lorsque les images contiennent un seul objet observé dans des conditions identiques. Nous travaillons actuellement à la généralisation du calcul d'histogrammes couleur spécifiques de manière à reconnaître des objets contenus dans des images qui représentent plusieurs objets. Cette méthode pourrait être alors exploitée par des systèmes de recherche d'images par le contenu, où les conditions d'acquisition des images considérées ne sont pas contrôlées.

4.7 Bibliographie

- [FBM98] B. Funt, K. Barnard et L. Martin. "Is machine colour constancy good enough?". In *Proc. of the 5th European Conf. on Computer Vision*, pages 445–459, Freiburg (Allemagne), 1998.
- [FHST03] G. Finlayson, S. Hordley, G. Schaefer et G. Tian. "Illuminant and device invariant colour using histogram equalisation". In *Proc. of the IS&T/SID Eleventh Color Imaging Conf.*, pages 205–211, Scottsdale (USA), 2003.
- [FS01] G. Finlayson et G. Schaefer. "Colour indexing across devices and viewing conditions". In *Proc. of the 2nd Int. Workshop on Content-based MultiMedia Indexing*, pages 215–221, Brescia (Italie), 2001.
- [GS96] T. Gevers et A. W. M. Smeulders. "A comparative study of several color models for color image invariant retrieval". In *Proc. of the First Int. Workshop on the Image Databases and Multi-Media Search*, pages 17–26, Amsterdam (Pays-Bas), 1996.
- [MGD98] P. Montesinos, V. Gouet et R. Deriche. "Differential invariants for color images". In *Proc. of the Int. Conf. on Pattern Recognition*, volume 1, pages 838–840, Brisbane (Australie), 1998.

- [MMtP03] D. Muselet, C. Motamed, **L. Macaire** et J. G. Postaire. "Co-occurrence matrices of color feature vectors for multi-camera vehicle identification". In *Proc. of the Int. Conf. on Advanced Concepts for Intelligent Vision Systems*, pages 22–29, Ghent (Belgique), 2003.
- [MtBP03] D. Muselet, **L. Macaire**, P. Bonnet et J. G. Postaire. "Reconnaissance d'objets grâce à l'analyse des composantes". Dans *Actes du GRETSI'03 Colloque sur le Traitement du Signal et des Images.*, volume 3, pages 30–33, Paris, 2003.
- [MtBP04] D. Muselet, **L. Macaire**, P. Bonnet et J. G. Postaire. "Reconnaissance d'objets sous éclairage non contrôlé par analyse d'histogrammes invariants et adaptés". Dans *Actes de CORESA'2004 Journées d'études et d'échanges COMpression et REprésentation des Signaux Audiovisuels*, pages 203–206, Villeneuve d'Ascq, 2004.
- [MtP02] D. Muselet, **L. Macaire** et J. G. Postaire. "Color person image indexing and retrieval". *Machine Graphics and Vision*, 11(2/3) :257–283, 2002.
- [MtPK02] D. Muselet, **L. Macaire**, J. G. Postaire et L. Khoudour. "Color invariant for person images indexing". In *Proc. of the first European Conf. on Colour in Graphics, Image and Vision*, pages 236–240, Poitiers, 2002.
- [StP00] M. Skrzypniak, **L. Macaire** et J. G. Postaire. "Indexation d'images de personnes par analyse de matrices de co-occurrences couleur". Dans *Actes de CORESA'2000 Journées d'études et d'échanges COMpression et REprésentation des Signaux Audiovisuels*, pages 411–418, Poitiers, 2000.
- [TFMB04] A. Trémeau, C. Fernandez-Maloigne et P. Bonton. *Image couleur : de l'acquisition au traitement*. Editions Dunod, Paris, 2004.
- [tM03] **L. Macaire** et D. Muselet. "Invariants couleur - méthodes et applications". Dans *Actes de l'Ecole de printemps du GDR ISIS Operation Imagerie Couleur*, Dijon, avril 2003.

Conclusion et perspectives

Ces pages ont retracé mon parcours depuis mon arrivée au Centre d'Automatique de Lille de l'Université des Sciences et Technologies de Lille en 1989. Pour présenter les contributions de notre travail (le terme notre s'impose car ce travail s'est effectué avec Jack-Gérard Postaire et en collaboration avec les doctorants que j'ai eu le plaisir de co-encadrer), trois thèmes choisis ont été détaillés. Ils s'intègrent dans une chaîne de traitements des images couleur et portent sur :

- la sélection de l'espace de représentation de la couleur approprié au type d'images à analyser,
- la segmentation des images par classification spatio-colorimétrique des pixels afin de reconstruire les régions représentant les objets présents dans les images,
- la reconnaissance des objets sous éclairage non contrôlé.

Nos travaux ont tenté de se démarquer des approches développées pour le traitement d'images en niveaux de gris. Les méthodes que nous avons proposées sont spécifiques à l'analyse d'images couleur. Ainsi, les thèmes présentés ont été abordés en considérant l'information couleur caractérisant chaque pixel comme une entité vectorielle à part entière.

La démarche suivie pour la sélection de l'espace de représentation de la couleur a été développée pour la segmentation d'images par classification supervisée des pixels. La solution proposée s'appuie sur une analyse exclusive des propriétés colorimétriques des pixels. Or, nous avons vu que la prise en compte simultanée des propriétés spatiales et colorimétriques des pixels est une approche très intéressante pour la segmentation des images. C'est pour cette raison qu'avec Jack-Gérard Postaire et Nicolas Vandenbroucke, nous poursuivons cette démarche en analysant les propriétés de connexité et d'homogénéité colorimétrique de sous-ensembles de pixels afin de sélectionner des espaces couleur hybrides adaptés aux types d'images considérés. Laurent Busin, doctorant au LAGIS depuis septembre 2003, travaille sur cette problématique qui est appliquée au contrôle-qualité de l'aspect de produits verriers par analyse d'images linéaires couleur. La très large gamme de motifs colorés fixés sur les verres observés par la caméra va permettre de tester les algorithmes développés pour sélectionner automatiquement des espaces couleur hybrides adaptés par l'analyse des propriétés spatio-colorimétriques de sous-ensembles de pixels.

De manière plus générale, la méthodologie de sélection automatique d'espaces couleur adaptés que nous avons développée dans le cadre de la segmentation mérite d'être étendue aux autres étapes constituant une chaîne de traitements, telles que notamment le filtrage des images ou la reconnaissance des objets.

Par ailleurs, les degrés de connexité et d'homogénéité de sous-ensembles de pixels constituent un outil original que nous avons utilisé pour la segmentation des images par classification spatio-colorimétrique des pixels. Cet outil pourrait être également exploité par des procédures de reconnaissance d'objets sous la forme d'une signature de chaque image qui tiendrait compte simultanément de la distribution des points-couleur et de la répartition spatiale des pixels. Comme les propriétés d'interaction spatiale entre les pixels sont relativement insensibles à un changement d'éclairage, il semble intéressant

d'étudier les conséquences d'un changement d'éclairage sur les degrés de connexité et d'homogénéité de sous-ensembles de pixels.

Tous les travaux cités dans ce mémoire s'appliquent à l'analyse d'images statiques. Un prolongement de nos travaux porte sur l'analyse de séquences d'images couleur pour la reconnaissance d'objets en mouvement. Cette reconnaissance nécessite la définition de signatures qui caractérisent le mouvement de l'objet et les propriétés spatio-colorimétriques des pixels le représentant dans les différentes images d'une même séquence. Des mesures spécifiques de comparaison de ces signatures caractérisant les séquences d'images couleur sont alors à développer.

Une des applications concerne la surveillance de zones sensibles de sites publics. La problématique est de retrouver les mêmes personnes dans différentes séquences d'images couleur saisies par des caméras observant ces différentes zones. Pour ce faire, nous proposons d'extraire des signatures qui caractérisent les personnes et qui permettent de les suivre au cours de leurs déplacements sous les différentes caméras. Par ailleurs, la comparaison des séquences d'images doit être la moins sensible possible aux conditions d'éclairage des différentes zones surveillées. Avec Jack-Gérard Postaire, nous travaillons actuellement sur ce problème en collaboration avec Louahdi Khoudour et Thomas Leclercq, tous deux respectivement chercheur et doctorant depuis 2002 au Laboratoire LEOST de l'INREST.

Ces quelques pistes de recherche sur lesquels porteront mes efforts à moyen terme montrent que l'exploitation de l'information couleur pour l'analyse d'images numériques est un axe de recherche qui soulève encore de nombreux problèmes à résoudre.

Publications choisies

Les publications écrites avec Aymeric Gillet, Vénérée Rakotomalala et Vincent Ultré et jointes à ce document, décrivent les contenus des travaux qui ne sont pas évoqués dans ce mémoire, et complètent ainsi notre présentation.

Chapitre de livre

1. A. Gillet, **L. Macaire**, C. Botte-Lecocq et J. G. Postaire. Chapitre "Color image segmentation by analysis of 3D histogram with fuzzy morphological filters", pages 154–177, Dans *Fuzzy filters for image processing- Studies in fuzziness and soft computing*, Authors of the book : M. Nachtgael, D. Van der Weken, D. Van De Ville and E.E. Kerre. Springer-Verlag Editor, 2002.

Articles dans des revues scientifiques avec comité de lecture

1. N. Vandenbroucke, **L. Macaire** et J. G. Postaire. "Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis.". *Computer Vision and Image Understanding*, 90(2) :190–216, 2003.
2. D. Muselet, **L. Macaire** et J. G. Postaire. "Color person image indexing and retrieval". *Machine Graphics and Vision*, 11(2/3) :257–283, 2002.
3. V. Rakotomalala, **L. Macaire**, J. G. Postaire et M. Valette. "Identification of retinal vessels by color image analysis". *Machine Graphics and Vision*, 7(4) :725–743, 1998.

Communication avec actes dans un congrès international

1. V. Ultré, **L. Macaire** et J. G. Postaire. "Determination of compatibility coefficients for color edge detection by relaxation". In *Proc. of the IEEE Int. Conf. on Image Processing*, volume 3, pages 1045–1048, Lausanne (Suisse), 1996.

Color image segmentation by analysis of 3D histogram with fuzzy morphological filters

A. Gillet, L. Macaire, C. Botte-Lecocq et J. G. Postaire

Chapitre du livre

Fuzzy filters for image processing- Studies in fuzziness and soft computing

Springer-Verlag Editor

pages 154-177 - 2002

Chapter 7

Color Image Segmentation by Analysis of 3D Histogram with Fuzzy Morphological Filters

Aymeric Gillet, Ludovic Macaire, Claudine Botte-Lecocq and Jack-Gerard Postaire

Laboratoire Automatique I3D
Universite Sciences et Technologies Lille 1
Cite scientifique - Bat P2 59 655 Villeneuve d'Ascq
France Tel : 00 33 - (0)3 20 43 41 69 Fax : 00 33 - (0)3 20 43 65 67
Ludovic.Macaire@univ-lille1.fr

Summary. This chapter presents a new color image segmentation method using fuzzy mathematical morphological filters applied to the 3D color histogram. Segmentation consists in detecting the different modes which are present in the 3D color histogram and associated with homogeneous regions. In order to detect these modes, we show how a color image can be considered as a fuzzy subset of pixels characterized by its membership function to a mode. This function is evaluated thanks to a concavity analysis of the 3D color histogram. Then, a fuzzy morphological transformation is applied to this membership function in order to enhance the modes. The effectiveness of our proposed fuzzy morphological approach is then illustrated with color images.

1 Introduction

A color is an additive combination of the three primary colors : red, green and blue. In color image processing, color features are evaluated from the stimuli R (red), G (green) and B (blue) [1]. The color of a pixel is usually given as three values corresponding to the (R, G, B) space.

Color image segmentation is defined as the partition of a color image into regions which are sets of connected pixels whose colors are homogeneous [2]. We consider that pixels, whose color feature levels are equal to R , G and B , are associated with points or observations in the three-dimensional color space (R, G, B) . We assume that all the pixels belonging to one region give rise in the color space to one cluster of points which can be considered as a pixel class. Furthermore, pixels which form several distinct regions of the same color, are assigned to the same pixel class. The color image segmentation method proposed in this paper is based on an original unsupervised pixel classification scheme by means of mode detection.

Many statistical procedures aim at mode detection by estimating the probability density function (*pdf*) underlying the distribution of the observa-

is completed either after a pre-specified maximum number of iterations or when any cluster can be dilated any more. The results depend both on the size of the structuring element used for morphological operations and on the choice of the color space.

Basic binary morphological filters have proved to be an efficient tool for cluster detection when the cluster shapes are rather spherical [4]. But when there is no a priori knowledge about the shapes of the clusters to be detected, it is interesting to introduce some adaptivity in the morphological filters [19].

One problem with traditional clustering techniques is that there are only two values, either 1 or 0, to indicate the membership degree of each observation to a cluster. This requires well-defined boundaries between the clusters, which is not the usual case for real images. The uncertainties arising in image segmentation can be solved by using the fuzzy set theory.

Fuzzy C-Means is a method which allows ambiguous boundaries between clusters, and has received much attention [8]. This iterative optimization method evaluates the membership degree of an observation to each of the clusters by means of a fuzzy distance between this observation and each of the cluster centers. At each iteration, these cluster centers are updated and an objective criterion is used to minimize the fuzzy intra-cluster distance.

During the coarse stage of a “coarse to fine” segmentation algorithm, Lim analyses the 1D color feature histograms in order to detect the cluster centers [10]. During the fine stage, he uses the Fuzzy C-Means method to assign the pixels which were not classified during the coarse stage to the closest detected cluster. The applicability of this technique, which tries to solve the problem of cluster validity by using a coarse to fine concept, is limited by the 1D histograms thresholding. The fuzzy C-Means method requires the a priori knowledge of the number of clusters to be detected, and its computational cost is quite high with large data sets.

For color image segmentation purpose, Pham proposes the mountain clustering which uses the fuzzy integral as a distance measure [13]. This method does not require initial estimates of cluster centers. Nevertheless, segmentation results depend on two main parameters which are difficult to adjust, namely the spatial resolution of the image, which specifies the degree of detail in the segmentation process, and an integration threshold.

In this paper, we propose to define fuzzy morphological erosion and dilation filters specifically adapted to the detection of the modes of the 3D color histogram. For this purpose, we describe in the second section how to extract, from the 3D histogram, a fuzzy subset defined by its membership function. In the third section, we present a fuzzy morphological transformation which is applied to this membership function in order to enhance the different modes. In the last section, we describe the pixel classification scheme which analyses the detected modes in order to yield the segmented image. Extensive experiments are carried out on color images in this section and the effectiveness of our color image segmentation method is assessed.

2 Fuzzy subset extracted from the 3D color histogram

2.1 Pdf approximation

As we have mentioned in the first section, many clustering procedures are based on the detection of the modes of the approximated underlying *pdf*. In our context, we propose to approximate this *pdf* by the 3D color histogram.

We consider that a pixel, whose color features are (R, G, B) , is associated with an observation X , whose coordinates are (R, G, B) in the 3D color space. All the observations X , which represent the different colors in the image, form the set Δ .

The color histogram is composed of bins which indicate the number $HISTO(X)$ of pixels with the corresponding (R, G, B) color features. This histogram, which is normalized between 0 to 1, can be considered as an approximation of the underlying *pdf*. For the segmentation purpose, the key problem is to assign the observations of Δ to the different classes identified by means of the detected modes.

This clustering problem is considered as a fuzzy decision process in order to cope with uncertainties in color quantization. Indeed, in the case of a real color image, the boundaries between the color clusters associated to the different regions constituting the image cannot be considered as well defined surfaces as when using the crisp decision theory.

Hence, we propose to implement a fuzzification phase. This step quantifies the degree with which each observation X belongs to a class associated with a mode. It allows to evaluate the confidence degree in the event “ X belongs to a class associated to a detected mode”. For the sake of simplicity, this event will be referred as “ X belongs to a mode” in what follows. This fuzzification requires the transformation of the crisp observation set Δ into a fuzzy subset M characterized by a membership function. This function, denoted μ_M , is evaluated thanks to a convexity analysis of the color histogram.

2.2 Membership function evaluation

The membership function defining the fuzzy subset M can be evaluated by considering that an observation X belongs to a mode if the *pdf* approximated at X is concave [14]. A convexity test locally assigns a “concave” or “convex” label to each observation of the color space. The local convexity of the approximated *pdf* at an observation X is evaluated by analyzing the variation of this *pdf* estimator when an observation domain grows around X . So, we first evaluate the estimator $p^1(X)$ of the *pdf* using a cubic observation domain D_X^1 which contains neighbors of X , and then, we compute another estimator $p^2(X)$ using the cubic observation domain D_X^2 , slightly larger than D_X^1 , such as

$$p^1(X) = \frac{\sum_{Y \in D_X^1} HISTO(Y)}{(2.m+1)^3} \quad (1)$$

and

$$p^2(X) = \frac{\sum_{Y \in D_X^2} HISTO(Y)}{(2.n+1)^3}. \quad (2)$$

The parameters $(2.m+1)^3$ and $(2.n+1)^3$ are respectively the volumes of the first and second observation domains, with $n > m$. By experiment, the sizes $(2.m+1)$ and $(2.n+1)$ are respectively taken equal to 3 and 5. In these conditions, if $p^2(X) < p^1(X)$, the approximated *pdf* is considered as locally concave and the observation X is labelled as “concave”, otherwise X is labelled as “convex”.

Now, we introduce a Π -fuzzification function which provides the confidence degree of the event “ X is a modal observation”. This degree depends on $\frac{k^1(X)}{(2.m+1)^3}$, where $k^1(X)$ is the number of concave observations which are included in D_X^1 . If the labels of most neighbors of X included in D_X^1 are concave (respectively convex), X is considered as a concave (respectively convex) observation, i.e as a modal (respectively non-modal) observation whose membership degree is close to 1 (respectively close to 0). So, the membership function μ_M is defined as:

$$\begin{aligned} &\text{if } \frac{k^1(X)}{(2.m+1)^3} \leq a, \quad \text{then } \mu_M(X) = 0, \\ &\text{if } \frac{k^1(X)}{(2.m+1)^3} > b, \quad \text{then } \mu_M(X) = 1, \\ &\text{if } a < \frac{k^1(X)}{(2.m+1)^3} < b, \text{ then} \end{aligned}$$

$$\mu_M(X) = \frac{1}{2} + \frac{1}{2} \cdot \sin\left(\frac{\pi}{b-a} \cdot \left(\frac{k^1(X)}{(2.m+1)^3} - \frac{a+b}{2}\right)\right). \quad (3)$$

In this chapter, the values of the parameters a and b are respectively taken equal to 0.2 and 0.8, so that we consider that an observation belongs to a mode if 80% of the neighboring observations in D_X^1 are labelled as “concave”. After this fuzzification phase, M is well defined by its membership function μ_M . However, even with this fuzzification scheme, experiments have shown that some clusters, associated with weak modes, tend to be ignored in the final clustering scheme. To cope with this limitation of the procedure, we propose to apply fuzzy morphological filters to the membership function μ_M in order to enhance the modes of M .

3 Mode detection by fuzzy morphological filters

The basic morphological filters are erosion and dilation [7]. The main effect of erosion is to enlarge the valleys, by eliminating the irregularities of the distribution, but this transformation also tends to shrink the modes. On the other hand, dilation is used to enlarge the modes, but it also tends to fill the valleys [16]. So, we propose to transform the membership function μ_M by means of fuzzy morphological filters which take advantage of these two basic operations without their drawbacks.

The aim of this procedure is to increase the contrast between modes and valleys. High values of μ_M correspond to observations X belonging to the modes of M while low values are associated with observations standing in the valleys between the modes. If an observation X is close to a mode, it is relevant to increase μ_M , i.e to enhance the dilation effect at the location of this observation and to limit the erosion effect in order to preserve the mode. Conversely, if the observation is far from all modes, it is probably located in a valley. In this case, it is useful to decrease μ_M , i.e to erode the membership function at the location of this observation and not to dilate it. So it would be pertinent to erode or to dilate the membership function more or less depending on the location of each observation in the color space. This location can be introduced in the definition of the structuring functions used by both fuzzy erosion and dilation filters.

In this section, after presenting the classical definition of fuzzy erosion and dilation filters, we describe new fuzzy morphological filters which include structuring functions specifically adapted to mode enhancement. Finally, we describe the morphological transformation which enhances the modes of M .

3.1 Classical fuzzy erosion and dilation filters

Among the available fuzzy erosion and dilation filters, we have selected the most classical one, which according to Bloch, corresponds to the greater fuzzy effect of dilation and erosion [3].

The chosen classical fuzzy erosion and fuzzy dilation filters at an observation X is defined as :

$$E_\gamma[\mu_M(X)] = \min_{Y \in S_X} (\max[\mu_M(Y), 1 - \gamma(X - Y)]) \quad (4)$$

$$D_\gamma[\mu_M(X)] = \max_{Y \in S_X} (\min[\mu_M(Y), \gamma(X - Y)]), \quad (5)$$

where S_X is the fuzzy structuring element centered at the observation X and defined by its membership function γ called the structuring function. This structuring function used by the erosion and dilation depends on the distances between X and each of the neighboring observations Y in S_X .

For the purpose of illustration, let us consider the 1D membership function μ_M of Fig. 1 and the 1D structuring function γ which is set to 1, for all the neighboring observations Y in S_X , the structuring element of size 3. The eroded and dilated membership functions $E_\gamma[\mu_M]$ and $D_\gamma[\mu_M]$ are presented in Figs. 2 and 3 respectively, where μ_M is represented by the dotted graph. This example shows that the fuzzy erosion tends to enlarge the valleys but tends to shrink the modes, whereas the fuzzy dilation tends to enlarge the modes while filling the valleys.

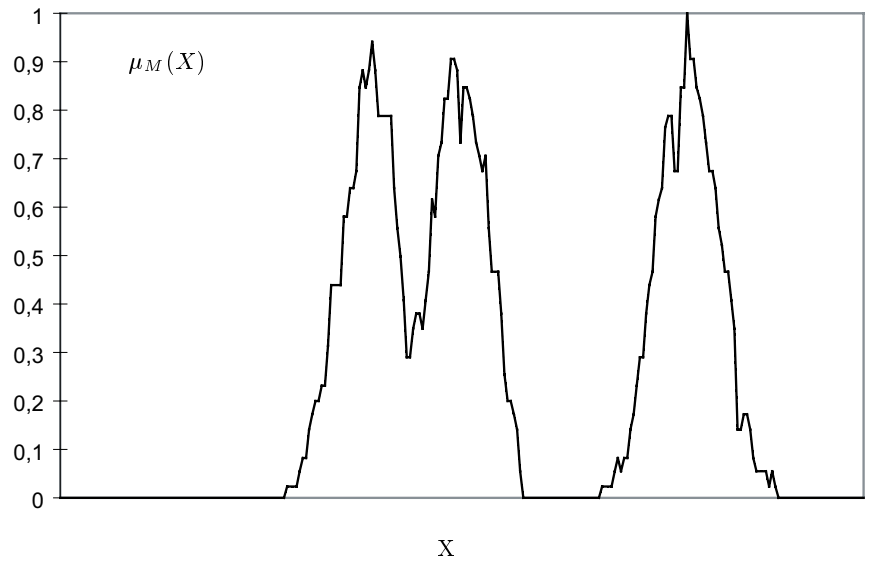


Fig. 1. Original 1D membership function μ_M .

For the purpose of mode detection, it would be pertinent to define new fuzzy morphological filters using different structuring functions for erosion and dilation.

3.2 Fuzzy erosion and dilation filters for mode detection

In the classical definitions Eqs. (4) and (5), the structuring function γ depends on the distance between the considered observation X and each of its neighbors Y , but does not depend on the membership degree of the neighbors. If a neighboring observation is characterized by an high membership degree, it would be interesting to take it into account only for the dilation purpose at X , but not for the erosion. On the other hand, if the membership degree of a neighbor is low, its influence should be higher for erosion than for dilation at X .

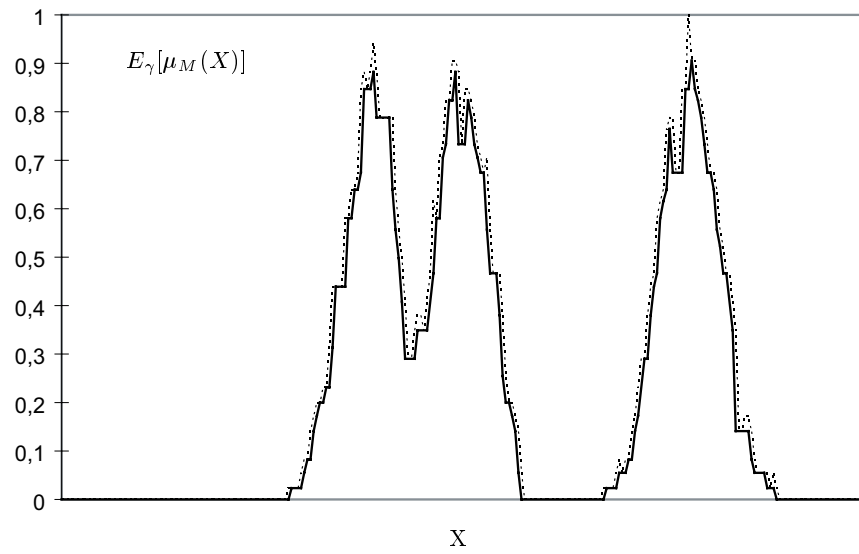


Fig. 2. Result of the 1D fuzzy erosion $E_\gamma[\mu_M]$.
Size of $S_X = 3$.

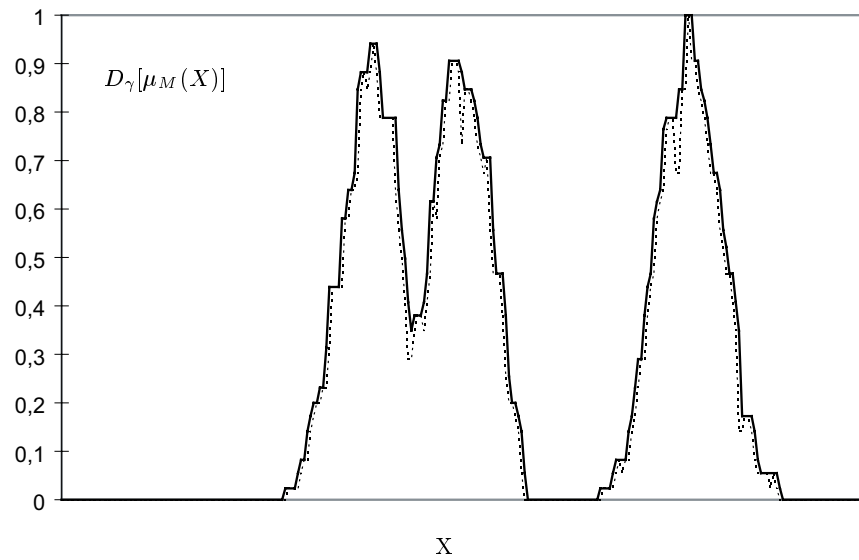


Fig. 3. Result of the 1D fuzzy dilation $D_\gamma[\mu_M]$.
Size of $S_X = 3$.

In order to erode and to dilate depending on the membership degrees of the neighbors in S_X , we define new morphological filters associated with the fuzzy structuring element S_X :

$$E'_{\gamma_E}[\mu_M(X)] = \min_{Y \in S_X} (\max[\mu_M(Y), 1 - \gamma_E(Y)]) \quad (6)$$

$$D'_{\gamma_D}[\mu_M(X)] = \max_{Y \in S_X} (\min[\mu_M(Y), \gamma_D(Y)]), \quad (7)$$

where γ_E and γ_D are the structuring functions respectively defined for fuzzy erosion and dilation and described in Sect. 3.3. These functions do not depend on the distance between the observation X and each of its neighbors Y in S_X but on the membership degrees of these neighbors.

In order to detect the modes, the structuring functions γ_E and γ_D have to be defined so that the erosion enlarges the valleys without shrinking the modes, while the dilation enlarges the modes without filling the valleys.

3.3 Structuring functions adapted to mode detection

In the color space, a neighboring observation Y belonging to the structuring element S_X can be located in a mode, in a valley, or in the border between a mode and a valley. If this observation is close to a mode, its membership degree $\mu_M(Y)$ is high, whereas if Y is close to a valley, its membership degree is low. Furthermore, a neighbor Y which is close to the border between a valley and a mode, is characterized by significant local spatial variations between its membership degree and the membership degrees of its neighbors. The combination of these two criteria is well suited to decide if an observation is located in a mode, a valley or a border. We present now how to introduce these criteria into the structuring functions for mode enhancement.

Structuring function associated with the fuzzy erosion. In order to detect the modes, we have to erode the valleys while not deepening the modes. This fuzzy erosion is performed with a structuring element S_X , centered at each observation X and using a structuring function γ_E defined for each of the neighboring observations Y . An observation X is considered all the more as belonging to a valley (respectively a mode) as its neighbors are also considered as belonging to a valley (respectively a mode). In this case, the value of the structuring function at each neighbor Y belonging to a valley will be set to 1, so that its effect will be high for the erosion operation. If the neighbor Y is located in a mode, its effect on the erosion will have to be very low.

To meet these requirements, we propose for fuzzy erosion expressed by Eq. (6), the structuring function defined as:

$$\begin{aligned} \gamma_E(Y) &= 1 && \text{if } [\mu_M(Y) \cdot (1 - g_{\mu_M}(Y))] \leq TH, \\ \gamma_E(Y) &= \mu_M(X) && \text{otherwise,} \end{aligned} \quad (8)$$

where TH is a threshold and $g_{\mu_M}(Y)$ is an approximation of the local spatial variations of $\mu_M(Y)$, evaluated as the mean of the local differences between the membership degree of Y and the membership degrees of its neighbors. To be more specific, $g_{\mu_M}(Y)$ is the normalized mean absolute difference between the membership degree of Y and the membership degrees of the neighboring observations which belong to the domain D_Y^1 centered at Y :

$$g_{\mu_M}(Y) = \frac{\sum_{Z \in D_Y^1} |\mu_M(Y) - \mu_M(Z)|}{(2.m + 1)^3}. \quad (9)$$

If the expression $[\mu_M(Y) \cdot (1 - g_{\mu_M}(Y))]$ is lower than the threshold TH , a neighbor Y is considered as being located in a valley or in a border close to a valley. In this case, $\gamma_E(Y)$ is set to 1 so that the effect of the neighbor Y on the fuzzy erosion at the observation X is the highest one. If $[\mu_M(Y) \cdot (1 - g_{\mu_M}(Y))]$ is higher than the threshold TH , Y is considered as being located in a modal observation or in a border close to a mode. In this case, $\gamma_E(Y)$ is set to $\mu_M(X)$, so that this neighboring observation Y has no effect on the fuzzy erosion at the observation X .

When considering all the neighboring observations Y in the structuring element S_X , the erosion of the membership function μ_M is performed with only the neighboring observations which are located in a valley or in a border close to a valley. This erosion using the structuring function γ_E tends to deepen the valleys and to preserve the modes.

Structuring function associated with the fuzzy dilation. In the same way, the structuring function associated with the fuzzy dilation expressed by Eq. (7) can be defined as :

$$\begin{aligned} \gamma_D(Y) &= 1 && \text{if } [\mu_M(Y) \cdot (1 - g_{\mu_M}(Y))] > TH, \\ \gamma_D(Y) &= \mu_M(X) && \text{otherwise.} \end{aligned} \quad (10)$$

If $[\mu_M(Y) \cdot (1 - g_{\mu_M}(Y))]$ is higher than TH , a neighbor Y , which belongs to the structuring element S_X , can be considered as being located in a mode or in a border close to a mode. Its effect on the fuzzy dilation at the observation X is the highest one. Conversely, if $[\mu_M(Y) \cdot (1 - g_{\mu_M}(Y))]$ is lower than TH , the observation Y is considered as being located in a valley or in a border close to a valley, and has no effect on the fuzzy dilation.

When considering all the neighbors Y in the structuring element S_X , the fuzzy dilation takes only into account the neighbors Y which are located in or near a mode. Dilating the membership function μ_M using this structuring function γ_D tends to enhance the modes without filling the valleys.

Figures 4 and 5 display the fuzzy eroded and dilated membership functions of Fig. 1 respectively, when the size of S_X is set to 3. The threshold TH is set to $0.5(1 - 0.1) = 0.45$, as to consider that an observation Y is located in a valley or in a border close to a valley if its membership degree is less than 0.5 and its local membership degree variation is higher than 0.1. Figure 4 shows that the membership function is only eroded at the observations X which are located in the valleys. Conversely, Fig. 5 shows that μ_M is dilated only at the observations which are located near the modes.

These results can be compared with those obtained by means of the classical morphological filters and illustrated by Figs. 2 and 3. We can see that the fuzzy classical erosion tends to shrink the modes while the proposed fuzzy erosion tends to only deepen the valleys. Furthermore, the classical fuzzy dilation tends to fill the valleys while the proposed fuzzy dilation tends to only enhance the modes. This example shows the improvement achieved with the proposed fuzzy morphological filters with respect to the using of the classical ones for mode enhancement.

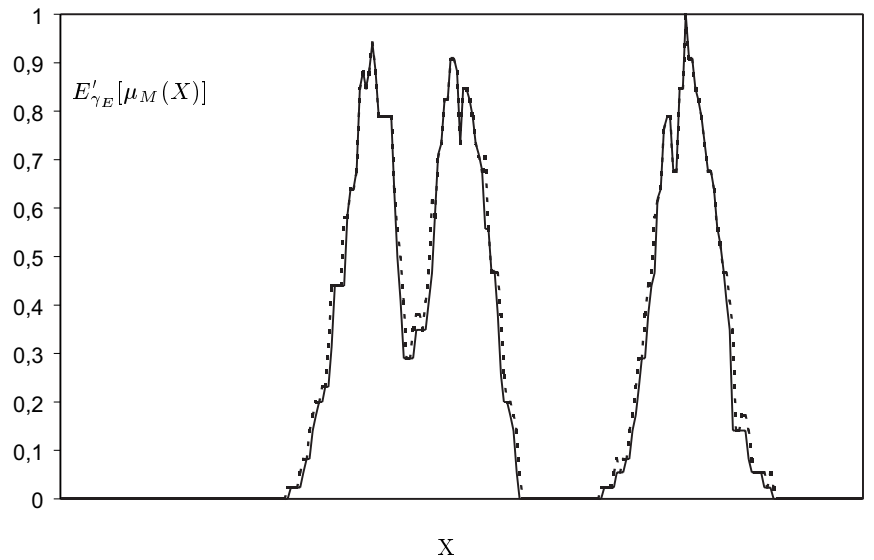


Fig. 4. Result of the 1D fuzzy erosion $E'_{\gamma_E}[\mu_M]$. Size of $S_X = 3$ and $TH = 0.5(1 - 0.1) = 0.45$.

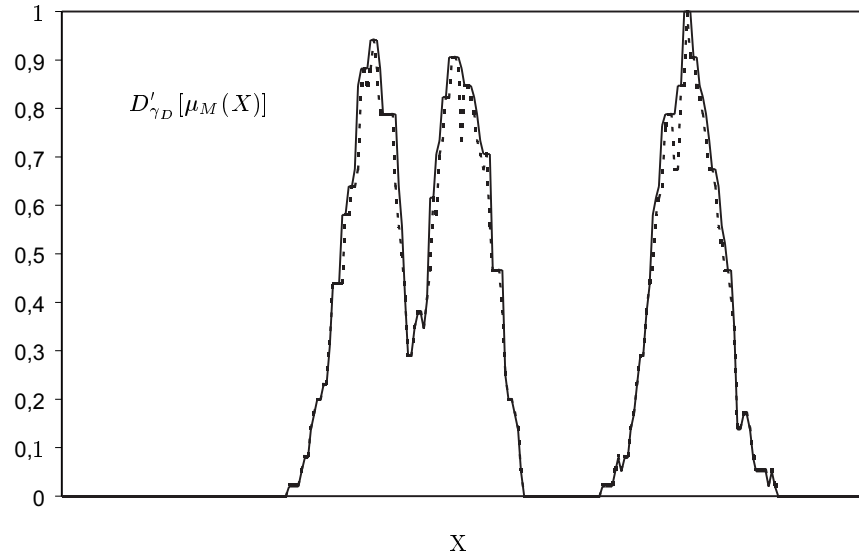


Fig. 5. Result of the 1D fuzzy dilation $D'_{\gamma_D}[\mu_M]$.
Size of $S_X = 3$ and $TH = 0.5(1 - 0.1) = 0.45$.

3.4 Fuzzy morphological transformation for mode extraction

In order to take advantage of the two fuzzy filters defined above, we propose to combine them into a fuzzy morphological transformation, denoted t , which performs a fuzzy erosion of the fuzzy subset M using the structuring function γ_E , followed by a fuzzy dilation of the resulting fuzzy subset using the structuring function γ_D [6].

This transformation yields the fuzzy subset M^t , characterized by its mode membership function μ_M^t , defined as :

$$\mu_M^t = t(\mu_M) = D'_{\gamma_D}[E'_{\gamma_E}(\mu_M)]. \quad (11)$$

Figure 6 presents the resulting membership function μ_M^t , displayed as a solid graph, of the function μ_M presented in Fig. 1 and displayed as a dotted graph. As required, the modes are enhanced while the main valleys are not filled. However, the effect of this transformation is rather weak. Hence, we propose to iterate it until the resulting mode membership function μ_M^T is stabilized, such as:

$$\mu_M^T = (\mu_M^t)^\infty \text{ with } (\mu_M^t)^j = t[(\mu_M^t)^{j-1}] = D'_{\gamma_D}[E'_{\gamma_E}[(\mu_M^t)^{j-1}]], j = 1, \dots, \infty. \quad (12)$$

The initial conditions are $(\mu_M^t)^0 = \mu_M$.

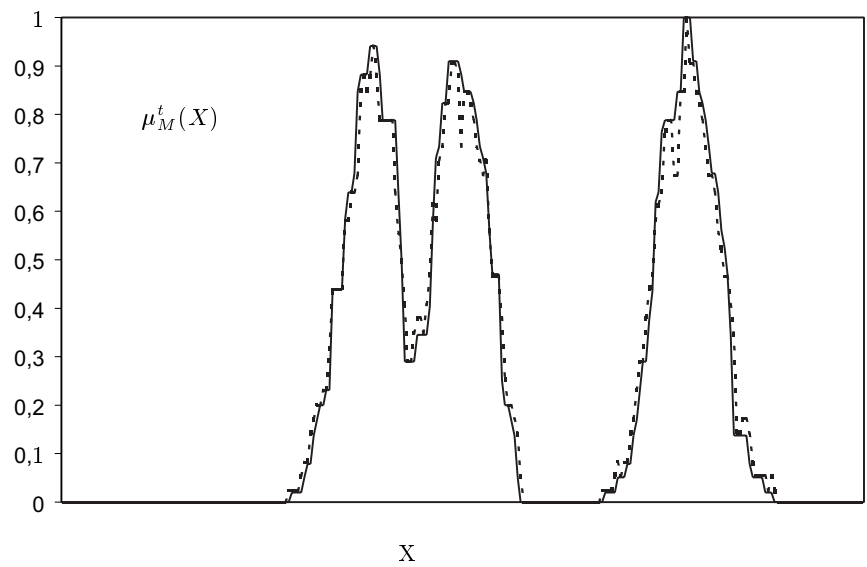


Fig. 6. Result of the transformation μ_M^t .
Size of $S_X = 3$ and $TH = 0.5(1 - 0.1) = 0.45$.

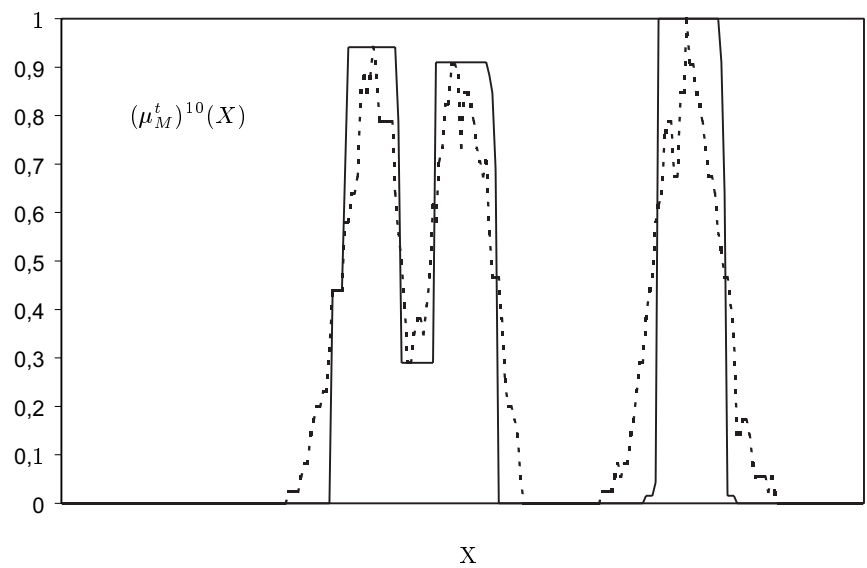


Fig. 7. Result of the transformation $(\mu_M^t)^{10}$.
Size of $S_X = 3$ and $TH = 0.5(1 - 0.1) = 0.45$.

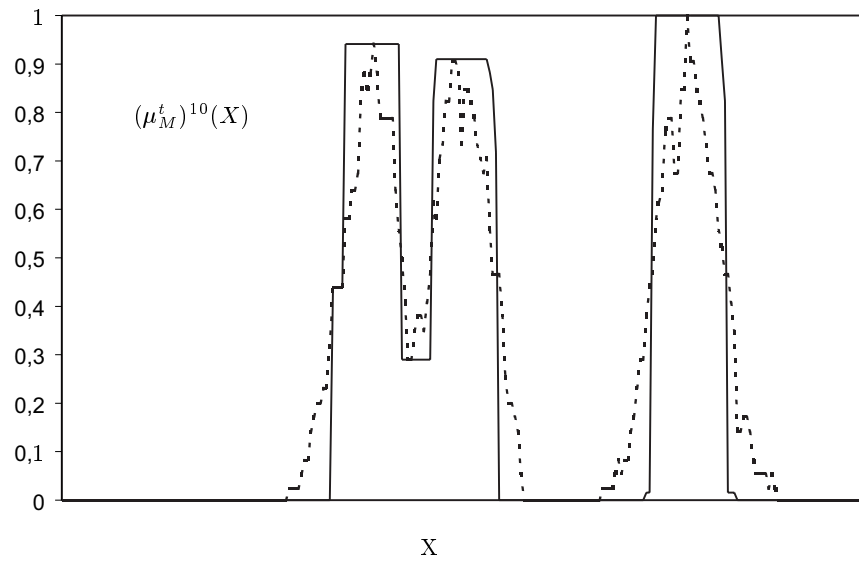


Fig. 8. Result of the transformation $(\mu_M^t)^{10}$.
Size of $S_X = 3$ and $TH = 0.5(1 - 0.2) = 0.40$.

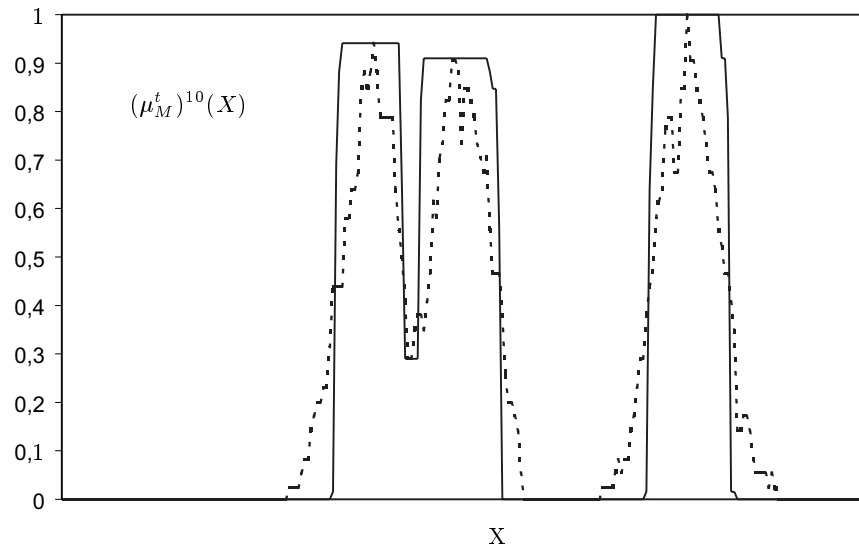


Fig. 9. Result of the transformation $(\mu_M^t)^{10}$.
Size of $S_X = 3$ and $TH = 0.5(1 - 0.3) = 0.35$.

Figure 7 shows the resulting membership function $(\mu_M^t)^{10}$ after 10 iterations. The three modes are well enhanced and it is important to notice that the higher and lower membership degrees associated to the modes and the valleys are preserved. Indeed the fuzzy dilation performed at the observation whose membership degree is the highest, propagates this high degree to its neighbors which belong to the structuring element centered in this observation and which are considered as being located in this mode or a border. Conversely, the proposed erosion performed at the observation whose degree is the lowest, propagates this low degree to the neighbors which belong to the structuring element centered in this observation and which are considered as being located in this valley or in a border.

The mode enhancement procedure is sensitive to the choice of the parameter TH used to decide if the neighboring observations act on the morphological filters.

Figures 7, 8, 9 display the transformed membership function $(\mu_M^t)^{10}$ with different values of TH , which are $0.5(1 - 0.1) = 0.45$, $0.5(1 - 0.2) = 0.40$ and $0.5(1 - 0.3) = 0.35$ respectively. These figures illustrate that lower the parameter TH is, the wider the modes are and the narrower the valleys are. This behavior of the procedure results from a relationship between the value of TH and the number of neighbors of X taken into account by the filtering process. The lower TH is, the stronger the effect of the dilation is.

The modes are easily detected thanks to the defuzzification of the transformed membership function μ_M^T , which is based on the convexity test described in the second section.

The local convexity of the transformed membership function μ_M^T at an observation X is evaluated by analyzing its variation when an observation domain grows around X . So, we first evaluate the estimator $p_*^1(X)$ of μ_M^T using the cubic observation domain D_X^1 , and then, we compute another estimator $p_*^2(X)$ using the cubic observation domain D_X^2 , slightly larger than D_X^1 , such as:

$$p_*^1(X) = \frac{\sum_{Y \in D_X^1} \mu_M^T(Y)}{(2.m + 1)^3} \quad (13)$$

and

$$p_*^2(X) = \frac{\sum_{Y \in D_X^2} \mu_M^T(Y)}{(2.n + 1)^3}. \quad (14)$$

As for the fuzzification phase by means of convexity test, we set the sizes $(2.m + 1)$ and $(2.n + 1)$ to 3 and 5 respectively. In these conditions, if $p_*^2(X) < p_*^1(X)$, the transformed membership function μ_M^T is considered as locally concave and so we decide that the observation X belongs to a mode.

4 Experimental results

In order to illustrate the behavior and the efficiency of the above mode detection procedure, we propose to segment the classical real image of the “house” (see Fig. 10) and one synthesis image (see Fig. 16). The segmentation of the house image is quite challenging because it is difficult to extract regions which take into account the shadow effect. Furthermore, some regions such as the roof and the walls present local color non-homogeneities.



Fig. 10. The “house” color test image.

As it is difficult to display the values of the bins of the 3D color histogram, we present the three 1D color feature histograms (see Fig. 11). They show that it is not easy to discriminate the pixel classes which correspond to the different regions.

In order to extract the modes from the analysis of the 3D color histogram of the image of Fig. 10, we evaluate the membership function associated with the fuzzy subset M . Figures 12 and 13 present the support of M (defined by $\mu_M(X) \neq 0$), and the kernel of M (defined by $\mu_M(X) = 1$), respectively.

We can see, in Fig. 12, that the different modes corresponding to the pixel classes cannot be easily discriminated using only the support of M . Furthermore, the kernel of M only contains three main modes and some insignificant ones (see Fig. 13). These two figures show that for this image, the fuzzification step is not sufficiently discriminating and does not provide well defined and separated modes.

The mode enhancement is achieved thanks to the iteration of the fuzzy morphological transformation t defined by Eq. 11. The threshold TH is set to $[0.5(1 - 0.1)] = 0.45$ and the size of the structuring element is set to 3. After 3 iterations, the resulting membership function μ_M^T is stabilized and the modes

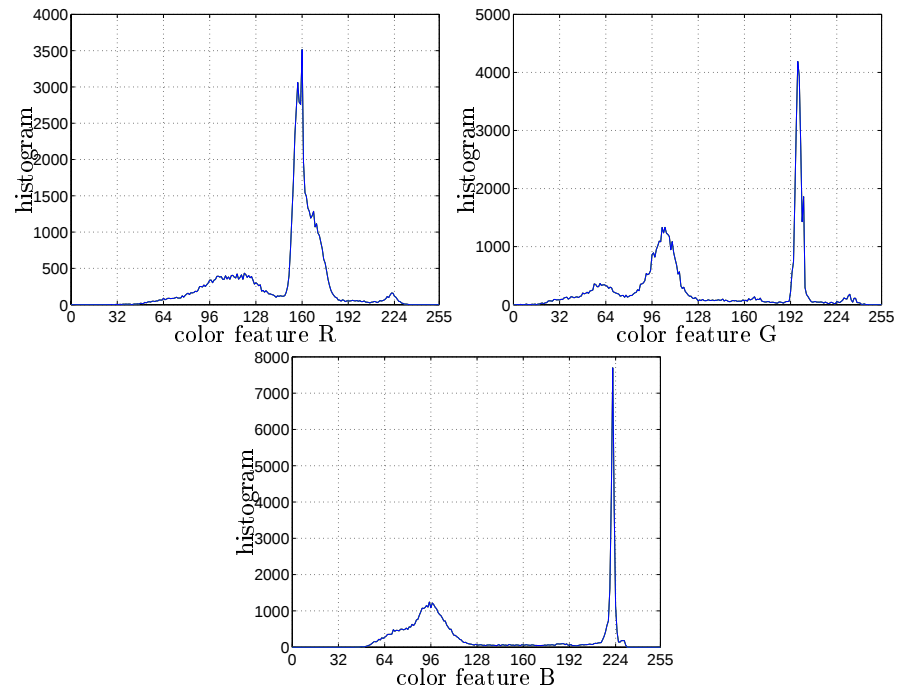


Fig. 11. 1D Color feature histograms of the image of Fig. 10.

are well enhanced. We propose to extract them through the defuzzification step. Figure 14 presents the six different modes which are extracted by means of our method. The modes are displayed as the subsets of observations that belong to each of them. A class of pixels is associated to each detected mode and is characterized by the gravity center of all the observations that belong to this mode. Each pixel of the analyzed color image is assigned to the class whose center is the nearest of the associated observation.

The pixels of the image of Fig. 15 are labelled with the colors defined by the coordinates of the centers of the classes to which they are assigned. This figure shows that the detected modes correspond to pixel classes which are not equiprobable. Indeed, the class of pixels which represents the wall of the house has a much larger population than the class which represents one of the windows. We can say that the number of classes determined by our scheme is relevant because the regions of the image of Fig. 10, such as the wall without the shadow, the wall with the shadow, the roof and the windows are well identified in the segmented image.

In order to evaluate the robustness of our approach, we propose to experiment a performance evaluation. It is based on the measure of the difference between the result of the segmentation and a reference segmentation. In order to make this reference segmentation available, we use the synthetic image

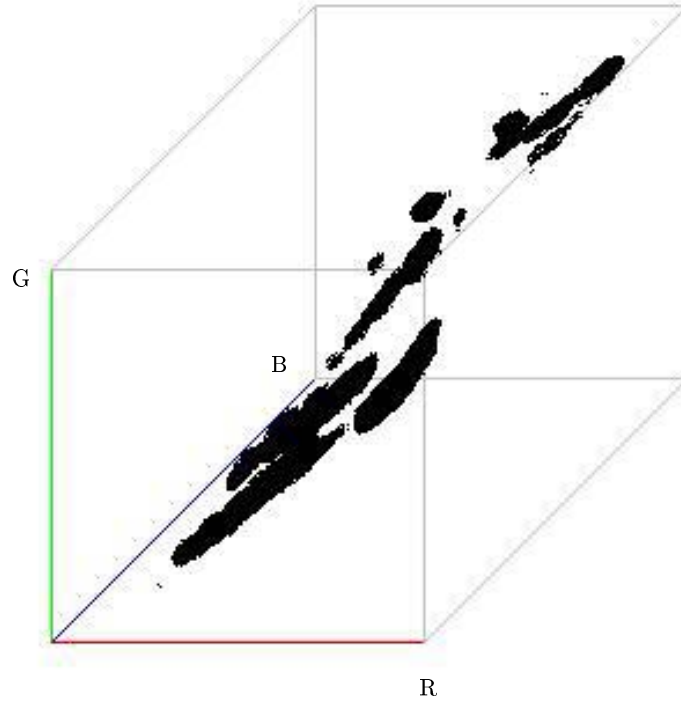


Fig. 12. Support of M extracted from the image of Fig. 10.

of Fig. 16. This synthetic image is composed of the 5 following regions with different shapes:

- a brown background R_1 ,
- an orange large square R_2 ,
- a red irregular shape R_3 ,
- a disc which is composed of 2 concentric regions:
 - a peripheral green ring R_4 and
 - a slightly darker green disc R_5 .

The regions R_4 and R_5 are built so that their red mean values are equal and the differences between their green and blue mean values are equal to 10 and 20 levels, respectively.

In order to analyze the reliability of the mode detection procedure, we evaluate the effects of the noise on the number of detected modes. For this purpose, the synthetic image of Fig. 16 is corrupted by a non-correlated gaussian noise with a standard deviation σ which is independently added to each of the three color features. Figure 17 shows the number of detected

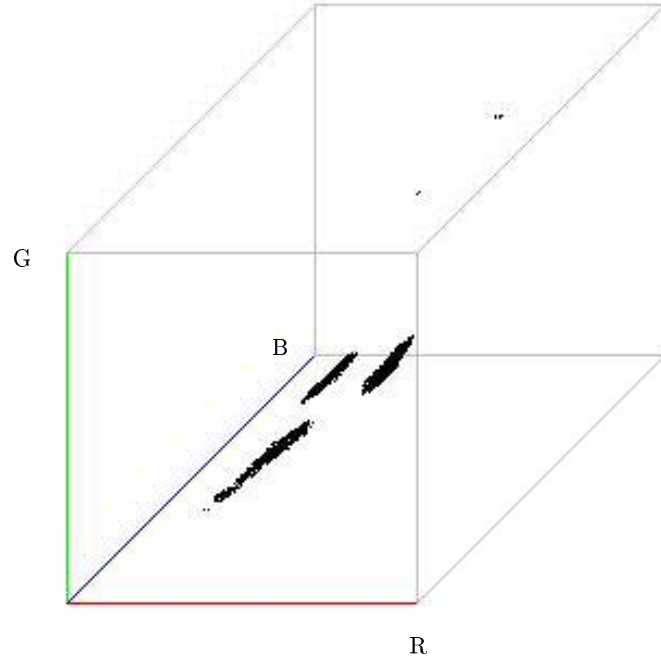


Fig. 13. Kernel of M extracted from the image of Fig. 10.

modes when the standard deviation ranges from 3 to 13, for a structuring element size set to 3. When σ is lower than 6, the number of detected modes corresponds exactly to the number of classes present in the pixel population of the image of Fig. 16. When σ is higher than 6 and lower than 13, only four modes are detected because the two modes corresponding to the classes associated to the regions R_4 and R_5 , i.e the two circular concentric regions, are identified as a single one. In this case, the pixels of the two circular regions are assigned to the same class. When σ is higher than 12, the noise is so strong that the classes corresponding to R_2 , R_4 and R_5 are merged into one detected mode.

In order to provide some insight into the behavior of the procedure, we consider the image of Fig. 16 corrupted for σ equal to 6 (see image of Fig. 18). Furthermore, by examining the three 1D color feature histograms presented by Fig. 19, we see that the distributions of the color features overlap. In order to detect the modes, we apply our approach with the same parameter values than the previous example. Figure 20 shows the five detected modes for σ equal to 6, which corresponds to the highest noise level for which the

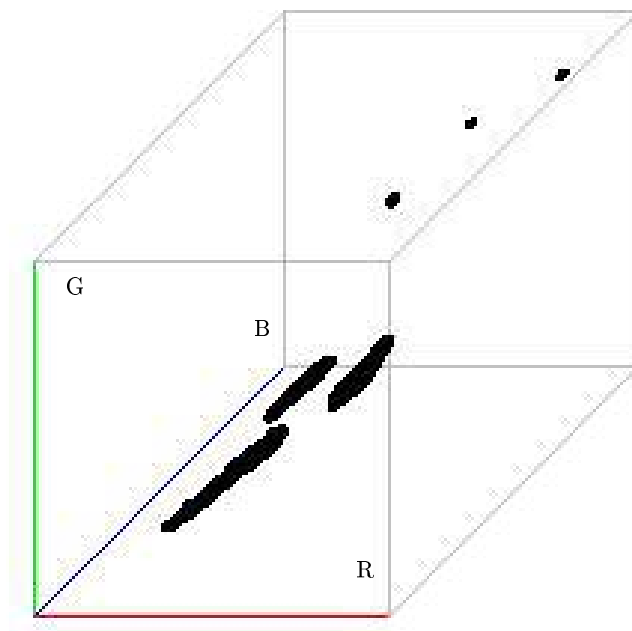


Fig. 14. Modes detected from the image of Fig. 10.
Size of $S_X = 3 \times 3 \times 3$ and $TH = 0.5(1 - 0.1) = 0.45$.



Fig. 15. Image of Fig. 10 segmented with the detected modes of Fig. 14.

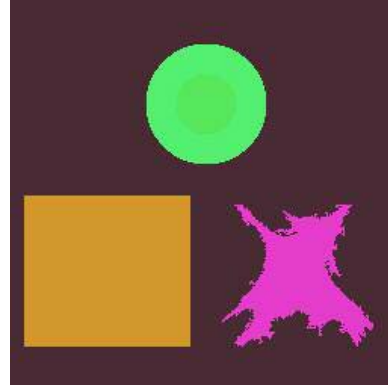


Fig. 16. Original synthetic image.

five modes are correctly identified. As for the real image, these modes are displayed as the subset of observations that belong to each of them.

The image of Fig. 21 shows how the pixels of the image of Fig. 16 are assigned to the five classes that correspond to the five regions of the original image. This result shows that our method is able to handle unequiprobable classes of pixels.

All the pixels of the regions R_1 , R_2 and R_3 have been correctly classified. Some misclassified pixels appear into regions R_4 and R_5 . However the error rate computed as the ratio of misclassified pixels to the total number constituting the two circular regions is equal to 3%.

The preceding evaluation of the mode stability with respect to the noise level has been performed with a structuring element size set to 3. We now evaluate the influence of this parameter.

Figure 17 shows that when the size of S_X is set to 5, any mode is detected until σ reaches 5. In this case, the size of the structuring element is so large that the erosion filtering erases all the modes of the membership function.

In conclusion, we have seen that the mode detection depends on two main parameters, namely the threshold TH and the size of the structuring element. TH is used for deciding if each neighbor of the center of the structuring element is taken or not into account for the erosion or dilation purpose. The examples of the third section (see Figs. 7, 8 and 9) show that the lower TH is, the stronger the effect of dilation is. If TH is set to a too lower value, the dilation filtering tends to merge neighboring modes. Finally, this section shows that the size of the structuring element deserves a great attention, because the erosion filtering with a too large structuring element whose size is large tends to erase the modes of the membership function.

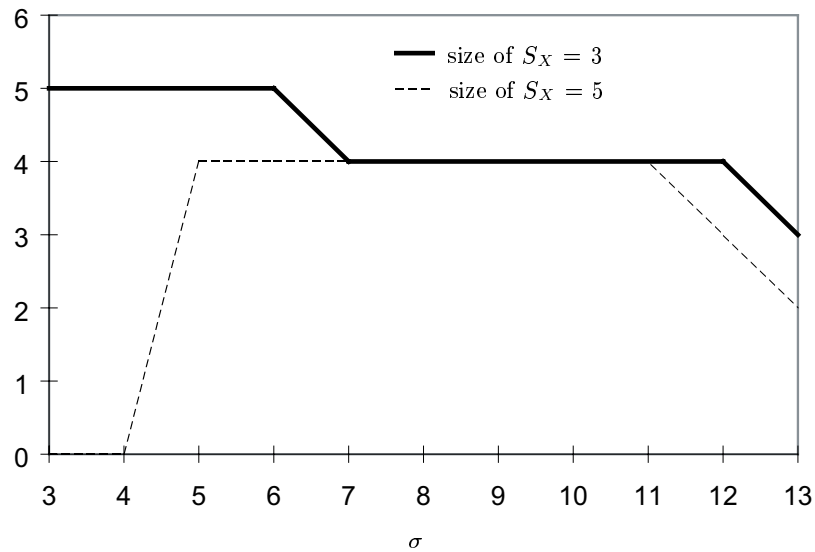


Fig. 17. Detected mode number with different noise standard deviations and structuring element sizes.

$$TH = 0.5(1 - 0.1) = 0.45.$$

5 Conclusion

In this paper, color image segmentation has been considered as a pixel classification problem, based on the detection of the modes of the 3D color histogram. Our scheme consists in associating each homogeneous region with a mode of the color histogram. This histogram is used to approximate the

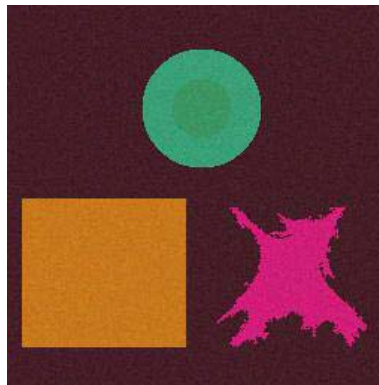


Fig. 18. Image of Fig. 16 corrupted with a gaussian noise for $\sigma = 6$.

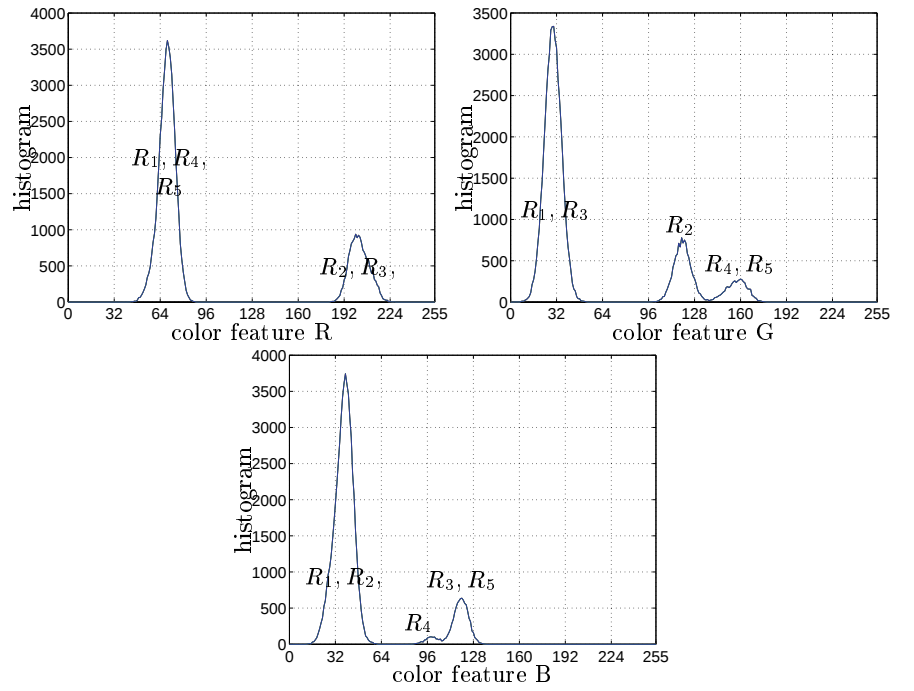


Fig. 19. 1D Color feature histograms of the image of Fig. 18.

underlying *pdf* whose concavity analysis yields a fuzzy subset defined by its membership function. A fuzzy morphological transformation, based on a combination of specific fuzzy erosions and dilations, is iteratively applied to this membership function. The pixels of the original color image are finally assigned to the classes associated with the detected modes.

This work leaves several possibilities of further improvement. First, the color histogram is a tool for approximating the *pdf*, which considers only the color properties of the pixels. The approximation of the *pdf* should be more pertinent in the context of image segmentation, if it was based on the co-occurrence matrices which take into account the colorimetric properties as well as the spatial interaction between the pixels. The second point concerns the evaluation of the structuring functions which should integrate a measure of the compatibility between the localization of the considered observation and those of its neighboring observations since the confidence in the assumption that an observation stands in a mode (respectively a valley) is an increasing function of the number of its neighbors that also stand in a mode (respectively a valley).

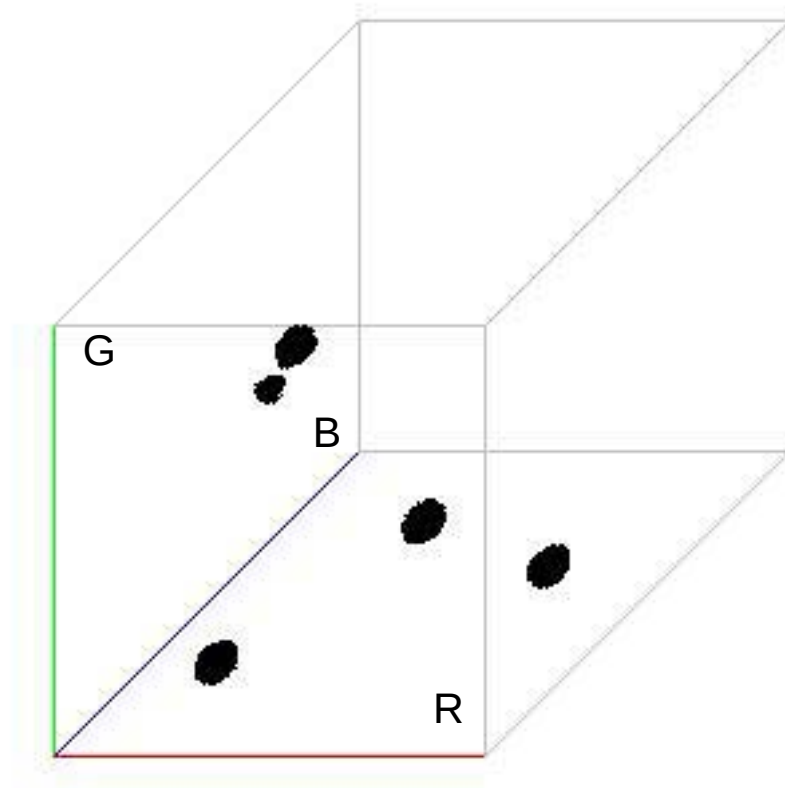


Fig. 20. Modes detected from the image of Fig. 18.
Size of $S_X = 3$ and $TH = 0.5(1 - 0.1) = 0.45$.

References

1. Sharma G., Trussell H.J., *Digital color imaging*, IEEE Transactions on Image Processing, IP7(6), 1997, pp. 901-932.
2. Cheng H.D., Jiang X.H., Sun Y., Wang J., *Color image segmentation: advances and prospects*, Pattern Recognition, PR-34(6), 2001, pp. 2259-2281.
3. Bloch I., Maitre H., *Fuzzy mathematical morphologies : a comparative study*, Pattern Recognition, PR-9(28), 1995, pp. 1341-1387.
4. Botte-Lecocq C., Zhang R.D., Postaire J.G., *Cluster analysis by binary morphology*, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-15(2), 1993, pp. 170-180.
5. Devijver P.A., Kittler J., *Pattern recognition : a statistical approach*, Prentice-Hall - Englewood Cliffs, New Jersey, 1982.
6. Gillet A., Botte-Lecocq C., Macaire L., Postaire J.G., *Application of fuzzy mathematical morphology for unsupervised color pixels classification*, Data Analysis, Classification and Related Methods, H.A.L. Kiers and all editor, Springer-Verlag, 2000, pp. 69-75.

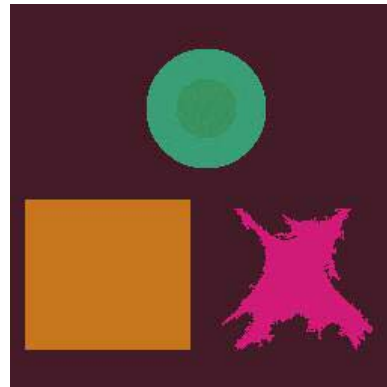


Fig. 21. Image of Fig. 18 segmented with the detected modes of Fig. 20.

7. Haralick R.M., Sternberg S.R., Zhuang X., *Image analysis using mathematical morphology*, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-9(4), 1987, pp. 532-550.
8. Huntsberger T.L., Jacobs C.L., Cannon R.L., *Iterative fuzzy image segmentation*, Pattern Recognition, PR 18(2), 1985, pp. 131-138.
9. Lambert P., Macaire L., *Filtering and segmentation: the specificity of color images*, CGIP'2000 - International Conference on Color in Graphics and Image Processing, Saint-Etienne - France, 2000, pp. 57-71.
10. Lim Y.W., Lee S.U., *On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques*, Pattern Recognition, PR(23)9, 1990, pp. 935-952.
11. Ohta Y.I., Kanade T., Sakai T., *Color information for region segmentation*, Computer Graphics and Image Processing, (13), 1980, pp. 222-241.
12. Park S.H., Yun I.D., Lee S.U., *Color image segmentation based on 3-D clustering : morphological approach*, Pattern Recognition, PR-31(8), 1998, pp. 1061-1076.
13. Pham T.D., Yan H., *Color image segmentation using fuzzy integral and mountain clustering*, Fuzzy Sets and Systems, FSS(107), 1999, pp. 121-130.
14. Postaire J.G., Vasseur C., *A convexity testing method for cluster analysis*, IEEE Transactions on Systems, Man and Cybernetics, SMC(10), 1980, pp. 145-149.
15. Schettini R., *Segmentation algorithm for color images*, Pattern Recognition Letters, PRL (14), 1993, pp. 499-506.
16. Serra J., *Image analysis and mathematical morphology*, Academic Press, 1988.
17. Shaffarenko L., Petrou M., Kittler J., *Histogram-based segmentation in a perceptually uniform color space*, IEEE Transactions on Image Processing, IP-7(9), 1998, pp. 1354-1358.
18. Tominaga S., *Color classification of natural color images*, Color Research and Application, CRA(17)4, 1992, pp. 230-239.
19. Turpin-Dhilly S., Botte-Lecocq C., *Application of fuzzy mathematical morphology for pattern classification*, Advances in Data Science and Classification, IFCS'98, 1998, pp. 125-130.

Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis

N. Vandenbroucke, L. Macaire et J. G. Postaire

Computer Vision and Image Understanding
volume 90, numéro 2
pages 190-216 - 2003



Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®

Computer Vision
and Image
Understanding

Computer Vision and Image Understanding 90 (2003) 190–216

www.elsevier.com/locate/cviu

Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis

Nicolas Vandenbroucke,^a Ludovic Macaire,^{b,*}
and Jack-Gérard Postaire^b

^a École d'Ingénieurs du Pas-de-Calais, Campus de la Malassise, BP39, 62967 Longuenesse, France

^b Laboratoire d'Automatique I³D, CNRS FRE 2497, Bâtiment P2, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq, France

Received 13 November 2001; accepted 6 February 2003

Abstract

In this paper, we propose an original approach in order to improve the results of color image segmentation by pixel classification. We define a new kind of color space by selecting a set of color components which can belong to any of the different classical color spaces. Such spaces, which have neither psycho-visual nor physical color significance, are named *hybrid color spaces*. We propose to classify pixels represented in the hybrid color space which is specifically designed to yield the best discrimination between the pixel classes. This space, which is called the *adapted hybrid color space*, is built by means of a sequential supervised feature selection scheme. This procedure determines the adapted hybrid color space associated with a given family of images. Its dimension is not always equal to three, as for classical color spaces. The effectiveness of our color segmentation method is assessed in the framework of soccer image analysis. The team of each player is identified by the colors of its soccer suit. The aim of the segmentation procedure is to extract meaningful regions representing the players and to recognize their teams.

© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Color images; Color spaces; Segmentation; Soccer

* Corresponding author.

E-mail addresses: nv@i3d.univ-lille1.fr (N. Vandenbroucke), Ludovic.Macaire@univ-lille1.fr (L. Macaire), Jack-Gerard.Postaire@univ-lille1.fr (J.-G. Postaire).

192 N. Vandenbroucke et al. / Computer Vision and Image Understanding 90 (2003) 190–216

1.1.2. Analysis in the color space

The methods based on the analysis of the pixel properties in the color space take advantage of the characterization of each pixel by its three trichromatic components, namely the red (R), the green (G) and the blue (B). In this approach, the pixels are represented by points in the three-dimensional (R, G, B) color space. It is assumed that homogeneous regions in the image plane give rise to clusters in the color space, each cluster corresponding to a class of pixels which share similar color properties. These clusters are generally identified by means of (a) an analysis of the color histogram or (b) a cluster analysis procedure and are mapped back to the original image plane to produce the segmentation.

(a) *Histogram analysis.* Pioneer works are based on marginal analysis of the distribution of the pixel colors in the three-dimensional color space, considering one-dimensional color component histograms. Ohta partitions the image into regions by a recursive analysis of these histograms [21]. Lim and Lee apply a “scale space filter” process to each of the one-dimensional color component histograms to determine modal intervals [14]. Then they partition the three-dimensional color space into hexahedra which are defined by the end points of the modal intervals. The non-empty hexahedra correspond to pixel classes. However, as pointed out by Park, well separated clusters in the three-dimensional color space may correspond to only one single mode in a one-dimensional color component histogram. To overcome this problem, Park analyzes the three-dimensional color histogram and search for the modes by thresholding the difference between two Gaussian smoothed three-dimensional histograms [23]. The so-extracted bins constitute initial seeds of the modes, which are constructed using adaptive morphological transformations. Shafarenko et al. [31] apply a watershed algorithm to a two-dimensional smoothed chromatic histogram to perform class construction. These histogram-based segmentation methods, which do not need a priori information on the image, are multi-thresholding techniques which present severe limitations due to their sensitivity to parameters adjustment.

(b) *Cluster analysis.* Considering the above-mentioned limitations, many authors propose to apply clustering techniques like the C-means and Isodata algorithms which are among the most widely used methods for cluster analysis [9,35]. However, the results achieved by these methods depend on the choice of the initial positions of the cluster centers. Uchiyama proposes to construct the pixel classes by means of an iterative process which is based on a competitive learning scheme [39]. This process, which does not require an initialization step, converges approximately toward the optimal solution. The original pixel clustering algorithm proposed by Scheunders is an hybrid approach combining a genetic algorithm with the C-means algorithm [29]. It seems to be superior to the C-means algorithm because it converges to a near-by global optimum in the least mean square sense, rather than to a local one.

1.2. Segmentation and color spaces

The analysis of the pixel color distribution in a color space is not restricted to the (R, G, B) space. Indeed, there exists a large number of color spaces which can be used

to represent the color of the pixels [32]. They are supported by specific physical, physiological and psycho-visual properties [28] and can be classified into a few non-exhaustive categories according to their definitions [25]. We propose to group the most classical color spaces into four main families, namely:

- The primary spaces which are based on the trichromatic theory, assuming that it is possible to match any color by mixing appropriate amounts of three primary colors.
- The luminance–chrominance spaces where one component represents the luminance and the two others the chrominance.
- The perceptual spaces which try to quantify the subjective human color perception using the intensity, the hue and the saturation.
- The independent axis spaces resulting from different statistical methods which provide as less correlated components as possible.

These four families are represented by the gray rectangles of Fig. 1. The dotted rectangles within these gray rectangles correspond to sub-families of color spaces [43].

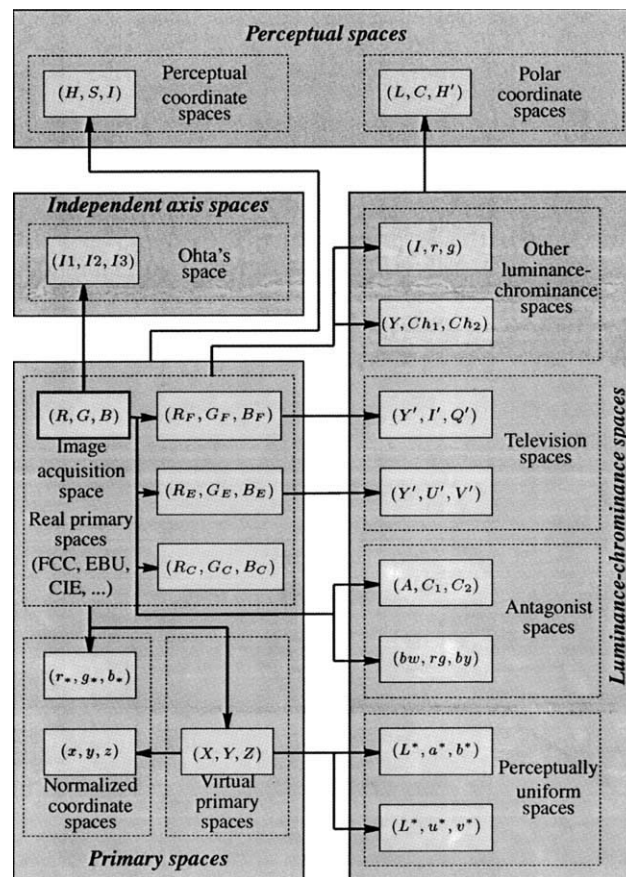


Fig. 1. The four main color space families.

In general, color images are acquired through the R, G, B components which define the (R, G, B) acquisition space. All the other color spaces are defined by linear or nonlinear transformations of this (R, G, B) space. Fig. 1 indicates how each color space is obtained by following the arrows starting from the (R, G, B) acquisition space.

The performance of an image segmentation procedure is known to depend on the choice of the color space. Many authors have tried to determine the color spaces which are the more appropriate for their specific color image segmentation problems.

Brunner discusses the properties of six color spaces to determine which are the best for detecting specific surface defects on Douglas-fir veneer. He shows that these defects are well detected when the clustering analysis is performed in the (R, G, B) space [2]. Littman applies two classification methods using different color spaces [16]. By means of a visual appreciation of the results, he also concludes that the (R, G, B) space is the best among all the considered color spaces. Ohta employs the Ohlander's method [20] in order to determine a new set of independent color components which result from a Karhunen–Loève transformation [21]. From eight kinds of color images, he determines a new color space $(I1, I2, I3)$ which is compared with different well known color spaces. He demonstrates that this space is the most adapted to natural scenes image segmentation. Lim and Lee [14] propose to segment images by means of the fuzzy C-means method. They visually compare the results with images whose color is represented with different color spaces and conclude that the (R, G, B) and $(I1, I2, I3)$ spaces provide the best results. Lee compares the correlation between the components of different color spaces [13]. According to his experiments, the $(I1, I2, I3)$ space provides the best segmentation results while poor results are obtained with the (R, G, B) space. Takahashi visually compares color images which are segmented by the Isodata algorithm with eight different color spaces and concludes that the perceptually uniform spaces (L^*, a^*, b^*) and (L^*, u^*, v^*) are the most adapted [35]. Park, who proposes a color image segmentation algorithm based on a morphological analysis of the color histogram, shows that segmentation results in five different color spaces are almost identical [23]. Liu applies his multiresolution segmentation method using six different color spaces [17]. He does not find any color space which provides satisfying results for the segmentation of all kinds of images.

This short and non-exhaustive bibliography shows that different authors provide contradictory conclusions about the pertinence of the available color spaces in the context of image segmentation.

Instead of searching the best classical color space for segmentation, we propose an original approach in order to improve the results of image segmentation by pixel classification. We define a new kind of color space by selecting a set of color components which can belong to any of the different classical color spaces listed in Fig. 1 [45]. Such spaces, which have neither psycho-visual nor physical color significance, are named *hybrid color spaces* in this paper.

We propose to classify pixels represented in the hybrid color space which is specifically designed to yield the best discrimination between the pixel classes [41]. This space, which is called the *adapted hybrid color space*, is built by means of a sequential

supervised feature selection scheme [5]. This procedure, presented in the second section of this paper, determines the adapted hybrid color space associated with a given family of images. Its dimension is not always equal to three, as for classical color spaces.

The performance of our approach is evaluated in the third section. The effectiveness of our color segmentation method is assessed in the framework of color soccer image analysis. The team of each player is characterized by the soccer suit colors which are not always homogeneous. The aim of the segmentation procedure is to extract meaningful regions representing the players and to identify their teams.

This application is well-suited to illustrate the efficiency of the proposed segmentation scheme, because extensive experiments can be carried out with a high number of soccer games, the color suits being specific to each soccer game [40–45]. One of the main problems is the determination of the adapted hybrid color space which yields the lowest segmentation error rate, since the sizes of the players are relatively small in the images. Furthermore, when two players get so close from each other that there is no space let between them, the teams of the touching players are difficult to identify in the images. When the segmentation scheme is the low level image processing stage used by a player tracking system, the quality of the segmentation is crucial.

2. Adapted hybrid color space construction

In this section, we present the construction of the hybrid color space which is adapted to the discrimination of pixel classes. For a given family of images or for a sequence of images, the adapted hybrid color space is determined in a supervised context. In the first part of this section, we consider the learning samples which contain representative pixels of the images. The components of an hybrid color space come from different color spaces and may be more or less correlated. As correlated color components are not a priori efficient for discriminating classes, a correlation measure is introduced in the second part of this section. With this measure, the adapted hybrid color space construction scheme takes into account only candidate hybrid color spaces constituted of color components which are as less correlated as possible. In the third part of this section, we present the sequential color components selection procedure which determines the adapted hybrid color space. It is assumed that the more the clusters corresponding to the pixel classes are well separated and compact in a candidate hybrid color space, the higher its discriminating power is. A suitable criterion is used to measure the discriminating power of the candidate hybrid color spaces. Finally, the determination of the dimension of the adapted hybrid color space, which is not always equal to three, is detailed in the last part of this section.

2.1. Learning sample

As stated before, our approach to color image segmentation by pixel classification is to look for the best hybrid color space for discriminating the N_{class} pixel classes constituting the image.

In order to determine the adapted hybrid color space we require, for each pixel class C_j , $j = 1, \dots, N_{\text{class}}$, a set of N_ω representative pixels $\omega_{i,j}$, $i = 1, \dots, N_\omega$. We will show, in the third section of this paper, how these representative pixels can be extracted from the images by means of an interactive procedure which is designed in the context of the segmentation of sequences of soccer images.

In a d -dimensional color space, we define, for each representative pixel $\omega_{i,j}$ of class C_j , a d -dimensional observation $X_{i,j} = [x_{i,j}^1, \dots, x_{i,j}^k, \dots, x_{i,j}^d]^T$, where $x_{i,j}^k$ is the level of the k^{th} color component. It is convenient to express this set of observations in the form of a matrix $\mathbf{X}$ defined as

$$\mathbf{X} = \begin{bmatrix} x_{1,1}^1 & \cdots & x_{N_\omega,1}^1 & \cdots & x_{i,j}^1 & \cdots & x_{1,N_{\text{class}}}^1 & \cdots & x_{N_\omega,N_{\text{class}}}^1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{1,1}^k & \cdots & x_{N_\omega,1}^k & \cdots & x_{i,j}^k & \cdots & x_{1,N_{\text{class}}}^k & \cdots & x_{N_\omega,N_{\text{class}}}^k \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{1,1}^d & \cdots & x_{N_\omega,1}^d & \cdots & x_{i,j}^d & \cdots & x_{1,N_{\text{class}}}^d & \cdots & x_{N_\omega,N_{\text{class}}}^d \end{bmatrix}. \quad (1)$$

The rows of the matrix $\mathbf{X}$ correspond to the $(N_\omega \times N_{\text{class}})$ available representative pixels while the columns correspond to the d color components of each representative pixel in the color space.

We can also write as

$$\mathbf{X} = \begin{bmatrix} X^1 \\ \vdots \\ X^k \\ \vdots \\ X^d \end{bmatrix}, \quad (2)$$

where $X^k = [x_{1,1}^k, \dots, x_{N_\omega,1}^k, x_{1,2}^k, \dots, x_{i,j}^k, \dots, x_{1,N_{\text{class}}}^k, \dots, x_{N_\omega,N_{\text{class}}}^k]$ contains the levels of the k^{th} color component of the $(N_\omega \times N_{\text{class}})$ representative pixels, or:

$$\mathbf{X} = [X_{1,1} \quad \cdots \quad X_{N_\omega,1} \quad X_{1,2} \quad \cdots \quad X_{i,j} \quad \cdots \quad X_{1,N_{\text{class}}} \quad \cdots \quad X_{N_\omega,N_{\text{class}}}], \quad (3)$$

where $X_{i,j}$ represents the d -dimensional observation associated with each representative pixel $\omega_{i,j}$.

2.2. Candidate hybrid color space

The construction of the adapted hybrid color space is based on a step by step color components selection. At each step d , this sequential procedure selects a new color component which is associated with the $(d - 1)$ previously selected color components, in order to constitute the *most discriminating d -dimensional hybrid color space*, denoted $\mathbf{S}^d$. The procedure stops at step $d = D$ when the stopping criterion described in the fourth part of this section is verified. The most discriminating D -dimensional hybrid color space is considered as the adapted hybrid color space which will be used to classify the pixels.

Let N_{comp} be the number of available color components, i.e., the number of different color components defining all the classical color spaces described in Fig. 1.

At each step d ($d \geq 2$) of the sequential selection procedure, we associate each of the $(N_{\text{comp}} - d + 1)$ remaining color components with the $(d - 1)$ previously selected color components in order to constitute $(N_{\text{comp}} - d + 1)$ d -dimensional hybrid color spaces denoted HS_a^d , with $a = 1, \dots, (N_{\text{comp}} - d + 1)$.

Among these components, some are more or less correlated to the $(d - 1)$ previously selected color components. When constructing a color space for pixel classification, Ohta and Tominaga have shown that looking for un-correlated components by means of the Karhunen–Loève transformation yields satisfying results [21,36]. Furthermore, Kouassi considers that color components have to be decorrelated so that each color component can be analyzed separately [11]. As we look for the most discriminating hybrid color space among a number of candidates, we first withdraw those with correlated components.

For this purpose, at each step d of the sequential selection procedure we measure the correlation between the components of each of the $(N_{\text{comp}} - d + 1)$ d -dimensional hybrid color spaces HS_a^d .

The d -dimensional hybrid color space HS_a^d contains the $(d - 1)$ color components which constitute the most discriminating $(d - 1)$ -dimensional hybrid color space S^{d-1} . A vector X^k , $k = 1, \dots, d - 1$, defined in Eq. (2), is associated with each of these components. The d^{th} color component of HS_a^d is one of the remaining $(N_{\text{comp}} - d + 1)$ color components. We denote X^d the vector which contains the values of this d^{th} color component for the $(N_{\omega} \times N_{\text{class}})$ representative pixels.

For measuring the correlation between the color components of HS_a^d , we consider the $(d - 1)$ couples of vectors constituted of X^d and each of the vectors X^k , $k = 1, \dots, d - 1$.

Let $\text{cor}(X^k, X^d)$ denote the correlation between a k^{th} color component, $k = 1, \dots, d - 1$, and the d^{th} color component, which is defined as

$$\text{cor}(X^k, X^d) = \frac{\text{cov}(X^k, X^d)}{\sigma^k \times \sigma^d}, \quad (4)$$

where the covariance $\text{cov}(X^k, X^d)$ between a k^{th} color component and the d^{th} color component is given by

$$\text{cov}(X^k, X^d) = \frac{1}{N_{\text{class}} \times N_{\omega}} \times \sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} (x_{i,j}^k - m^k) \times (x_{i,j}^d - m^d), \quad (5)$$

where m^k , m^d , σ^k , and σ^d are the means and the standard deviations of the k^{th} and the d^{th} color components, respectively. For the k^{th} color component, they are defined as

$$m^k = \frac{1}{N_{\text{class}} \times N_{\omega}} \times \sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} x_{i,j}^k \quad (6)$$

and

$$\sigma^k = \frac{1}{\sqrt{N_{\text{class}} \times N_{\omega}}} \times \sqrt{\sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} (x_{i,j}^k - m^k)^2}. \quad (7)$$

The correlation values range from 0 to 1. The closer is the correlation to 1, the more correlated are the two color components.

The correlation measure $J_{\text{cor}}(HS_a^d)$ between the color components of the d -dimensional hybrid color space HS_a^d is evaluated as the maximum of the correlation measures between the d^{th} color component and all the other $(d - 1)$ color components

$$J_{\text{cor}}(HS_a^d) = \max_{k=1}^{d-1} (\text{cor}(X^d, X^k)). \quad (8)$$

We consider that only the d -dimensional hybrid color spaces with a correlation measure lower than a given threshold can be taken into account as interesting candidates for the most discriminating hybrid color space construction procedure. By means of this thresholding procedure, which is achieved at each step d of the color component selection, we consider only a limited number N_{cand}^d of interesting d -dimensional candidate hybrid color spaces S_l^d , with $l = 1, \dots, N_{\text{cand}}^d$, among all the $(N_{\text{comp}} - d + 1)$ available hybrid color spaces HS_a^d .

2.3. Adapted hybrid color space construction

As it would be unrealistic to evaluate the discriminating power of all the possible combinations of color components, we use a sub-optimal sequential forward selection algorithm for finding the adapted hybrid color space [6]. More sophisticated strategies exist, such as the sequential adaptive floating search method which selects features by means of successive forward and backward search steps [33]. A comparison of the performances of different feature selection algorithms has shown that the results obtained with different sequential search methods are generally similar for the first five selected features [24]. Hence, for computational efficiency, we have used the simplest sequential forward selection process.

At each step d of the construction procedure, we consider the set of N_{cand}^d d -dimensional candidate hybrid color spaces S_l^d , $l = 1, \dots, N_{\text{cand}}^d$. We evaluate their discriminating powers with a suitable criterion and select the best one, which is the most discriminating d -dimensional hybrid color space S^d . This criterion, denoted $J_{\text{dis}}(S_l^d)$, measures the scatter of the observations associated with the representative pixels in each of the d -dimensional candidate hybrid color spaces. To be more specific, the problem is to find an hybrid color space in which the colors of the pixels which belong to each region of the image are similar, while the colors of pixels belonging to distinct regions are significantly different.

At the beginning of the construction procedure ($d = 1$), we consider the one-dimensional candidate spaces defined by each of the N_{comp} available color components. The selected one-dimensional space S^1 is that which maximizes the discriminating power $J_{\text{dis}}(S_l^1)$. At the second step of the procedure ($d = 2$), two-dimensional hybrid color spaces are constituted by associating the color component of S^1 with each of $(N_{\text{comp}} - 1)$ remaining color components. Among the N_{cand}^2 two-dimensional hybrid color spaces S_l^2 , $l = 1, \dots, N_{\text{cand}}^2$, that satisfy the criterion based on the correlation measure, we select the one that maximizes $J_{\text{dis}}(S_l^2)$ as the best in terms of discriminating power between the pixel classes. This selected space is the most discriminating

two-dimensional hybrid color space $\mathbf{S}^2$. The procedure is iterated until a stopping criterion, denoted J_{stop}^d and detailed in the next section, is verified. The number $d = D$ of the last step indicates the dimension D of the adapted hybrid color space $\mathbf{S}^D$. Fig. 2 illustrates the procedure which constructs the adapted hybrid color space.

The discriminating power of a candidate hybrid color space is evaluated by means of a measure of class separability and compactness. We use the within-cluster scatter matrix $\mathbf{W}(S_l^d)$ to evaluate the compactness of each class C_j in each d -dimensional candidate hybrid color space S_l^d , $l = 1, \dots, N_{\text{cand}}^d$

$$\mathbf{W}(S_l^d) = \frac{1}{N_{\omega} \times N_{\text{class}}} \times \sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} (X_{i,j} - M_j)(X_{i,j} - M_j)^T, \quad (9)$$

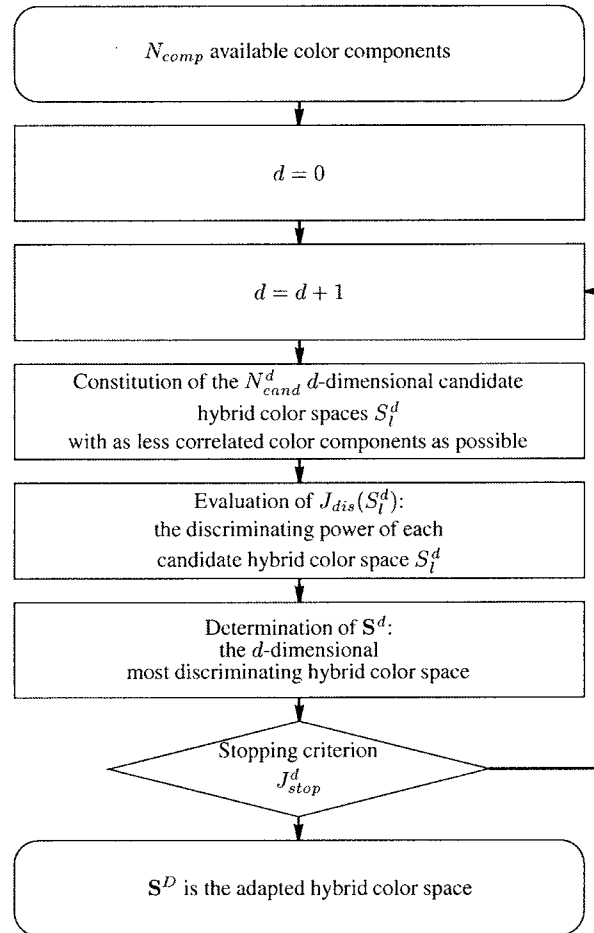


Fig. 2. Sequential procedure of the adapted hybrid color space construction.

200 *N. Vandenbroucke et al. / Computer Vision and Image Understanding 90 (2003) 190–216*

where $X_{i,j}$ is defined in Eq. (3), and $M_j = [m_j^1, \dots, m_j^k, \dots, m_j^d]^T$ is the d -dimensional mean vector of the N_o observations $X_{i,j}$ used to define the class C_j , with

$$m_j^k = \frac{1}{N_o} \sum_{i=1}^{N_o} x_{i,j}^k. \quad (10)$$

To measure the class separability, we use the between-cluster scatter matrix $\mathbf{B}(S_l^d)$

$$\mathbf{B}(S_l^d) = \frac{1}{N_{\text{class}}} \times \sum_{j=1}^{N_{\text{class}}} (M_j - M)(M_j - M)^T, \quad (11)$$

where $M = [m^1, \dots, m^k, \dots, m^d]^T$ is the d -dimensional mean vector of all the available observations, with m^k defined in Eq. (6).

For each candidate hybrid color space S_l^d , the discriminating power, denoted $J_{\text{dis}}(S_l^d)$, is measured by the well-known trace criterion [8]. In these conditions, $J_{\text{dis}}(S_l^d)$ can be written as

$$J_{\text{dis}}(S_l^d) = \text{trace} \left((\mathbf{W}(S_l^d) + \mathbf{B}(S_l^d))^{-1} \mathbf{W}(S_l^d) \right) = \text{trace} \left((\mathbf{T}(S_l^d))^{-1} \mathbf{W}(S_l^d) \right), \quad (12)$$

where $\mathbf{T}(S_l^d) = \mathbf{W}(S_l^d) + \mathbf{B}(S_l^d)$ is the total scatter matrix.

At each step d of the construction procedure, the most discriminating d -dimensional hybrid color space $\mathbf{S}^d$ is that which maximizes the discriminating criterion

$$\mathbf{S}^d = \left\{ S_{l_0}^d \mid J_{\text{dis}}(S_{l_0}^d) = \max_{l=1}^{N_{\text{cand}}^d} (J_{\text{dis}}(S_l^d)) \right\}, \quad (13)$$

where l_0 indicates the number of the selected candidate hybrid color space.

2.4. Stopping criterion

The criterion given in Eq. (12) is used to compare the discriminating powers of different sets of color components. However, it does not indicate if the d -dimensional hybrid color space $\mathbf{S}^d$ which is determined at step d is more discriminating than $\mathbf{S}^{d-1}$, i.e., the $(d-1)$ -dimensional hybrid color space selected at step $(d-1)$. We propose to use an information criterion in order to stop the selection procedure when the d^{th} color component of $\mathbf{S}^d$ which is selected at step d does not significantly improve the discrimination.

This criterion measures the scatter of the observations in the one-dimensional space constituted by the d^{th} component of $\mathbf{S}^d$, given the scatter of the observations in the most discriminating $(d-1)$ -dimensional space $\mathbf{S}^{d-1}$ [26]. Rao defines the conditional total and within-cluster scatter matrices with respect to the d^{th} component of $\mathbf{S}^d$, given the $(d-1)$ components of $\mathbf{S}^{d-1}$. For this purpose, he divides the d components of $\mathbf{S}^d$ into two groups, namely the $(d-1)$ components of $\mathbf{S}^{d-1}$ and the d^{th} component associated with $\mathbf{S}^d$. The within-cluster scatter $\mathbf{W}(\mathbf{S}^d)$ is a $(d \times d)$ -dimensional matrix which can be written as

$$\mathbf{W}(\mathbf{S}^d) = \frac{1}{N_{\text{class}} \times N_{\omega}} \times \begin{bmatrix} \sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} (x_{i,j}^1 - m_j^1)^2 & \cdots & \sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} (x_{i,j}^1 - m_j^1) \times (x_{i,j}^d - m_j^d) \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} (x_{i,j}^d - m_j^d) \times (x_{i,j}^1 - m_j^1) & \cdots & \sum_{j=1}^{N_{\text{class}}} \sum_{i=1}^{N_{\omega}} (x_{i,j}^d - m_j^d)^2 \end{bmatrix}. \quad (14)$$

The elements of the last row and the last column of this matrix are associated with the d^{th} selected color component. So, the matrix $\mathbf{W}(\mathbf{S}^d)$ can be written as

$$\mathbf{W}(\mathbf{S}^d) = \begin{bmatrix} \mathbf{W}(\mathbf{S}^{d-1}) & \mathbf{W}_{d-1,1}(\mathbf{S}^d) \\ \mathbf{W}_{1,d-1}(\mathbf{S}^d) & \mathbf{W}_{1,1}(\mathbf{S}^d) \end{bmatrix}, \quad (15)$$

where $\mathbf{W}(\mathbf{S}^{d-1})$ is the $(d-1) \times (d-1)$ -dimensional within-cluster scatter matrix of the observations in the space $\mathbf{S}^{d-1}$. The two $(d-1)$ -dimensional vectors $\mathbf{W}_{1,d-1}(\mathbf{S}^d)$ and $\mathbf{W}_{d-1,1}(\mathbf{S}^d)$ and the scalar $\mathbf{W}_{1,1}(\mathbf{S}^d)$ provide a measure of the contribution of the d^{th} color component in terms of discriminating power. A similar decomposition can be applied to the matrix $\mathbf{T}(\mathbf{S}^d)$.

By means of this decomposition, the conditional within-cluster scatter matrix, denoted $\varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1})$, is defined as

$$\varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1}) = \mathbf{W}_{1,1}(\mathbf{S}^d) - \mathbf{W}_{1,d-1}(\mathbf{S}^d)(\mathbf{W}(\mathbf{S}^{d-1}))^{-1}\mathbf{W}_{d-1,1}(\mathbf{S}^d). \quad (16)$$

Similarly, the conditional total scatter matrix, denoted $\varsigma_T(\mathbf{S}^d/\mathbf{S}^{d-1})$, is defined as

$$\varsigma_T(\mathbf{S}^d/\mathbf{S}^{d-1}) = \mathbf{T}_{1,1}(\mathbf{S}^d) - \mathbf{T}_{1,d-1}(\mathbf{S}^d)(\mathbf{T}(\mathbf{S}^{d-1}))^{-1}\mathbf{T}_{d-1,1}(\mathbf{S}^d). \quad (17)$$

As the difference between the dimensions of $\mathbf{S}^d$ and $\mathbf{S}^{d-1}$ is equal to one, the conditional within-cluster and total scatter matrices are reduced to scalars called, hereafter, conditional scatter measures. The stopping criterion, which has been proposed by Romeder [27], is evaluated with the relative difference of the conditional scatter measures $\varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1})$ and $\varsigma_T(\mathbf{S}^d/\mathbf{S}^{d-1})$

$$J_{\text{stop}}^d = \frac{\varsigma_T(\mathbf{S}^d/\mathbf{S}^{d-1}) - \varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1})}{\varsigma_W(\mathbf{S}^d/\mathbf{S}^{d-1})}. \quad (18)$$

With this criterion, it is possible to quantify the contribution of the selected hybrid color space $\mathbf{S}^d$ with respect to the preceding space $\mathbf{S}^{d-1}$. As long as the criterion J_{stop}^d increases, the most discriminating hybrid color spaces should improve the classification results. The evolution of J_{stop}^d is analyzed in order to detect any decreasing phase which indicates that the new most discriminating hybrid color space $\mathbf{S}^d$ should yield lower performance in terms of discrimination between the pixel classes than the preceding one $\mathbf{S}^{d-1}$. Hence, the procedure stops when the value of this criterion decreases. Let D denote the number of steps of the procedure which precedes the first decreasing phase of the stopping criterion. This strategy yields the dimension D of the adapted hybrid color space $\mathbf{S}^D$.

202 N. Vandenbroucke et al. / *Computer Vision and Image Understanding* 90 (2003) 190–216

In order to evaluate the performance of our approach, we propose to extract the players from color soccer images. The aim of the procedure is to identify their teams by means of the analysis of the colors of their suits, which may be heterogeneous. This application has been chosen to show how psycho-visually heterogeneous regions in the (R, G, B) color space can be represented as sets of pixels sharing similar properties in the adapted hybrid color space.

3. Application to soccer color image analysis

Many authors have tried to extract soccer players from video sequences for tracking purpose. They apply either the multi-thresholding technique which is based on the analysis of the histograms of the three trichromatic components [10,30,19], either the bloc matching method [7] or the difference between the current image and a reference image which only contains the ground [1,34]. Although good results have been reported with these segmentation techniques, they tend to fail in cases of occlusion and touching players.

We apply our proposed classification scheme in the adapted hybrid color space to extract the soccer players and identify their teams in sequences of color images acquired by a fixed camera. This approach is designed to be integrated in a tracking player system.

Let us consider four images of such a sequence (see Fig. 3). The aim of the analysis of these images is to extract the players and the referee. In this context, there are three pixel classes of interest, i.e., the pixels representing the players of the two opposing teams and the referee, which are identified by the colors of their suits.

First, the ground is withdrawn by means of the Ohlander's multi-thresholding scheme [20] in order to extract the player pixels, i.e., the pixels which represent the players (see Fig. 4). This procedure provides pertinent thresholds by means of the analysis of the one-dimensional histograms of the trichromatic components.

The soccer player identification scheme is divided into two steps which are described in the two first parts of this section. First, the adapted hybrid color space is built off-line using a supervised learning procedure. For this purpose, a learning sample is interactively selected from images of the game. Then, an on-line player pixel classification scheme is used to perform a real-time segmentation of the successive

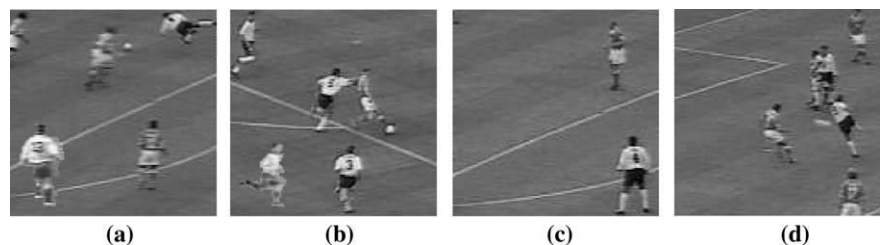


Fig. 3. Soccer images (125×125 pixels).

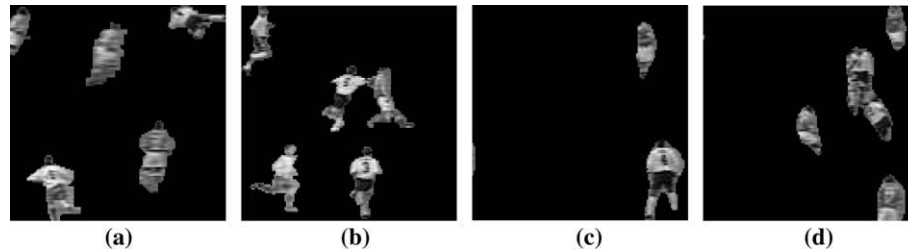


Fig. 4. Player pixels extracted from the images of Fig. 3.

images whose colors are represented in the so-built adapted hybrid color space. For this purpose, each player pixel is assigned to the class with the nearest center to its color in the adapted hybrid color space. Experimental results presented in the last part of this section demonstrate the effectiveness of our approach.

3.1. Supervised learning scheme

In order to determine the adapted hybrid color space for player pixel classification, it is necessary to construct a learning sample representative of the player pixel classes. Using the notations defined in Section 2.1, the player pixels are supposed to belong to a set of N_{class} classes C_j , $j = 1, \dots, N_{\text{class}}$. The learning sample is constituted of N_{ω} representative player pixels $\omega_{i,j}$, $i = 1, \dots, N_{\omega}$, for each class C_j , $j = 1, \dots, N_{\text{class}}$.

We consider the pre-segmented learning images from which the background has been eliminated as previously explained (see Fig. 4). In each of these learning images, bounding boxes are interactively adjusted on the players, so as to extract rectangular player windows, each one containing the image of one player (see Fig. 5). These player windows are extracted from a series of images where players are in different situations (running, jumping, pushing the ball, dribbling, etc.) and in different positions with respect to the camera (facing the camera, turning their back toward the camera, far or near the camera, etc.).

We select the same number N_W of player windows for each class (see Fig. 6). In order to constitute the set of the N_{ω} representative player pixels for each class C_j , $j = 1, \dots, N_{\text{class}}$, N_{ω}/N_W player pixels are automatically and randomly selected in each of the N_W player windows (see Fig. 7).

In order to analyze the discriminating properties of a d -dimensional candidate hybrid color space, the construction procedure described in the preceding section associates a d -dimensional observation $X_{i,j} = [x_{i,j}^1, \dots, x_{i,j}^k, \dots, x_{i,j}^d]^T$ with each representative player pixel $\omega_{i,j}$.

If we consider the d color components of the representative player pixel $\omega_{i,j}$, the observations corresponding to the same player would be sensitive to non-significant local variations of his suit color. In order to reduce the effects of such variations, each color component is evaluated in a neighborhood centered on the representative player pixel $\omega_{i,j}$. To be more specific, the k^{th} color component of the observation $X_{i,j}$

204 N. Vandenbroucke et al. / Computer Vision and Image Understanding 90 (2003) 190–216

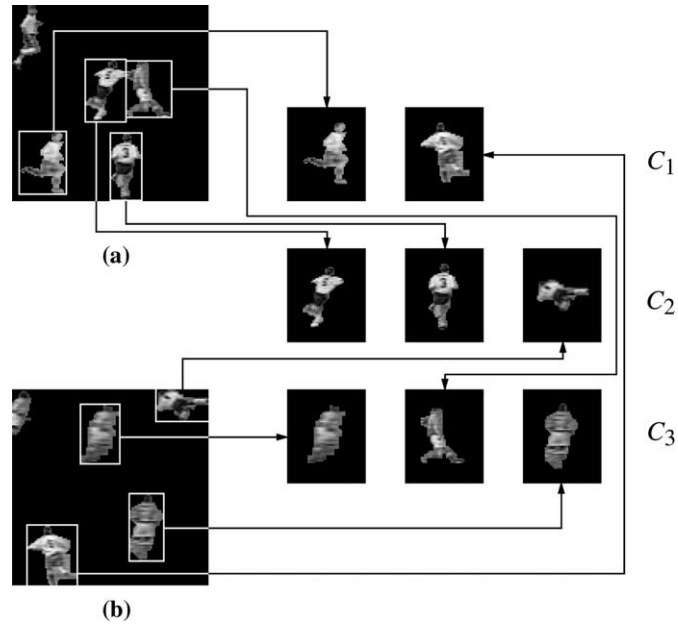


Fig. 5. Player windows interactively extracted from learning images.

is computed as the mean value of this component for all the player pixels falling into a neighborhood centered on the representative player pixel $\omega_{i,j}$.

The construction of the adapted hybrid color space is based on the analysis of this learning set, which is constituted of the $(N_{\omega} \times N_{\text{class}})$ vectors $X_{i,j}$. For this example, the resulting adapted hybrid color space S^D is the plane ($D = 2$) which is defined by the a^* color component of the (L^*, a^*, b^*) space and the C_{UV} chroma component evaluated in the (Y', U', V') space.

Fig. 8 shows the evolution of the stopping criterion of the construction procedure which decreases after the second step of this sequential procedure and therefore, leads to this two-dimensional space.

3.2. Player pixel classification

Once the off-line supervised learning scheme has provided the adapted hybrid color space S^D , the segmentation of a sequence of color images can be achieved in real-time by means of the following classification scheme. First, the Ohlander's multi-thresholding extracts the player pixels of the color images. In order to classify the so-extracted player pixels, we compute a feature vector $X_P = [x_P^1, \dots, x_P^D]^T$ for each player pixel P in the adapted hybrid color space S^D . As during the learning step, the coordinates of X_P are the mean color components of the player pixels falling into the previously defined neighborhood centered on the player pixel P (see Fig. 9). Since only player pixels are taken into account for this mean evaluation, this procedure reduces non-significant local variations of the player suit colors.

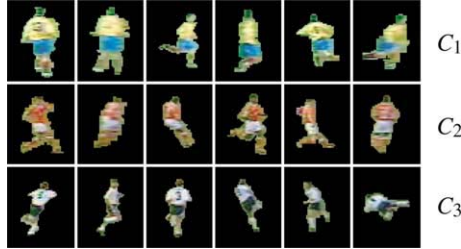


Fig. 6. Player windows.

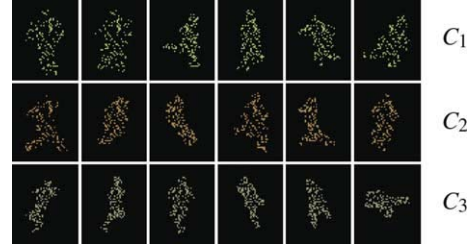


Fig. 7. Representative player pixels.

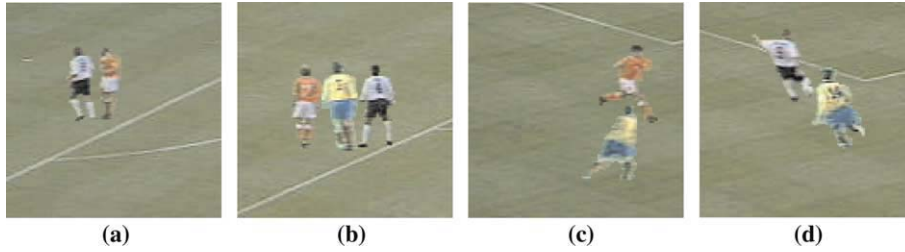


Fig. 10. Color soccer images (125 × 125 pixels).

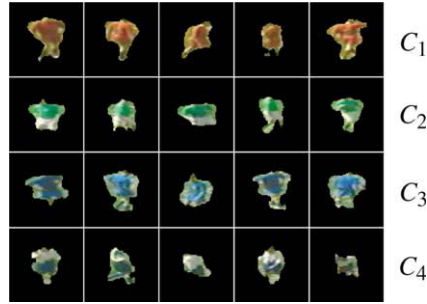


Fig. 16. Player windows.

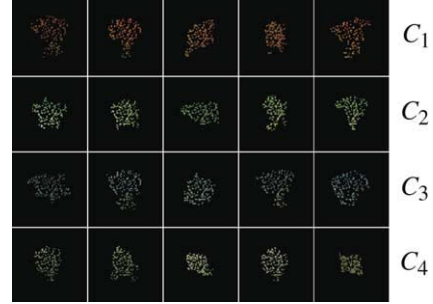


Fig. 17. Representative player pixels.

The simple minimum distance decision rule is used to assign the player pixel P to the class with the nearest center in the adapted hybrid color space. The Euclidean distance D_j^P between X_P and the center of each class C_j , defined as the mean vector $M_j = [m_j^1, \dots, m_j^D]^T$ in the adapted hybrid color space $\mathbf{S}^D$, is evaluated as

$$D_j^P = \|X_P - M_j\| = \sqrt{\sum_{d=1}^D (x_P^d - m_j^d)^2}. \quad (19)$$

206 *N. Vandenbroucke et al. / Computer Vision and Image Understanding 90 (2003) 190–216*

The player pixel is assigned to the class C_l with the smallest distance

$$\{C_l = C_{l_0} \mid D_{l_0}^p = \min_{j=1}^{N_{\text{class}}}(D_j^p)\}. \quad (20)$$

As our purpose is to compare the results achieved with the same classification procedure but in different color spaces, we have not tried to use more sophisticated classification schemes such as those based on neural networks [15] or Markov random fields [18], which could improve the results, but which are beyond the scope of this paper.

3.3. Results

3.3.1. Example 1

In order to compare the segmentation achieved in the adapted hybrid color space with the segmentation using classical color spaces, we propose to segment four color images extracted from a sequence with two opposing players and the referee (see Fig. 10).

The images of Fig. 10 constitute a test sample extracted from the same sequence as the images of Fig. 3 which have been used to build the learning sample. These im-

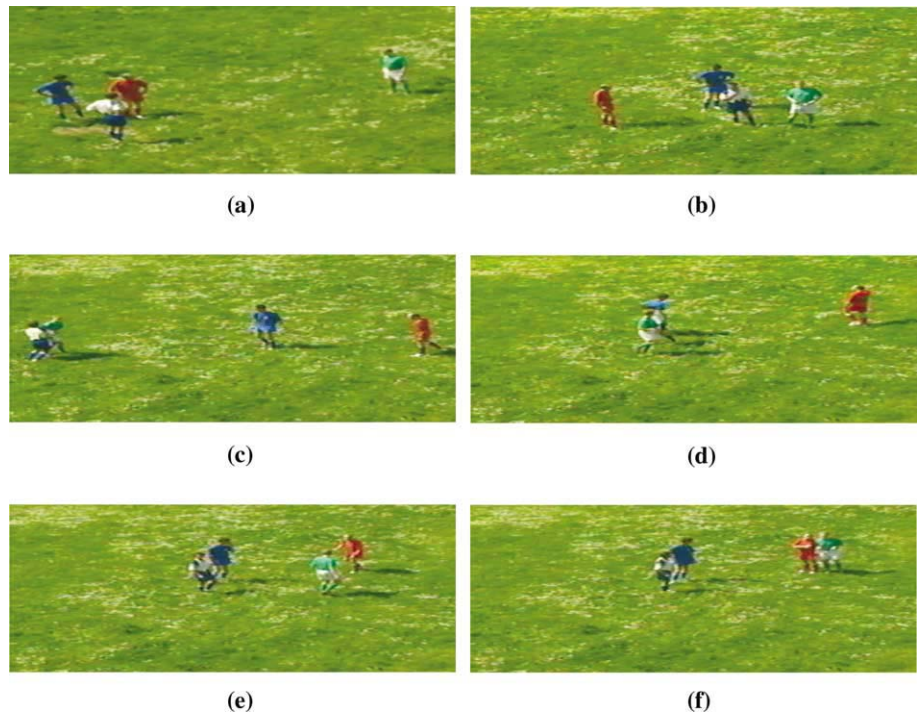


Fig. 14. Color soccer images (400×150 pixels).

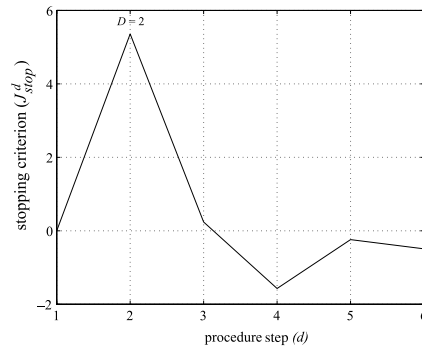


Fig. 8. Stopping criterion evolution.

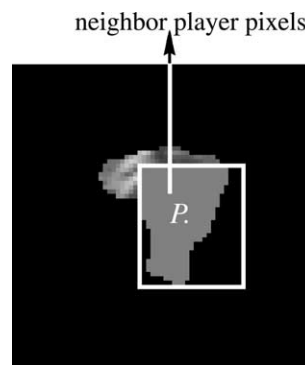


Fig. 9. Player pixel neighborhood.

ages, which are different from those of Fig. 3, are not easy to segment because each of them contains at least two touching players.

The player pixels which are extracted from the images of Fig. 10 by means of the Ohlander's multi-thresholding scheme are displayed in the images of Fig. 11.

The images of Fig. 12 show how the so-extracted player pixels have been classified in the previously determined adapted hybrid color space (a^* , C_{UV}). The player pixels assigned to the same class are labelled with the same color. The connected player pixels with the same label constitute regions which correspond to the different soccer players.

Note that in the images of Figs. 11(a) and (b), the colors of the legs or socks of these players are quite similar to the colors of the suits of opposing players. This explains why the player pixels in the images of Figs. 12(a) and (b), which represent the legs or the socks of players are misclassified.

In order to demonstrate the discriminating properties of the adapted hybrid color space, we compare the segmentation results shown in Fig. 12 with those achieved when using the following 22 classical color spaces described in Fig. 1, namely:

208 N. Vandenbroucke et al. / Computer Vision and Image Understanding 90 (2003) 190–216

- the four primary spaces: (R, G, B) , (r, g, b) , (X, Y, Z) , (x, y, z) ,
- the independent axis space: $(I1, I2, I3)$,
- the eight luminance–chrominance spaces: (L^*, u^*, v^*) , (L^*, a^*, b^*) , (bw, rg, by) , $(A, C1, C2)$, (Y', U', V') , (Y', I', Q') , (Y, Ch_1, Ch_2) , (I, r, g) ,
- the nine perceptual spaces including the eight (L, C, H') polar coordinate spaces deduced from the eight luminance–chrominance spaces and the (H, S, I) perceptual coordinate space.

For this purpose, we propose to measure the classification error rate, denoted ε , by comparison of automatically segmented images with interactively segmented ones (see Fig. 13). Although the interactive assignment of the player pixels to the different classes is rather tedious, there is no other solution to assess the results.

The (r, g, b) , the (H, S, I) and the (Y', C_{UV}, h_{UV}) spaces are the three classical color spaces leading to the lowest error rates. Table 1 contains these classification error rates for all the images of Fig. 11. They are compared to the error rate obtained with the (a^*, C_{UV}) space, which is the adapted hybrid color space. The player pixel classification which is achieved in the adapted hybrid color space provides the lowest error rates for the images of Figs. 11(b)–(d). For the image of Fig. 11(a), the classification error rate in the adapted hybrid color space remains close to the lowest one achieved in the (H, S, I) space. Furthermore, we have added in Table 1 the error rates obtained with the (R, G, B) acquisition space to show that they are significantly higher than those obtained with the adapted hybrid color space.

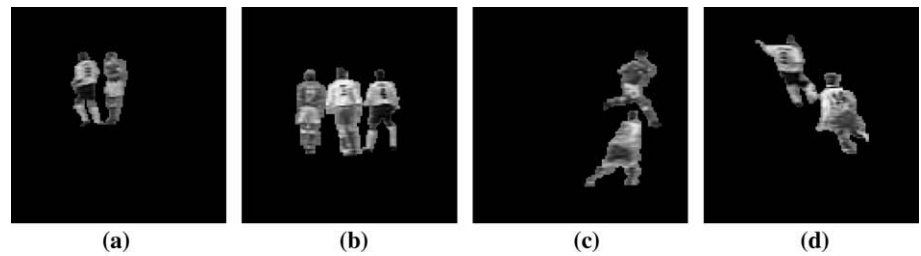


Fig. 11. Player pixels extracted from the images of Fig. 10.

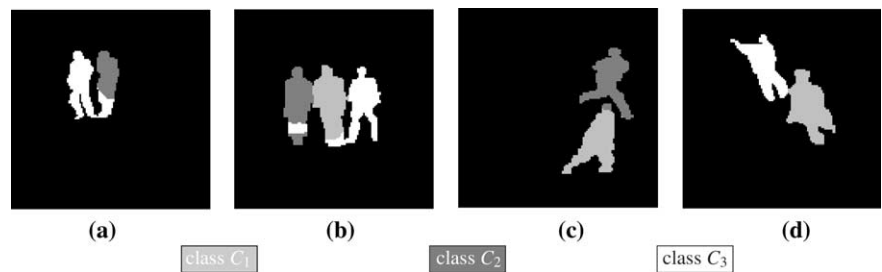


Fig. 12. Player pixels of the images of Fig. 11 classified in the adapted hybrid color space (a^*, C_{UV}) .

As the mean classification error rate associated with the adapted hybrid color space is the lowest one, we conclude that the classification in this two-dimensional hybrid color space provides better results than the classification in the three classical three-dimensional color spaces leading to the lowest error rates.

3.3.2. Example 2

In order to demonstrate the performance of our method in non-trivial situations, we propose to segment six color images extracted from a sequence with two opposing players, one goalkeeper and the referee (see Fig. 14). This second example has been selected to get an insight into the behavior of our segmentation scheme in rather complex situations, such as partially hidden players in the images of Figs. 14(c) and (d) and touching players in the images of Figs. 14(a), (b), (e), and (f).

The images of Fig. 15 show the player pixels which are extracted from the images of Fig. 14 by means of the Ohlander's multi-thresholding scheme.

During the supervised learning scheme, we have selected the set of different player windows presented in Fig. 16. These player windows are selected in images of the same game, but they are different from those of Fig. 14. Fig. 17 contains the set of the selected representative player pixels of the four classes corresponding to the two teams, one of the two goalkeepers and the referee.

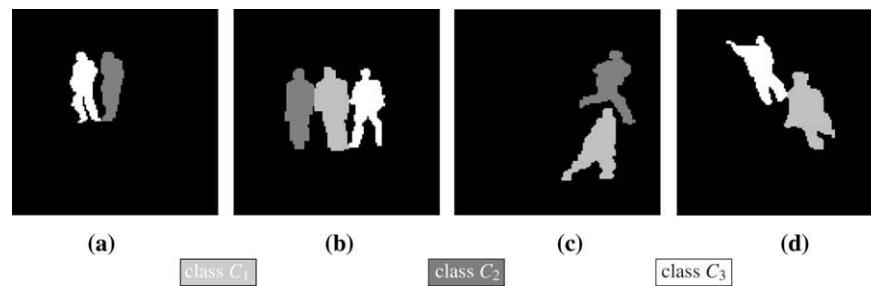


Fig. 13. Player pixels of the images of Fig. 11 interactively classified.

Table 1
Classification error rates of player pixels of Fig. 11

Color space	Classification error rates ε				Mean error rate $\bar{\varepsilon}$
	Image (a)	Image (b)	Image (c)	Image (d)	
(r, g, b)	32.63	10.01	5.19	0.65	10.24
(H, S, I)	8.63	27.77	11.63	1.81	14.39
(Y', C_{UV}, h_{UV})	12.28	17.15	23.49	6.14	15.08
(a^*, C_{UV})	11.5	7.35	1.86	0.29	4.98
(R, G, B)	24.12	29.25	44.57	48.23	36.73

210 N. Vandenbroucke et al. / Computer Vision and Image Understanding 90 (2003) 190–216

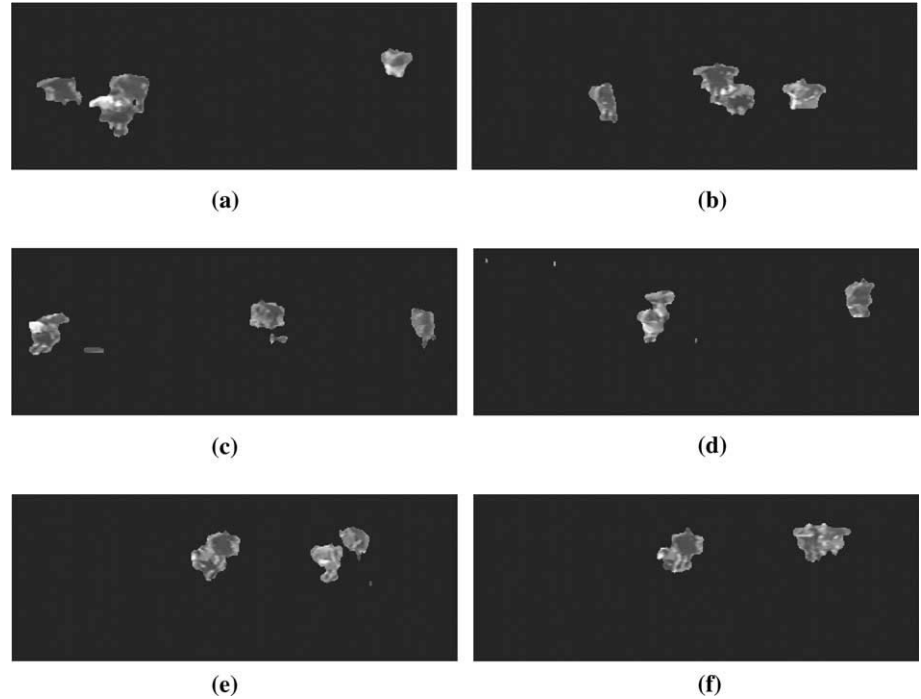


Fig. 15. Player pixels extracted from the images of Fig. 14.

The adapted hybrid color space resulting from the analysis of these representative player pixels is the plane defined by the CIE x chromaticity coordinate and the Ch_2 color component of the (Y, Ch_1, Ch_2) Carron's space [12].

The images of Fig. 18 show how the player pixels of the images of Fig. 15 are classified in this adapted hybrid color space. Note that the three players of the image of Fig. 18(d) are identified without ambiguity whereas they are visually difficult to distinguish in the image of Fig. 14(d), except if we look at their shadows. Furthermore, touching players in the images are represented by distinct regions (see the images of Figs. 18(a)–(c) and (e)).

There is only one error which appears in the image of Fig. 18(f) where the white label region standing between the red and green labels regions does not correspond to any soccer player.

As for the first example, we compare the results of player pixel classification achieved in the adapted hybrid color space and in different classical color spaces. For this purpose, the player pixels are interactively assigned to the different classes (see Fig. 19).

The classification is achieved for all the images of Fig. 15 in all the classical color spaces used in the first example (see Section 3.3.1). The (L^*, C_{ab}^*, h_{ab}) , the (L^*, C_{uv}^*, h_{uv}) and the (L^*, a^*, b^*) spaces are the three classical color spaces which yield the lowest error rates. The classification error rates of these spaces for all the images of Fig. 15 are

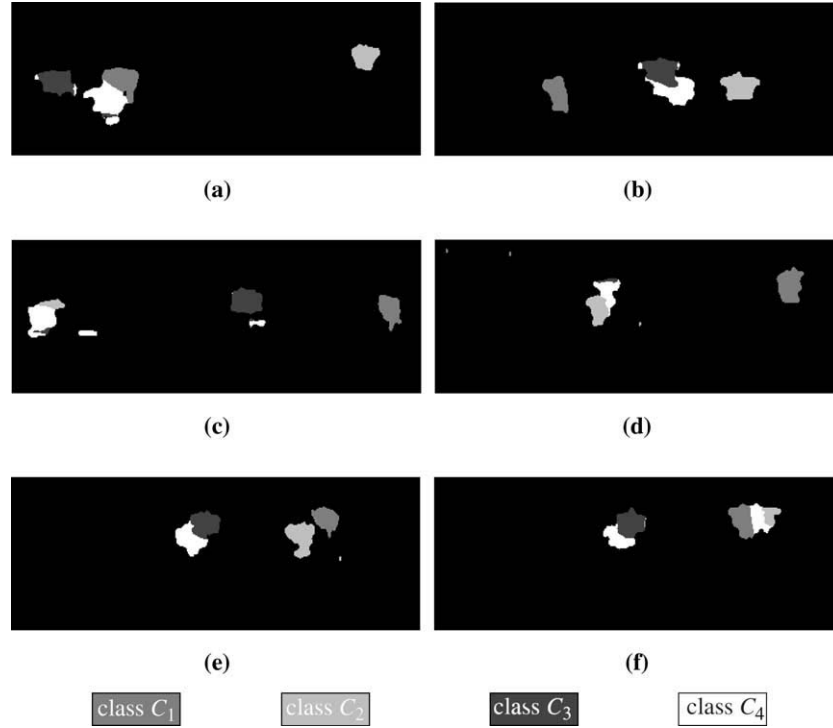


Fig. 18. Player pixels of the images of Fig. 15 classified in the adapted hybrid color space (x, Ch_2) .

given in Table 2. The error rates achieved with these three color spaces are compared with those obtained with the (x, Ch_2) space, which is the adapted hybrid color space.

For all the images of Fig. 15, this two-dimensional adapted hybrid color space provides better classification results than the (R, G, B) acquisition space. The classification error rates achieved in the adapted hybrid color space are higher for the images of Figs. 15(b) and (f), similar for the image of Fig. 15(e), and lower for the images of Figs. 15(a), (c), and (d) than the error rates associated with the three classical color spaces leading to the lowest error rates. So, the use of the adapted hybrid color space improves the classification results for three images among the six ones of Fig. 15. Let us consider the three images of Figs. 15(b), (e), and (f) for which the adapted hybrid color space does not provide the lowest error rate. These three images are so difficult to segment, that there does not exist one single classical color space which is adapted to all of them. Indeed, the (L^*, C_{uv}^*, h_{uv}) space is the best one for the images of Figs. 15(b) and (f), while the (L^*, C_{ab}^*, h_{ab}) space provides the best results for the image of Fig. 15(e).

3.3.3. Dimension of the adapted hybrid color space

For these two examples, the two adapted hybrid color spaces are planes and their dimensions are not equal to three, as for classical color spaces.

212 N. Vandenbroucke et al. / Computer Vision and Image Understanding 90 (2003) 190–216

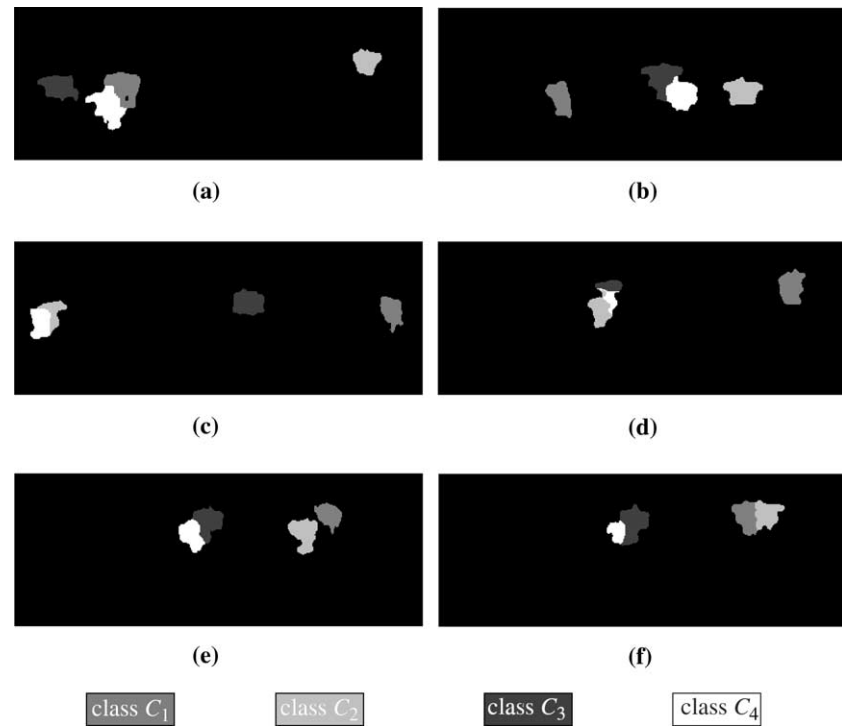


Fig. 19. Player pixels of the images of Fig. 15 interactively classified.

Table 2
Classification error rates of player pixels of Fig. 15

Color space	Classification error rates ε						Mean error rate $\bar{\varepsilon}$
	Image (a)	Image (b)	Image (c)	Image (d)	Image (e)	Image (f)	
$(L^*, C_{ab}^*, h_{ab}^*)$	7.44	5.19	16.82	22.14	4.58	12.00	11.36
$(L^*, C_{uv}^*, h_{uv}^*)$	12.59	4.65	16.92	24.53	4.71	7.97	11.89
(L^*, a^*, b^*)	8.41	9.43	13.30	16.72	4.98	21.79	12.44
(x, Ch_2)	7.09	11.13	11.40	14.01	4.76	25.07	12.24
(R, G, B)	23.37	16.78	32.03	23.82	10.69	33.60	23.38

As above mentioned, the sequential procedure of the adapted hybrid color space construction stops when the stopping criterion decreases. This strategy yields the dimension D of the adapted hybrid color space. In order to evaluate the relevance of this dimension, the images of the two examples are segmented using the six first most discriminating hybrid color spaces selected during the sequential construction procedure.

Figs. 20 and 21 display, for the examples 1 and 2, the evolution of (a) the stopping criterion and (b) the classification error rate with respect to d , the dimension of the most discriminating hybrid color space selected at step d of the construction procedure. These two figures show that when the value of the stopping criterion J_{stop}^d begins to decrease, the classification error rate ε begins to increase.

In the first example, the stopping criterion reaches a local maximum at the second step of the construction procedure which generates a two-dimensional adapted hybrid color space (see Fig. 20(a)). Fig. 20(b) shows that the mean classification error rate associated with the so-determined two-dimensional adapted hybrid color space is the lowest one. This example illustrates the relevance of the stopping criterion which is integrated in the selection procedure.

In the second example, the stopping criterion decreases after the second step of the construction procedure, but it reaches its maximum value at the fourth step (see Fig. 21(a)). Fig. 21(b) indicates that the two-dimensional adapted hybrid color

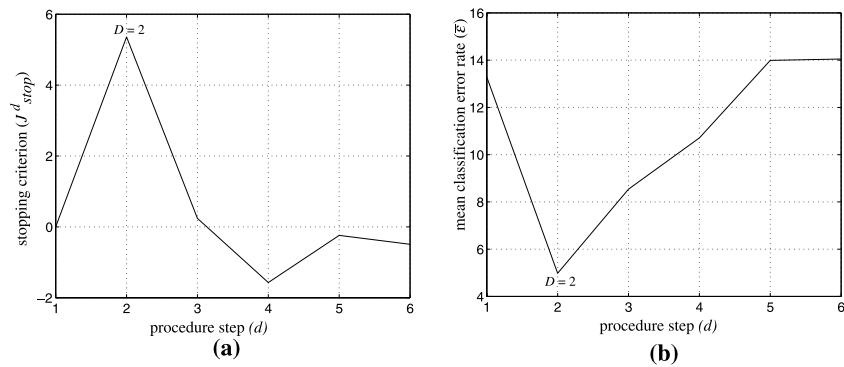


Fig. 20. Stopping criterion and mean classification error rate for example 1.

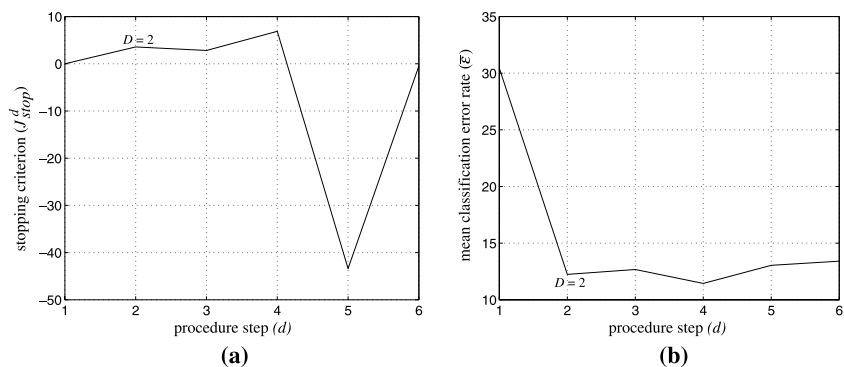


Fig. 21. Stopping criterion and mean classification error rate for example 2.

214 N. Vandenbroucke et al. / *Computer Vision and Image Understanding* 90 (2003) 190–216

space provides a slightly higher mean classification error rate than the four-dimensional most discriminating hybrid color space. However, in the context of a real-time implementation of the segmentation scheme, the classification in a four-dimensional space is more time consuming than the segmentation achieved in the two-dimensional adapted hybrid color space. So, in this example, the slight improvement of the classification results achieved in the four-dimensional most discriminating hybrid color space is not justified in terms of computing time.

Other experiments reported in [40–45] have shown that the adapted hybrid color space rarely coincides with one of the classical color spaces, even when its dimension is equal to three. Indeed, the adapted hybrid color space is determined by means of an analysis of the data discriminating properties whereas the classical color spaces respect psycho-visual and physical properties. This is one of the reasons why the color components of the adapted hybrid color space do not belong necessarily to the same classical color space. Furthermore, the colors of pixels belonging to the same region are more homogeneous when they are analyzed in the adapted hybrid color space than when they are analyzed in one of the classical color spaces.

4. Conclusion

In this paper, we have underlined the influence of the color spaces on color image segmentation results. As there does not exist any classical color space which provides satisfying results for the segmentation of all kinds of images, we have proposed to classify pixels in hybrid color spaces which are specifically designed to yield the best discrimination. These spaces, called the adapted hybrid color spaces, are built by means of a sequential procedure in a supervised context. They are constituted of a small number of color components which are as less correlated as possible and which can belong to different classical color spaces.

In order to evaluate the performance of our approach, we have applied this adapted hybrid color space representation to soccer color image analysis. When the colors of the players are heterogeneous, the segmentation achieved in the (R, G, B) acquisition space yields poor results. On the other hand, when analyzing their colors in the adapted hybrid color space, the discriminating properties of the selected color components are enhanced due to the better homogeneity and separability of the different player pixel classes. This explains why regions with heterogeneous colors in the (R, G, B) color space become more homogeneous when they are analyzed in the adapted hybrid color space. Hence, a very simple classification scheme achieved in the adapted hybrid color space provides satisfying results.

These experiments show that there does not exist any classical color space which is well adapted to the discrimination of all the player pixel classes. The adapted hybrid color space, which is designed to discriminate the considered player pixel classes, is a relevant solution for the extraction of soccer players and the identification of their teams.

The adapted hybrid color spaces which are selected in these two chosen examples are planes that provide better results than the three-dimensional classical color

spaces. In these cases, the classification in this hybrid color space is less time consuming than in one of the classical color spaces. In terms of computation complexity and segmentation results, the use of the adapted hybrid color space is a relevant solution for real-time color image segmentation.

We are presently working on the generalization of our approach to the selection of the most discriminating color texture features for image segmentation [42,44].

References

- [1] T. Bebie, H. Bieri, Soccerman: reconstructing soccer games from video sequences, in: IEEE Int. Conf. on Image Processing, Chicago, USA, 1998, pp. 898–902.
- [2] C.C. Brunner, A.G. Maristany, D.A. Butler, D. Vanleuween, J.W. Funck, An evaluation of color spaces for detecting defects in Douglas-fir veneer, *Ind. Metrol.* 2 (3 and 4) (1992) 169–184.
- [3] H.-D. Cheng, X.-H. Jiang, Y. Sun, J. Wang, Color image segmentation: advances and prospects, *Pattern Recognition* 34 (12) (2001) 2259–2281.
- [4] A. Cumani, A note on the gradient of multi-image edge detection in multispectral images, *Comput. Vision Graphics Image Process.* 53 (1991) 40–51.
- [5] M. Dash, H. Liu, Feature selection for classification, *Intelligent Data Anal.* 1 (3) (1997) 131–156.
- [6] C. Firmin, D. Hamad, J. Postaire, R.D. Zhang, Feature extraction and selection for fault detection in production of glass bottles, *Machine Graphics Vision Int. J.* 5 (1) (1996) 77–86.
- [7] Y.H. Gong, L.T. Sin, C.H. Chuan, H. Zhang, M. Sakauchi, Automatic parsing of TV soccer programs, in: *Int. Conf. on Multimedia Computing and Systems*, Washington, DC, USA, 1995, pp. 167–174.
- [8] J.F. Hair, R.E. Anderson, R.L. Tatham, W. Black, *Multivariate Data Analysis*, Prentice-Hall/Location Englewood Cliffs, NJ, 1998.
- [9] M.A. Ismail, M.S. Kamel, Multidimensional data clustering utilizing hybrid search strategies, *Pattern Recognition* 22 (1989) 75–89.
- [10] T. Kawashima, K. Yoshino, Y. Aoki, Qualitative image analysis of group behaviour, in: *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Seattle, USA, 1994, pp. 690–693.
- [11] R. Kouassi, P. Gouton, M. Paindavoine, Approximation of the Karhunen–Loève transformation and its application to colour images, *Signal Processing: Image Commun.* 16 (6) (2001) 541–551.
- [12] P. Lambert, T. Carron, Symbolic fusion of luminance–hue–chroma features for region segmentation, *Pattern Recognition* 32 (11) (1999) 1857–1872.
- [13] J.H. Lee, B.H. Chang, S.D. Kim, Comparison of colour transformations for image segmentation, *Electron. Lett.* 30 (20) (1994) 1660–1661.
- [14] Y.W. Lim, S.U. Lee, On the color image segmentation algorithm based on the thresholding and the fuzzy *c*-means techniques, *Pattern Recognition* 23 (9) (1990) 935–952.
- [15] E. Littman, E. Ritter, Colour image segmentation: a comparison of neural and statistical methods, *IEEE Trans. Neural Networks* 8 (1) (1997) 175–185.
- [16] E. Littmann, H. Ritter, Adaptive color segmentation. A comparison of neural and statistical methods, *IEEE Trans. Neural Networks* 8 (1) (1997) 175–185.
- [17] J. Liu, Y. Yang, Multiresolution color image segmentation, *IEEE Trans. Pattern Anal. Machine Intell.* 16 (7) (1994) 689–700.
- [18] J. Mukherjee, Mrf clustering for segmentation of color images, *Pattern Recognition* 23 (2002) 917–929.
- [19] M. Naemura, A. Fukuda, Y. Mizutani, Y. Izumi, Y. Tanaka, K. Enami, Morphological segmentation of sport scenes using color information, *IEEE Trans. Broadcasting* 46 (3) (2000) 181–188.
- [20] R. Ohlander, K. Price, D.R. Reddy, Picture segmentation using a recursive region splitting method, *Comput. Graphics Image Process.* 8 (1978) 313–333.
- [21] Y.I. Ohta, T. Kanade, T. Sakai, Color information for region segmentation, *Comput. Graphics Image Process.* 13 (1980) 222–241.

216 N. Vandenbroucke et al. / *Computer Vision and Image Understanding* 90 (2003) 190–216

- [22] D.K. Panjwani, G. Healey, Markov random field models for unsupervised segmentation of textured color images, *IEEE Trans. Pattern Anal. Machine Intell.* 17 (10) (1995) 939–954.
- [23] S.H. Park, I.D. Yun, S.U. Lee, Color image segmentation based on 3-D clustering: morphological approach, *Pattern Recognition* 31 (8) (1998) 1061–1076.
- [24] F. Pernkopf, P. O'Leary, Automatic inspection system for detection and classification of flaws on turned parts, in: *IAPR Int. Conf. on Quality Control by Artificial Vision*, Le Creusot, France, 2001, pp. 359–364.
- [25] C.A. Poynton A guided tour of color space, in: *SMPTE Advanced Television and Electronic Imaging Conf.*, San Francisco, 1995, 167–180.
- [26] C.R. Rao, *Linear Statistical Inference and its Applications*, Wiley, New York, 1973.
- [27] J. Romeder, *Mé thodes et programmes d'analyse discriminante*, Dunod, Paris, 1973, ISBN 2-04-007719-7.
- [28] S.J. Sangwine, R.E.N. Horne, *The Colour Image Processing Handbook*, Chapman & Hall, London, 1998.
- [29] P. Scheunders, A genetic *c*-means clustering algorithm applied to color image quantization, *Pattern Recognition* 30 (6) (1997) 859–866.
- [30] Y. Seo, S. Choi, H. Kim, K. Hong, Where are the ball and players? Soccer game analysis with color-based tracking and image mosaick, in: *IAPR Int. Conf. on Image Analysis and Processing*, Firenze, Italy, 1997, 196–203.
- [31] L. Shafarenko, M. Petrou, J.V. Kittler, Histogram based segmentation in a perceptually uniform color space, *IEEE Trans. Image Process.* 7 (9) (1998) 1354–1358.
- [32] G. Sharma, H.J. Trussell, Digital color imaging, *IEEE Trans. Image Process.* 6 (7) (1997) 901–932.
- [33] P. Somol, P. Pudil, J. Novovicova, P. Paclik, Adaptive floating search methods in feature selection, *Pattern Recognition Lett.* 20 (8) (1999) 1157–1163.
- [34] S. Sudo, S. Ozawa, Scene analysis of soccer game, in: *IAPR Int. Conf. on Quality Control by Artificial Vision*, Trois-Rivi res, Canada, 1999, 119–123.
- [35] K. Takahashi, H. Nakatani, K. Abe, Color image segmentation using ISODATA clustering method, in: *Second Asian Conf. on Computer Vision*, Singapore, vol. 1, 1995, 523–527.
- [36] S. Tominaga, Color classification of natural color images, *Color Res. Appl.* 17 (4) (1992) 230–239.
- [37] A. Trémeau, N. Borel, A region growing and merging algorithm to color segmentation, *Pattern Recognition* 30 (7) (1997) 1191–1203.
- [38] A. Trémeau, P. Colantoni, Regions adjacency graph applied to color image segmentation, *IEEE Trans. Image Process.* 9 (4) (2000) 735–744.
- [39] T. Uchiyama, M.A. Arbib, Color image segmentation using competitive learning, *IEEE Trans. Pattern Anal. Machine Intell.* 16 (12) (1994) 1197–1206.
- [40] N. Vandenbroucke, L. Macaire, J. Postaire, Soccer player recognition by pixels classification in an hybrid color space, in: *Multispectral and Hyperspectral Imagery III*, Orlando, USA, vol. 3071, 1997, pp. 23–33.
- [41] N. Vandenbroucke, L. Macaire, J. Postaire, Color pixels classification in an hybrid color space, in: *IEEE Int. Conf. on Image Processing*, Chicago, USA, vol. 180, 1998, pp. 176–180.
- [42] N. Vandenbroucke, L. Macaire, J. Postaire, Color image segmentation by supervised pixel classification in a color texture feature space. Application to soccer image segmentation, in: *IAPR Int. Conf. on Pattern Recognition*, Barcelona, Espana, vol. 3, 2000, pp. 625–628.
- [43] N. Vandenbroucke, L. Macaire, J. Postaire, Color systems coding for color image processing, in: *Int. Conf. on Color in Graphics and Image Processing*, Saint-Etienne, France, vol. 1, 2000, pp. 180–185.
- [44] N. Vandenbroucke, L. Macaire, J. Postaire, Unsupervised color texture features extraction and selection for soccer images segmentation, in: *IEEE Int. Conf. on Image Processing*, Vancouver, Canada, vol. 2, 2000, pp. 800–803.
- [45] N. Vandenbroucke, L. Macaire, C. Vieren, J. Postaire, Contribution of color classification to soccer players tracking with snakes, in: *IEEE Int. Conf. on System, Man, and Cybernetics*, Orlando, USA, vol. 4, 1997, pp. 3660–3665.
- [46] S.D. Zenzo, A note on the gradient of a multi-image, *Comput. Vision Graphics Image Process.* 33 (1986) 116–125.

Color person image indexing and retrieval

D. Muselet, L. Macaire et J. G. Postaire

Machine Graphics and Vision

volume 11, numéro 2/3

pages 257-283 - 2002

COLOR PERSON IMAGE INDEXING AND RETRIEVAL

Damien Muselet, Ludovic Macaire and Jack-G rard Postaire

Laboratoire d'Automatique I³D

Universit  des Sciences et Technologies de Lille

Cit  Scientifique - B timent P2 - 59655

Villeneuve d'Ascq - FRANCE

dm@i3d.univ-lille1.fr

Abstract

In the context of image indexing for retrieval purpose, colored object recognition methods tend to fail when the illumination of the objects varies from an image to another. A new approach for indexing images of persons is proposed, which copes with the variations of the lighting conditions. We consider that illumination changes can be described using a simple linear transform. For comparing two images, we transform the target one according to the query one by means of an original color histogram specification based on color invariant evaluation. For the retrieval purpose, we evaluate invariant color signatures of the query image and the transformed target image, through the use of color co-occurrence matrices. Test on real images are very encouraging, with substantially better results than those obtained with other well-established indexing and retrieval schemes.

Key words: Color, Illumination, Invariant features, Indexing, Retrieval, Highlights, Shadows, Color co-occurrence matrices.

1 Introduction

Monitoring of crowd flows in public sites, like airports, multimodal stations in public transport systems, supermarkets, factories... requires a system that determines, for each person, the entrance and exit points as well as the duration of her stay within the site under surveillance. For this purpose, cameras are installed above the entrances and the exits, so as to acquire top view color images of the persons when they enter or leave the controlled area (see image of figure 1). The key problem is then to recognize a leaving person, by comparing her image with all the images of other persons who have entered the site and who have not yet left it.

In practice, each image contains only one person, but as the persons are moving, the images of the same person are different in shape. Furthermore, as the lighting conditions are different for all the cameras, the color components of pixels that correspond to the same parts of the persons are also different.

As the aim of this multi-camera system is to recognize a person who walks under one of the different cameras under uncontrolled lighting conditions and in an uncontrollable position in the field of view of the camera, we propose in this paper an original approach for indexing and retrieving images acquired by means of different color cameras.

In this framework, this recognition problem can be stated in terms of looking for target images among all the images acquired by all the cameras in order to find the images that contain the same person as the one represented in the query

2

image.



Figure 1: Example of person image

This purpose requires a complex image indexing scheme, which consists in extracting robust and efficient characteristic features from target images and from the query image. These image features are typically derived from shape, texture or color properties of the query and target images. Person image retrieval is performed by means of a matching scheme which compares the features of the query person image with those of the target person images. The matching scheme is based on a similarity measure between these image features. The target person images are ranked according to their similarity measures with the query person image, in order to determine which cameras have observed the same person than the person contained in the query image.

There are a lot of works about image indexing based on color features [1], which can be divided into two families. Methods of the first family assume that the target and query images are acquired under the same lighting conditions. Approaches of the second family propose color invariants so that the retrieving schemes are not sensitive to illumination changes.

When the lighting conditions are controlled, color provides powerful and robust information for color image retrieval. In this context, Swain [2] uses the color histogram as an image feature which contains bins, whose values indicate the distribution of the colors in the image. Swain shows that color histograms of images containing multicolored objects are robust and efficient image features. Furthermore, color histograms are stable object representations in case of object occlusion. Swain proposes an image retrieval scheme based on the histogram intersection, which matches the target and query image histograms.

The main drawback of histograms is that they do not represent the spatial distribution of the colors in an image. Then, Huang [3] uses the color correlogram as the image feature, which evaluates the spatial correlation between the colors. Experimental results show that this image feature is efficient for content based image retrieval. However, the color correlogram is quite sensitive to the spatial resolution of the image and so requires a spatial normalization of the similarity measures.

Pass [4] proposes an alternative to color histograms called a joint histogram which contains bins with multidimensional coordinates, namely the color, the edge density, the texturedness, the gradient magnitude and a rank. Since the coordinates of the bins are multidimensional, this data structure requires a very large amount of memory and these joint histograms contain a lot of empty cells.

More and more indexing methodologies integrate color and spatial information as the Spatial-Chromatic-Histogram proposed by Cinque [5] which indicates the spatial location and repartition of pixels having the same color. As this spatial

location is represented by the coordinates of the baricenter of the pixels having the same color, this feature is sensitive to any spatial translation in the image.

For the purpose of indexation, Park [6] proposes two original image features, the Modified Color Adjacency Graph (MCAG) and the Spatial Variance Graph (SVG), which take into account only the colors associated to the largest pixel populations. The MCAG and the SVG indicate the spatial adjacency and the spatial spread of pixels having one of these considered colors, respectively. In order to evaluate a similarity measure between the query and one of the target images, Park expands the histogram intersection defined by Swain into the graph intersection. Even if these image features provide very good results, this scheme is too complex to be implemented for real-time retrieval.

In order to reduce the image feature sizes, many authors propose to first segment the images, so as to extract image features from the resulting regions.

Kankanhalli [7] segments the images thanks to a pixel classification scheme. The vectorial image feature contains the mean color and the population of pixels assigned to the each class. The retrieval, which is based on a specific disparity measure in order to compare these vectorial features, does not take into account the spatial interaction between the pixels.

Fuertes [8] presents a similarity measure which analyzes the color and the shape of the regions of homogeneous colors. The hierarchical retrieval scheme is divided into two successive steps. First, the retrieval procedure matches regions of the target and query image of similar color means. Second, the shape similarities between the matched target and query image regions are measured. Thanks to the model of shape deformation proposed by Nastar [9], he evaluates the deformation he would have to apply to a region of the query image in order to transform it into a matched region of the target image.

These schemes, which depend on the segmentation results, are restricted to images acquired under strictly controlled lighting conditions. Otherwise, the regions extracted from each image would not be similar, even if the scenes observed by the cameras were similar.

Since all these above presented features are quite sensitive to variations of color and intensity of the illumination, Funt [10] proposes to transform the pixel color components into color invariant features, which are analyzed by the indexing scheme. The color invariant features are based on the two following assumptions. The spectral response of each *RGB* sensor is sufficiently narrow so that it can be considered as being only sensitive to one of the three wavelengths characterizing these primary colors. Secondly, he assumes that the illumination of a surface represented by neighboring pixels is constant. So he proposes color invariants which takes into account the interaction between neighboring pixels. These color invariant features are the derivative (Laplacian or first directional derivatives) of the logarithm of the color components, which can be considered as the neighborhood-based ratios. These ratios are relatively insensitive to changes in the incident illumination. For the purpose of retrieval, these color invariants are taken into account by the Swain's method.

Adjehro [11] explains the performance of such a neighborhood-based ratios approach in image retrieval. He shows that since the color ratio distribution follows the Cauchy probability law, such ratios are efficient even when the colors are coded in different color representation spaces such as YIQ and $I_1 I_2 I_3$.

Drew [12] studies the effects of the illumination changes on the performance of retrieval systems. He assumes that the color components of a pixel of the query image can be deduced from the color components of the pixel of the target image which corresponds to the same observed surface of the scene by means of the simple least-square estimator. This linear transformation requires that the content of the two images are strictly identical, so that each pixel of the query image and the pixel of the target image with the same spatial coordinates represent the same observed surface. These conditions are too restrictive for real image analysis.

Gevers [13] proposes different retrieval systems which analyze color images containing only one object. Therefore, he presents a set of color invariant features which are based on the same assumptions than those used by Funt. They are composed of a hue and three color ratios which are designed to be robust to any change in the direction and the intensity of the illumination. Then, he defines color image features based on thresholded histograms of these specific color invariant features in order to be invariant to the spatial resolution of the image as well as to the view-point. Since the features are sensitive to illumination color changes, Gevers describes and proposes new color neighborhood ratios [14]. Extensive experiments are carried out on color images in order to evaluate their invariance to variations of image acquisition conditions, such as viewing direction, highlight presence, illumination intensity variations and illumination colors. These experimental results show that the proposed neighborhood ratios are the most appropriate

color components for the purpose of image indexing when illuminations are unknown and non-controlled. To improve these results, Gevers proposes to combine color and shape invariant features and builds an unified high-dimensional invariant feature space [15]. The shape invariant is based on invariant edges, which are detected by means of the local Di Zenzo operator [16]. Finally, the proposed similarity measure analyses a 4D histogram $H(i, j, k, l)$ which counts the number of invariant edge triples, whose gradient values are i, j, k respectively, and which generates the angle l in the image.

Kim assumes that the color contents of an image can be represented by the chromaticity coordinates of the significant bins of a chromaticity histogram [17]. He transforms the color components into two ratios which form the so-called 2D perspective chromaticity space. In order to discount lighting changes, he defines a vectorial image feature by means of the normalized chromaticity coordinates of the significant bins. Therefore, the results show that normalized coordinates of the significant bins are not sensitive to changes in illumination color. However, the results of this procedure depend on the significant bin detection which is reliable only with simple content images.

In this paper, we propose an original method for indexing and retrieval person images acquired by different cameras under non-controlled and different illumination conditions. So, the proposed image features have to be invariant to the following differences between the query and target images which contain the same person:

- the spatial resolution of images,
- the spatial position and shape of the observed person in the image,
- the color, the intensity and the geometry of the incident illumination.

As different combinations of heads and shoulders will appear in the images, the retrieval scheme is based on a separate comparison of the head and shoulder pixels of the query and target images. After elimination of the background pixels that represent the pavement, the remaining pixels are assigned to a head class and a shoulder class, thanks to a color texture analysis (section 2).

Since the images are affected by shadows and bright areas, hereafter called shadows and highlights, due to the geometrical characteristics of the lighting devices, we propose, in section 3, a preprocessing strategy that eliminates these shadows and highlights.

A query image and the target image of the same person can be acquired by different cameras under different lighting conditions. Consequently, the colors of the head and shoulder pixels in the query image can be different from the colors of these pixels in the target image. To cope with this problem, we consider the images by pairs which are composed of the query image and one of the target images of the database. In order to compare these different pairs of images, the color components of the pixels of the target image are transformed using the color histogram specification scheme which is presented in section 4.

For indexing purpose, invariant color features of the different pairs of images are evaluated by means of specific color co-occurrence matrices described in section 5.

Our approach is compared with well-established procedures in the context of person image retrieval for access control purpose. The experiments described in section 6 demonstrate the efficiency of our scheme.

2 Color texture-based pixel classification

The person images contain three pixel classes which represent the background (i.e the pavement), the head and the shoulders of the observed person (see image of figure 1).

Generally, simple features can characterize the area representing the pavement under each camera. In our case, the observed persons are walking on a dark pavement, so that the background pixels can be extracted thanks to a simple thresholding procedure which takes into account the intensity of the pixels [18]. So, black lines could appear on some images because of the background extraction procedure, which does not discriminate background pixels from brown person pixels (see images of figure 2). We consider now only the pixels which represent the person.

Taking advantage of the a priori knowledge about the content of the images, it is pertinent to distinguish the shoulder from the head pixels in order to characterize a person.

Since the images are acquired under non-controlled lighting conditions, a color-based classification scheme would provide poor results. That leads us to consider the spatial homogeneity of pixel colors and propose to classify pixels into two classes, the head and the shoulder classes.

Engelson [19] proposes a grey-level texture signature of each pixel which is the number of its neighboring pixels whose intensities differ from its intensity by more than a fixed value. We generalize this approach to color image analysis.

For this purpose, we evaluate a spatial local variation between the (R, G, B) color components of each pixel and those of its 3x3 neighboring pixels. Let us consider a pixel P of the color image I , whose color is $\vec{c}(P) = (R(P), G(P), B(P))$. In order to cope with intensity non uniformity of the illumination, the spatial local variations $V_R(P)$, $V_G(P)$ and $V_B(P)$ are evaluated as normalized mean square differences:

$$\begin{cases} V_R(P) = \frac{\sum_{i=1}^8 \sqrt{(R(P) - R(P_i))^2}}{R(P)} \\ V_G(P) = \frac{\sum_{i=1}^8 \sqrt{(G(P) - G(P_i))^2}}{G(P)} \\ V_B(P) = \frac{\sum_{i=1}^8 \sqrt{(B(P) - B(P_i))^2}}{B(P)} \end{cases} \quad (1)$$

where P_i is one of the 3x3-neighboring pixels of P .

A binary image, denoted HV contains only high variation pixels:

$$\begin{cases} HV(P) = 1 \text{ if } (V_R(P) > TH_R) \\ \quad \text{and } (V_G(P) > TH_G) \\ \quad \text{and } (V_B(P) > TH_B) \\ = 0 \text{ otherwise} \end{cases} \quad (2)$$

where the mean values of the spatial local variations TH_R , TH_G and TH_B are expressed as:

$$\begin{cases} TH_R = \frac{\sum_{P \in I} V_R(P)}{Card(I)} \\ TH_G = \frac{\sum_{P \in I} V_G(P)}{Card(I)} \\ TH_B = \frac{\sum_{P \in I} V_B(P)}{Card(I)} \end{cases} \quad (3)$$

and where $Card(I)$ is the size of the image I .

The labellisation step consists in assigning to the head class, denoted w^{He} , the pixels whose 9x9 neighborhood contains a number of high variation pixels ($HV(P) = 1$) which is higher than a pre-specified number, set experimentally to 40.

Since the region of the image constituted by the set of pixels assigned to the head class w^{He} can contain some holes, all the remaining pixels which are located between any head pixel of w^{He} and the spatial center of gravity of these head pixels are assigned to w^{He} . The so obtained set of head pixels represents the head of the observed person. The other pixels are assigned to the shoulder class, denoted w^{Sh} .

In order to illustrate the performance of this segmentation scheme, we consider three person images (a), (d) and (g) of figure 2 where the background pixels have been extracted and displayed in black. The size of these images is 700x700.

6

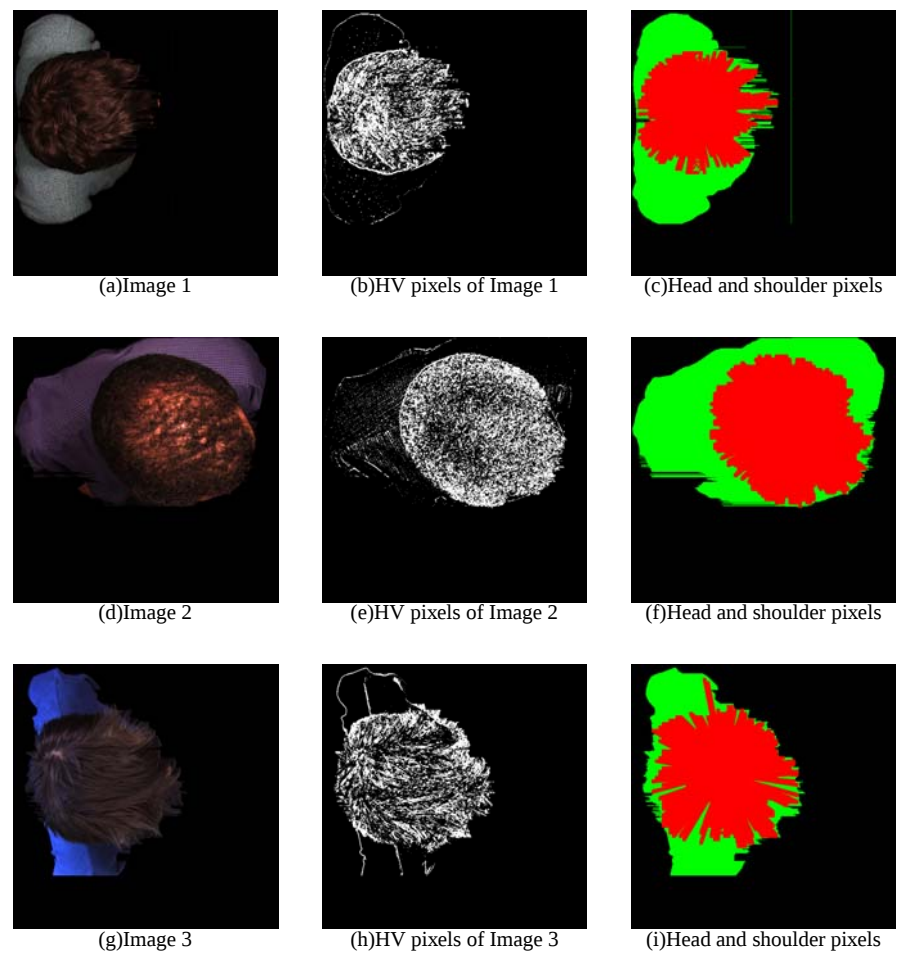


Figure 2: Person image pixel classification

The images (b), (e) and (h) of figure 2 which display the processed binary images HV , show that the head and shoulder pixels can be discriminated thanks to the proposed spatial density measure of the HV pixels.

The spatial repartitions of the pixels assigned to classes w^{He} and w^{Sh} are displayed in the images (c), (f) and (i) of figure 2. The two first images show that the two classes correspond effectively to the head and the shoulders of the observed person.

The third image shows a case where the head is poorly identified. However, the purpose of this classification is not to segment the image and we will see that these artifacts do not affect the retrieval performance.

3 Illumination correction

Let us consider the image of figure 3 which is obviously affected by highlights and shadows, here due to the non-uniformity of the lighting device. Since image features are generally sensitive to these kinds of artifacts, we propose to remove them. This strategy is preferred to the introduction of restrictive and strong assumptions about the uniformity of the intra image illumination, which would be unrealistic for an indexing system using a database constituted of real color images of persons acquired under uncontrollable lighting conditions.

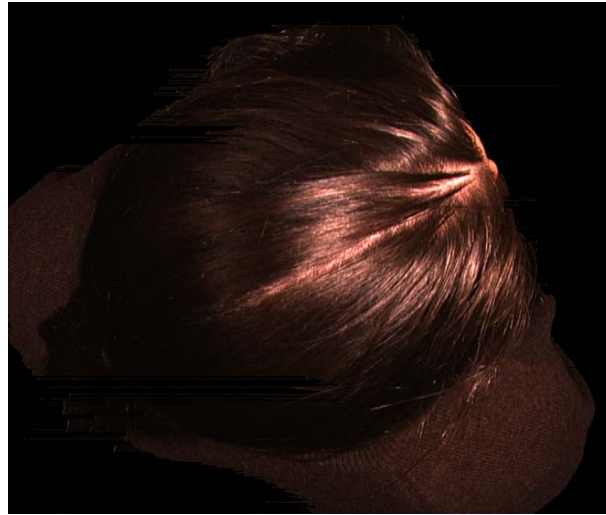


Figure 3: Example of non-uniform illuminated person image

The procedure is divided into two successive steps. The pixels which are affected by shadows or highlights are detected thanks to the analysis described in the first part of this section. The second part presents the illumination model that allows to transform the pixels affected by the detected highlights and shadows, as described in the third part. A highlight (shadow, respectively) is composed of bright (brown, respectively) areas due to the non-uniform spatial repartition of the incident light.

3.1 Highlight and shadow detection

Our proposition consists first in detecting pixels which are affected by highlights or shadows by considering separately the head and the shoulder pixels.

Let us consider the set of points in a 3D color space defined by the colors $\vec{c}(P) = (R(P), G(P), B(P))$ of the pixels P of an image. The points representing the pixels assigned to class w^{He} and to class w^{Sh} , as introduced in section 2, give rise to two clusters in the color space. When the regions of the image corresponding to the head and to the shoulders of the person are uniform, the two clusters of points corresponding to the two classes tend to be rather compact. On the other hand, in case of color variations due to highlights or shadows, the points representing the colors of the pixels of each class will tend to scatter more widely through the color space.

In order to discriminate the pixels that correspond to shadows or to highlights in each class, we propose to analyze the spread of the colors which are projected on the axis corresponding to the first eigenvector $\vec{v}^i$ of the variance-covariance matrix $\sum^i$ associated to the class w^i , $i = \{He, Sh\}$.

The first step consists in evaluating, for each class w^i , the mean color $\vec{m}^i$, $i = \{He, Sh\}$, of the pixels which are assigned to the head and shoulder classes w^{He} and w^{Sh} as:

$$\vec{m}^i = \frac{\sum_{P \in w^i} \vec{c}(P)}{\text{card}(w^i)}. \quad (4)$$

Then, the variance-covariance matrix is expressed as:

$$\sum^i = \frac{\sum_{P \in w^i} (\vec{c}(P) - \vec{m}^i)^t (\vec{c}(P) - \vec{m}^i)}{(\text{card}(w^i))^2}. \quad (5)$$

The color $\vec{c}(P)$ of each pixel P belonging to w^i is projected on the axis corresponding to the first eigenvector $\vec{v}^i$ of this matrix in order to evaluate the difference $D^i(P)$ between the mean color $\vec{m}^i$ of the class w^i and its projected color as:

$$D^i(P) = \vec{v}^i \cdot (\vec{m}^i - \vec{c}(P)), \quad P \in w^i. \quad (6)$$

For extracting highlight and shadow pixels, we consider separately the pixels which belong to the two classes. We consider that pixels which are assigned to the class w^i and which verify equation (7), or equation (8), are affected by shadows, or highlights, respectively.

$$P \in w^{i,HIG} : \quad P \text{ is an highlight pixel if } D^i(P) > Th^i, \quad (7)$$

$$P \in w^{i,SHA} : \quad P \text{ is a shadow pixel if } D^i(P) < -Th^i, \quad (8)$$

$$P \in w^{i,NOR} : \quad P \text{ is a normal pixel otherwise,} \quad (9)$$

where the thresholds Th^i are experimentally adjusted.

Each class w^i is therefore divided into three subclasses of pixels, namely the class $w^{i,HIG}$ associated to the highlight pixels, the class $w^{i,SHA}$ that contains the shadow pixels and the class $w^{i,NOR}$ which is composed of normal pixels that are affected neither by shadows, neither by highlights.

3.2 Illumination model

The next step consists in modifying the color of shadow and highlight pixels in order to obtain pixels with the color they would have if these artifacts did not affect them. As we have no a priori knowledge about the characteristics of the illumination, we use a simplified model of the reflection.

The color $\vec{c}(P)$ of the elementary surface observed by a camera and represented by the pixel P depends on the illumination $E(\lambda)$, the surface spectral reflectance factor $\beta_P(\lambda)$ and $R(\lambda)$, $G(\lambda)$, $B(\lambda)$ which are the three sensor responses of the camera [20]. In these conditions, the color vector $\vec{c}(P)$ can be written as:

$$\vec{c}(P) = \begin{pmatrix} R(P) \\ G(P) \\ B(P) \end{pmatrix} = \begin{pmatrix} \int_{\lambda=380}^{\lambda=780} R(\lambda)\beta_P(\lambda)E(\lambda)d\lambda \\ \int_{\lambda=380}^{\lambda=780} G(\lambda)\beta_P(\lambda)E(\lambda)d\lambda \\ \int_{\lambda=380}^{\lambda=780} B(\lambda)\beta_P(\lambda)E(\lambda)d\lambda \end{pmatrix}. \quad (10)$$

First, we assume that the three sensors of the camera are sensitive to specific non-overlapping wavelength intervals. The three intervals $[\lambda_0; \lambda_1]$, $[\lambda_1; \lambda_2]$ and $[\lambda_2; \lambda_3]$ are associated with the blue, green and red sensor sensitivities respectively with $380\text{nm} \leq \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq 780\text{nm}$ (see figure 4).

The figure 4 displays the spectral colorimetric functions $R(\lambda)$, $G(\lambda)$ and $B(\lambda)$ of the CIE-1931 [21], which are quite different from the spectral color responses of the color cameras. This figure shows that our assumption neglects the sensibilities of the sensors to the extrema wavelengths.

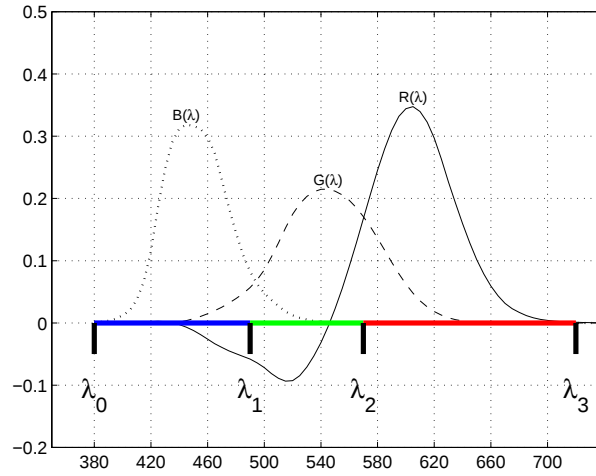


Figure 4: Camera sensors and the three wavelength intervals.

So the color of a pixel is defined as:

$$\vec{c}(P) = \begin{pmatrix} R(P) \\ G(P) \\ B(P) \end{pmatrix} = \begin{pmatrix} \int_{\lambda_2}^{\lambda_3} R(\lambda)\beta_P(\lambda)E(\lambda)d\lambda \\ \int_{\lambda_1}^{\lambda_2} G(\lambda)\beta_P(\lambda)E(\lambda)d\lambda \\ \int_{\lambda_0}^{\lambda_1} B(\lambda)\beta_P(\lambda)E(\lambda)d\lambda \end{pmatrix}. \quad (11)$$

Secondly, we assume that for each sensor wavelength interval, the illumination function $E(\lambda)$ can be expressed by three constants e_R , e_G and e_B associated with the three wavelength intervals, so that:

$$\begin{cases} R = e_R \int_{\lambda_2}^{\lambda_3} R(\lambda)\beta_P(\lambda)d\lambda, \\ G = e_G \int_{\lambda_1}^{\lambda_2} G(\lambda)\beta_P(\lambda)d\lambda, \\ B = e_B \int_{\lambda_0}^{\lambda_1} B(\lambda)\beta_P(\lambda)d\lambda. \end{cases} \quad (12)$$

These assumptions about the illumination model will now allow us to transform the colors of shadow and highlight pixels in order to eliminate abrupt illumination variations in an image.

10

3.3 Color transformation of the highlight and shadow pixels

In the context of highlight and shadow pixel detection, we assume that the pixels assigned to the class w^i represent surfaces lighted by an uniform illumination $E^{i,j}(\lambda)$ characterized by $e_R^{i,j}$, $e_G^{i,j}$ and $e_B^{i,j}$, $i = \{He, Sh\}$ and $j = \{HIG, SHA, NOR\}$, respectively.

We have to modify the color of each shadow and highlight pixel belonging to $w^{i,SHA}$ and $w^{i,HIG}$, so that its new color is that of the surface seen by this pixel if it was placed under a normal illumination, denoted $E^{i,NOR}(\lambda)$.

Considering equation (12), the color of a pixel P belonging to $w^{i,j}$ is expressed as:

$$\vec{c}(P) = \begin{pmatrix} e_R^{i,j} \int_{\lambda_2}^{\lambda_3} R(\lambda) \beta_P(\lambda) d\lambda \\ e_G^{i,j} \int_{\lambda_1}^{\lambda_2} G(\lambda) \beta_P(\lambda) d\lambda \\ e_B^{i,j} \int_{\lambda_0}^{\lambda_1} B(\lambda) \beta_P(\lambda) d\lambda \end{pmatrix} = \begin{pmatrix} R(P) \\ G(P) \\ B(P) \end{pmatrix}. \quad (13)$$

If the surface represented by this pixel P was lighted by an illumination $E^{i,k}(\lambda)$, its color would be $\vec{c}^j(P)$:

$$\vec{c}^j(P) = \begin{pmatrix} e_R^{i,k} \int_{\lambda_2}^{\lambda_3} R(\lambda) \beta_P(\lambda) d\lambda \\ e_G^{i,k} \int_{\lambda_1}^{\lambda_2} G(\lambda) \beta_P(\lambda) d\lambda \\ e_B^{i,k} \int_{\lambda_0}^{\lambda_1} B(\lambda) \beta_P(\lambda) d\lambda \end{pmatrix}. \quad (14)$$

From equations (13) and (14), we can deduce $\vec{c}^j(P)$ from $\vec{c}(P)$ thanks to a linear transformation:

$$\vec{c}^j(P) = \begin{pmatrix} R(P) \frac{e_R^{i,k}}{e_R^{i,j}} \\ G(P) \frac{e_G^{i,k}}{e_G^{i,j}} \\ B(P) \frac{e_B^{i,k}}{e_B^{i,j}} \end{pmatrix} = \begin{pmatrix} R(P) \alpha_R^{i,j,k} \\ G(P) \alpha_G^{i,j,k} \\ B(P) \alpha_B^{i,j,k} \end{pmatrix}, \quad (15)$$

where $\alpha_R^{i,j,k}$, $\alpha_G^{i,j,k}$ and $\alpha_B^{i,j,k}$ are the transformation coefficients.

Equation (15) indicates the linear transform which is applied to the color of a pixel belonging to $w^{i,j}$ by modifying the illumination condition from illumination $E^{i,j}(\lambda)$ to illumination $E^{i,k}(\lambda)$.

Thanks to this equation, we are able to transform the colors of the highlight and shadow pixels belonging to $w^{i,HIG}$ and $w^{i,SHA}$ respectively, so that their surfaces are lighted by the illumination $E^{i,NOR}(\lambda)$. For this purpose, we have to determine the transform coefficients $\alpha_R^{i,HIG,NOR}$, $\alpha_G^{i,HIG,NOR}$, $\alpha_B^{i,HIG,NOR}$, $\alpha_R^{i,SHA,NOR}$, $\alpha_G^{i,SHA,NOR}$ and $\alpha_B^{i,SHA,NOR}$ which are:

$$\alpha_X^{i,j,NOR} = \frac{e_X^{i,NOR}}{e_X^{i,j}} \quad (16)$$

with $X = \{R, G, B\}$, $i = \{HE, SH\}$ and $j = \{HIG, SHA\}$.

In order to determine these coefficients, the color mean values $\vec{m}^{i,j}$ of pixels belonging to w^i are evaluated as:

$$\vec{m}^{i,j} = \begin{pmatrix} m_R^{i,j} \\ m_G^{i,j} \\ m_B^{i,j} \end{pmatrix} = \begin{pmatrix} \frac{\sum_{P \in w^{i,j}} (e_R^{i,j} \int_{\lambda_2}^{\lambda_3} R(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,j})} \\ \frac{\sum_{P \in w^{i,j}} (e_G^{i,j} \int_{\lambda_1}^{\lambda_2} G(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,j})} \\ \frac{\sum_{P \in w^{i,j}} (e_B^{i,j} \int_{\lambda_0}^{\lambda_1} B(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,j})} \end{pmatrix} \quad (17)$$

11

where $i = \{He, Sh\}$, $j = \{HIG, SHA, NOR\}$ and where $card(w^{i,j})$ represents the number of pixels belonging to $w^{i,j}$.

We obtain:

$$\vec{m}^{i,j} = \begin{pmatrix} e_R^{i,j} \frac{\sum_{P \in w^{i,j}} (\int_{\lambda_2}^{\lambda_3} R(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,j})} \\ e_G^{i,j} \frac{\sum_{P \in w^{i,j}} (\int_{\lambda_1}^{\lambda_2} G(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,j})} \\ e_B^{i,j} \frac{\sum_{P \in w^{i,j}} (\int_{\lambda_0}^{\lambda_1} B(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,j})} \end{pmatrix}. \quad (18)$$

We now assume that the pixels P which are assigned to the class w^i represent surfaces whose mean spectral reflectances $\beta^i(\lambda)$ can be estimated with only the highlight, shadow or normal pixels which belong to w^i :

$$\beta^i(\lambda) = \frac{\sum_{P \in w^{i,j}} \beta_P(\lambda)}{card(w^{i,j})} = \frac{\sum_{P \in w^{i,k}} \beta_P(\lambda)}{card(w^{i,k})} \quad (19)$$

with $i = \{He, Sh\}$ and $(j, k) = \{(HIG, NOR), (SHA, NOR)\}$.

From equation (19) we deduce:

$$\begin{cases} \frac{\sum_{P \in w^{i,j}} (\int_{\lambda_2}^{\lambda_3} R(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,j})} = \frac{\sum_{P \in w^{i,k}} (\int_{\lambda_2}^{\lambda_3} R(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,k})} \\ \frac{\sum_{P \in w^{i,j}} (\int_{\lambda_1}^{\lambda_2} G(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,j})} = \frac{\sum_{P \in w^{i,k}} (\int_{\lambda_1}^{\lambda_2} G(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,k})} \\ \frac{\sum_{P \in w^{i,j}} (\int_{\lambda_0}^{\lambda_1} B(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,j})} = \frac{\sum_{P \in w^{i,k}} (\int_{\lambda_0}^{\lambda_1} B(\lambda) \beta_P(\lambda) d\lambda)}{card(w^{i,k})} \end{cases} \quad (20)$$

with $i = \{He, Sh\}$ and $(j, k) = \{(HIG, NOR), (SHA, NOR)\}$.

Considering simultaneously equations (18) and (20), we obtain:

$$\begin{pmatrix} \frac{m_R^{i,k}}{m_R^{i,j}} \\ \frac{m_G^{i,k}}{m_G^{i,j}} \\ \frac{m_B^{i,k}}{m_B^{i,j}} \end{pmatrix} = \begin{pmatrix} \frac{e_R^{i,k}}{e_R^{i,j}} \\ \frac{e_G^{i,k}}{e_G^{i,j}} \\ \frac{e_B^{i,k}}{e_B^{i,j}} \end{pmatrix} = \begin{pmatrix} C_R^{i,j,k} \\ C_G^{i,j,k} \\ C_B^{i,j,k} \end{pmatrix} \quad (21)$$

with $i = \{He, Sh\}$ and $(j, k) = \{(HIG, NOR), (SHA, NOR)\}$.

Equation (21) provides the transform coefficients used in equation (15).

Finally, we consider the color $\vec{c}(P)$ of a highlight pixel (respectively a shadow) which belongs to the class w^i . We know that all the highlight pixels (respectively the shadow) of the same class w^i are characterized by their mean value $m^{i, \vec{HIG}}$ (respectively $m^{i, \vec{SHA}}$) and are lighted by an uniform illumination $E^{i, HIG}(\lambda)$ (respectively $E^{i, SHA}(\lambda)$). To eliminate the effects of these different illuminations, the original color $\vec{c}(P)$ of this pixel is replaced by the color $\vec{c}'(P)$ such as:

$$\begin{cases} \vec{c}'(P) = \begin{pmatrix} R'(P) \\ G'(P) \\ B'(P) \end{pmatrix} = \begin{pmatrix} \frac{m_R^{i, NOR}}{m_R^{i,j}} R(P) \\ \frac{m_G^{i, NOR}}{m_G^{i,j}} G(P) \\ \frac{m_B^{i, NOR}}{m_B^{i,j}} B(P) \end{pmatrix} \\ \text{if } P \text{ belongs to } w^{i,j}, \end{cases} \quad (22)$$

12

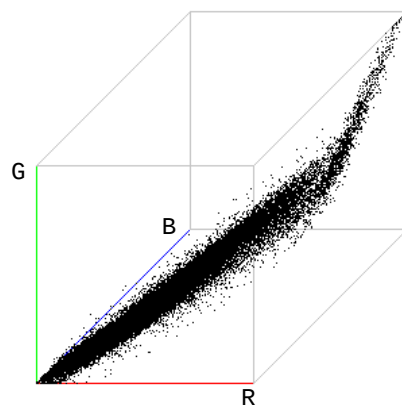
where $i = \{He, Sh\}$ and $j = \{HIG, SHA\}$. This new color corresponds to the color of the same pixel if the observed surface was lighted by the illumination $E^{i,NOR}(\lambda)$.

Finally, the pre-processing of an image is divided into three successive steps:

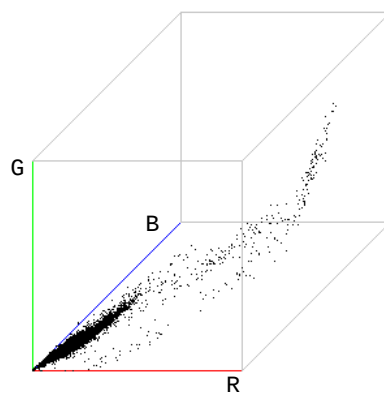
- color texture-based pixel classification which assigns pixels to the shoulder and head classes.
- The pixels which are assigned to one of these two classes are divided into highlight, normal and shadow pixels.
- the color of the highlight and shadow pixels are modified thanks to equation (22).

To illustrate this preprocessing, we apply these successive steps to the image of figure 3 where the pixels that represent the background are displayed as black.

Figure 5 which shows the distributions of the points corresponding to the colors of the head (figure 5(a)) or shoulder (figure 5(b)) pixels in the color space, is a good illustration of the highlights and shadows problem. As some head pixels of image 3 represent surfaces which are affected by highlights and shadows, the spread of the head class is large. On the other hand, as the shoulder pixels don't present any illumination discontinuities, the distribution of the points representing the color of the shoulder pixels is much more concentrated.



(a) Color distribution of head pixels



(b) Color distribution of shoulder pixels

Figure 5: Color distribution of the pixels of the two classes of the image of figure 3

Figure 6 presents the results of the successive steps of the preprocessing applied to the image of figure 3.

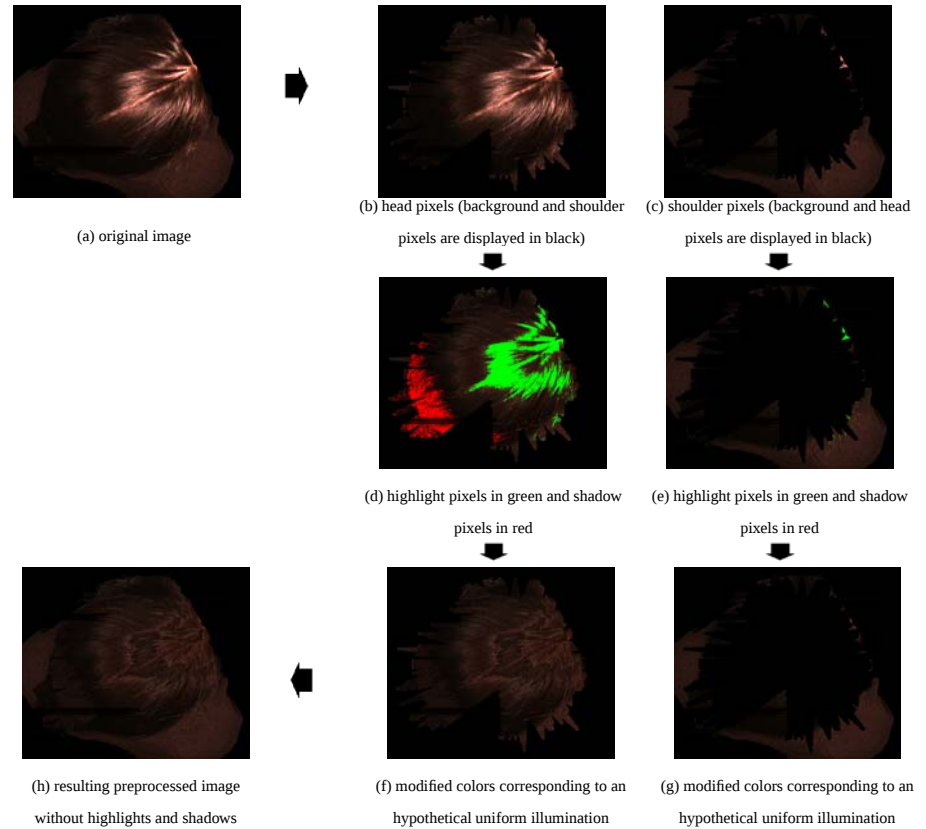


Figure 6: Preprocessing of the image of figure 3

The images of figures 6(b) and 6(c) show that the classification step has generated a class of head pixels and a class of shoulder pixels which roughly correspond to the actual structure of the image. Only a few percentage of pixels have been misclassified, but these errors will not perturb the indexing and retrieval processes. Figures 6(d) and 6(e) display highlight and shadow pixels detected thanks to the thresholding step as green and red respectively. Their colors are then modified so that their new colors correspond to an hypothetical uniform illumination without shadows nor highlights (see figures 6(f) and 6(g)). The image of figure 6(h) contains the preprocessed obtained color image of image of figure 6(a).

Figure 7 represents the points corresponding to the colors of the head pixels after transformation and which are displayed on the image (f) of the figure 6. This distribution has to be compared with that represented by figure 5(a) which represent the colors of the same pixels before transformation. The spread along the main axis is drastically reduced, since the initial large spread was essentially due to highlights and shadows in the image.

As the shoulders of the observed person are not affected by any highlight or shadow (see image (e) of figure 6), the detection step extracts a very few highlight or shadow pixels and so, the color of a very few shoulder pixels are modified.

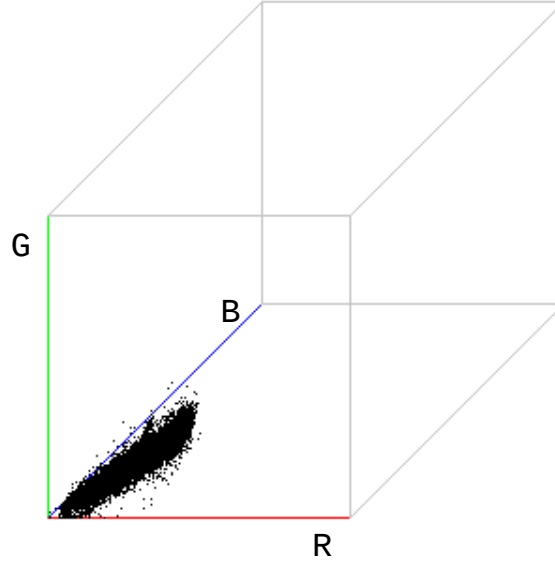


Figure 7: Color distribution of pixels of image (f) of figure 6

Applying this efficient pre-processing to all the images of our data base, we consider that each resulting image represents a scene lighted by an uniform illumination. Since the illumination conditions are specific to each camera, we propose color invariant in the following section.

4 Color invariant by trichromatic histogram specification

The aim of histogram modeling techniques [22], as histogram equalization, is to modify the dynamic range and the contrast of a grey-level image so that its intensity histogram has a desired shape.

Sometimes, it is desirable to control the shape of the output histogram in order to enhance some specific intensity levels in an image. This can be performed by means of an histogram specification procedure.

For the purpose of color image indexing using color invariant, the three 1D trichromatic histograms of the analyzed images are transformed by means of a 1D-histogram specification scheme.

The trichromatic histogram specification consists in transforming each color component of each pixel of a target image, so that its transformed color component tends to be similar to the color component of the pixel of the query image which represents the same surface.

Due to the pre-processing step described in the third section, we consider that all the surfaces represented by pixels of a processed image are lighted by the same illumination. So, we consider together the shoulder and head pixels of the images during the histogram specification procedure.

Let us denote I_{que} the query image acquired under an uniform illumination $E_{que}(\lambda)$ characterized by the three constants e_{que_R} , e_{que_G} and e_{que_B} .

From equation (12), the color $\tilde{c}(P_{que})$ of a pixel P_{que} of the query image I_{que} can be expressed as:

16

$$\tilde{c}(P_que) = \begin{pmatrix} R(P_que) \\ G(P_que) \\ B(P_que) \end{pmatrix} = \begin{pmatrix} e_que_R \int_{\lambda_2}^{\lambda_3} R(\lambda) \beta_{P_que}(\lambda) d\lambda \\ e_que_G \int_{\lambda_1}^{\lambda_2} G(\lambda) \beta_{P_que}(\lambda) d\lambda \\ e_que_B \int_{\lambda_0}^{\lambda_1} B(\lambda) \beta_{P_que}(\lambda) d\lambda \end{pmatrix}. \quad (23)$$

Let us now consider a target image I_tar_0 , which represents the same person as that of the image I_que , but which is acquired under the illumination $E_tar_0(\lambda)$ characterized by the three constants $e_tar_0_R$, $e_tar_0_G$ and $e_tar_0_B$. If the images I_que and I_tar_0 contain the same person, we assume that we can find the pixel P_tar_0 which represents the same observed surface as the pixel P_que . Its color $\tilde{c}(P_tar_0)$ is expressed as:

$$\tilde{c}(P_tar_0) = \begin{pmatrix} R(P_tar_0) \\ G(P_tar_0) \\ B(P_tar_0) \end{pmatrix} = \begin{pmatrix} e_tar_0_R \int_{\lambda_2}^{\lambda_3} R(\lambda) \beta_{P_tar_0}(\lambda) d\lambda \\ e_tar_0_G \int_{\lambda_1}^{\lambda_2} G(\lambda) \beta_{P_tar_0}(\lambda) d\lambda \\ e_tar_0_B \int_{\lambda_0}^{\lambda_1} B(\lambda) \beta_{P_tar_0}(\lambda) d\lambda \end{pmatrix}. \quad (24)$$

As we consider that P_tar_0 and P_que represent the same observed surface, the spectral reflectance factors $\beta_{P_tar_0}(\lambda)$ and $\beta_{P_que}(\lambda)$ are equal.

Furthermore, we assume that the three spectral sensitivity coefficients $R(\lambda)$, $G(\lambda)$ and $B(\lambda)$ are the same for all the cameras which control the public site.

Considering equations (23) and (24) and applying the approach described in the above section, we can express the color $\tilde{c}(P_tar_0)$ of the pixel P_tar_0 representing a surface lighted by the illumination $E_tar_0(\lambda)$, with the color $\tilde{c}(P_que)$ of the pixel P_que associated with the same surface lighted by the illumination $E_que(\lambda)$, thanks to the following linear equations:

$$\tilde{c}(P_tar_0) = \begin{pmatrix} \frac{e_tar_0_R}{e_que_R} R(P_que) \\ \frac{e_tar_0_G}{e_que_G} G(P_que) \\ \frac{e_tar_0_B}{e_que_B} B(P_que) \end{pmatrix} = \begin{pmatrix} \alpha_R^{tar_0, que} R(P_que) \\ \alpha_G^{tar_0, que} G(P_que) \\ \alpha_B^{tar_0, que} B(P_que) \end{pmatrix}. \quad (25)$$

We assume that if the target and query images represent the same person, we can extend equation (25) to all the pixels of these two images. We determine these coefficients by using the ratio of the mean color components of the two images:

$$\begin{pmatrix} \alpha_R^{tar_0, que} \\ \alpha_G^{tar_0, que} \\ \alpha_B^{tar_0, que} \end{pmatrix} = \begin{pmatrix} \frac{\sum_{P \in I_que} (R(P_que))}{card(I_que)} \\ \frac{\sum_{P \in I_tar_0} (R(P_tar_0))}{card(I_tar_0)} \\ \frac{\sum_{P \in I_que} (G(P_que))}{card(I_que)} \\ \frac{\sum_{P \in I_tar_0} (G(P_tar_0))}{card(I_tar_0)} \\ \frac{\sum_{P \in I_que} (B(P_que))}{card(I_que)} \\ \frac{\sum_{P \in I_tar_0} (B(P_tar_0))}{card(I_tar_0)} \end{pmatrix} \quad (26)$$

where $card(I)$ indicates the number of person pixels of the image I .

In order to compare these two images, we propose to specify the trichromatic histograms of I_tar_0 with the trichromatic histograms of I_que , and subsequently to construct the transformed image I_tar_1 .

The first step consists in determining the three histogram specification coefficients $\alpha_R^{tar_0, que}$, $\alpha_G^{tar_0, que}$ et $\alpha_B^{tar_0, que}$ used to transform the target image I_tar_0 into the image I_tar_1 which will be suitable for comparison with the query image I_que .

For each pair of pixels (P_tar_1, P_tar_0) with the same spatial coordinates, we evaluate the color $\tilde{c}(P_tar_1)$ by transforming the color $\tilde{c}(P_tar_0)$ with the previously determined coefficients as:

$$\tilde{c}(P_{tar_1}) = \begin{pmatrix} R(P_{tar_1}) \\ G(P_{tar_1}) \\ B(P_{tar_1}) \end{pmatrix} = \begin{pmatrix} \alpha_R^{tar_0, que} R(P_{tar_0}) \\ \alpha_G^{tar_0, que} G(P_{tar_0}) \\ \alpha_B^{tar_0, que} B(P_{tar_0}) \end{pmatrix}. \quad (27)$$

These three coefficients are suitable if the two images I_{tar_0} and I_{que} represent the same person, because each pixel of the image I_{tar_0} can be associated with one pixel of the image I_{que} which represents the same observed surface. In this case, the colors of the pixels representing the head and the shoulders in the two images I_{tar_1} and I_{que} would be quite similar and the color features extracted from these two images would be similar. On the other hand, when the two images I_{tar_0} and I_{que} do not represent the same person, the colors of the pixels of I_{tar_1} would be different from the colors of the pixels of I_{que} . It is therefore interesting to consider each pair of images (I_{tar_0} , I_{que}) with a suitable measure of similarity in order to discriminate the two above cases.

The indexing method, presented in the next section, extracts image features in order to compare the query image I_{que} with the target image I_{tar_1} which is the result of the transformation by means of this trichromatic histogram specification.

5 Indexing and retrieval using color co-occurrence matrices

For indexing and retrieval purpose, we use color image features based on color co-occurrence matrices which take into account the spatial and colorimetric interactions between pixels [23]. In this context the color co-occurrence matrices of two images containing the same person are expected to be similar, while they are likely to be different when the images represent different persons.

Since these images contain two pixel classes, it is appealing to consider separately the interaction between the pixels assigned to the head and the shoulder classes.

In order to reduce the computational complexity, the proposed color co-occurrence matrices are computed for two color components selected among the three trichromatic components R , G and B . Let denote c_1 and c_2 these two color components. We evaluate the co-occurrence matrix $Mat_{c_1, c_2}^{I, Sh}$ which measures the interaction between the components c_1 and c_2 of the shoulder pixels of the image I and $Mat_{c_1, c_2}^{I, He}$ which measures the interaction between the same components of the head pixels of the image I .

Let $Mat_{c_1, c_2}^{I, Sh}(u, v)$ be the number of times a shoulder pixel of image I , with a level of the color component c_2 equal to v is in the 8-neighborhood of a shoulder pixel of the same image with a level of the color component c_1 equal to u .

The color co-occurrence matrix is a more suitable tool than the histogram for characterizing the content of an image, since it takes into account the color as well as the spatial interactions between the pixels. Furthermore, the storage of the co-occurrence matrices require $6n^2$ cells, whereas the color histogram needs n^3 cells.

The intersection between the co-occurrence matrices of the pixels which are assigned to the class w^k ($k = \{He, Sh\}$) of the query image I_{que} and the transformed target image I_{tar_1} is an extension of histogram intersection presented by Swain [2] and is expressed as:

$$Inter(Mat_{c_1, c_2}^{I_{que}, k}, Mat_{c_1, c_2}^{I_{tar_1}, k}) = \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} \min(m^{I_{que}, k}(u, v), m^{I_{tar_1}, k}(u, v)) \quad (28)$$

$$\text{where } m^{I, k}(u, v) = \frac{Mat_{c_1, c_2}^{I, k}(u, v)}{\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} Mat_{c_1, c_2}^{I, k}(i, j)}$$

and where n is the number of levels used to quantify the color components. The denominators are simple normalization terms so that this feature is invariant to the spatial resolution of the images.

The more similar the spatial and color interactions between pixels assigned to a same class are, the closer to one is the evaluated intersection between two co-occurrence matrices.

18

As the maximal value of the intersection between two matrices is equal to one, the similarity measure between the query and the transformed target images is the sum of the differences between one and the intersections between the respective co-occurrence matrices of the head and shoulder pixels of the two considered images, such as:

$$SM(I_{que}, I_{tar_1}) = \sum_{k=Sh, He} \sum_{c_1=R, G, B} \sum_{c_2=R, G, B} (1 - Inter(Mat_{C_1, C_2}^{que, k}, Mat_{C_1, C_2}^{tar_1, k}))^2 \quad (29)$$

The lower the value of SM is, the more similar the two compared images are.

6 Experimental results

This indexation approach has been applied to a database composed of images of persons entering or leaving a controlled area of a public site. In order to evaluate the performance of the indexation, 20 persons have been asked to walk 3 times under cameras with different lighting devices and parameters settings, so as to acquire 3 images of each of them under different lighting conditions. We simulate the monitoring of the entrances and exits of the area by looking for target images of this database which contain the same person as the person represented in the query image. Each query image is compared with 59 target images, among which only two represent the person of the query image.

The highlight and shadow pixels are eliminated from the images by means of the procedure described in the third section. The colors of the pixels of target images are transformed thanks to the histogram specifications according to the query image. The intersections between the co-occurrence matrices of the query image and the transformed target images are evaluated. Finally, the target images are sorted according to decreasing values of their similarity measures expressed by equation (29).

Matching performance of a retrieval method can be evaluated by means of two indicators: the rank and the correct match percentile [2]. The rank is the position of one of the two target images $T(I_{que})$ which represents the same person as the query image I_{que} in the sorted list of matching values. The correct match percentile M for each query image is then calculated as:

$$M(I_{que}) = \sum_{T(I_{que})} \frac{n - rank(T(I_{que}))}{n - 1} \quad (30)$$

where n is the number of target images of the database which is equal to 59.

The total match percentile MT is the average of $M(I_{que})$ over all the 20 query images in the experiment. A value of 1 indicates perfect matching and a value of 0.5 indicates the match selection is not better than random.

To demonstrate the efficiency of the preprocessing scheme described in the section 3, we compare the results of our method with and without this preprocessing (see table 1). This comparison shows that the preprocessing improves the total match percentile from 93.60 to 94.74.

Algorithms	Rank 1	Rank 2	Rank 3	> Rank 3	MT
<i>without preprocessing</i>	16	3	6	15	93.60
<i>with preprocessing</i>	12	9	8	11	94.74

Table 1: Impacts of the preprocessing on the matching results. Each query image is compared with 59 target images.

The performance of the proposed color invariant is then compared with classical invariants based on neighborhood ratios such as those of Funt [10] or Gevers [14]. As we want to compare only the invariants and not the indexing methods, we add our classification step to these two classical methods. That is to say that, for each method, we compare the head

pixels (respectively the shoulder pixels) of the target image with the head pixels (respectively the shoulder pixels) of the query image.

Furthermore, as the invariants of Funt and Gevers are based on neighborhood ratios, it is not judicious to index them thanks to co-occurrence matrices which are based on spatial neighborhood interactions. For the indexing and retrieving steps, the most appropriate image features are the color histograms, as proposed by Swain [2], which are moreover recommended by Funt [10]. Experimental results show that the histogram provides better results than co-occurrence matrices when applied to these two ratio based methods.

To test Funt's method, we compute for each pixel, the Laplacian of the logarithm of the color components, while the tested invariants of Gevers are the logarithms of (m_1, m_2, m_3) which are expressed as:

$$\begin{cases} m_1 = \frac{R \times G^{mean}}{R^{mean} \times G}, \\ m_2 = \frac{R \times B^{mean}}{R^{mean} \times B}, \\ m_3 = \frac{G \times B^{mean}}{G^{mean} \times B}. \end{cases} \quad (31)$$

where $X = \{R, G, B\}$ represents one of the color components of the considered pixel and X^{mean} is the mean value of the X component among the neighbors of this pixel.

The table 2 shows that our method provides the best matching performance with these images acquired under different lighting conditions. The Funt's and Gevers's methods provide quite equal matching results.

Algorithms	Rank 1	Rank 2	Rank 3	> Rank 3	MT
Funt	3	1	2	34	70.78
Gevers	6	0	0	34	72.54
<i>Invariant by specification</i>	12	9	8	11	94.74

Table 2: Matching results for three algorithms. Each query image is compared with 59 target images.

Instead of defining an invariant transformation for all images, the proposed invariant is processed thanks to a transformation of one of the target images relatively to the query image. So, the presented invariant is not based only on illumination hypothesis which are whatever too restrictive in real cases.

7 Conclusion

In this paper, we have proposed an original approach for color invariant feature evaluation. First all images are pre-processed in order to transform them as if the scene was observed under a uniform illumination, without shadows and highlights. Then, instead of defining the same invariant for all the images, as well known methods proposed by Gevers [14] or by Funt [10], we consider each couple formed by the query image and one of the target images. If these two images contain the same person, the color co-occurrence matrices of the query image and the transformed target image thanks to the color histogram specification are expected to be similar. Otherwise, these matrices should be different.

The simulation of the monitoring of the entrances and exits of the area, which consists in taking each of these 60 images as a query image in order to retrieve the person among the 59 other images of the database, considered as target images, has shown that our approach works better than two other well established retrieval procedures.

References

- [1] R. Schettini, G. Ciocca, and S. Zuffi, "Color in databases: indexation and similarity," in *International Conference on Color in Graphics and Image Processing*, Saint-Etienne, 2000, vol. 1, pp. 244–249.

- [2] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7(1), pp. 11–32, 1991.
- [3] J. Huang, S. Ravi Kumar, M. Mitra, W. Zhu, and R. Zabih, "Image indexing using color correlogram," *IEEE Conf. Computer Vision Pattern Recognition*, pp. 762–768, 1997.
- [4] G. Pass and R. Zabih, "Comparing images using joint histograms," *Multimedia Systems*, vol. 7(3), pp. 234–240, 1999.
- [5] L. Cinque, G. Ciocca, S. Levialdi, A. Pellicano, and R. Schettini, "Color-based image retrieval using spatial-chromatic histograms," *Image and Vision Computing*, vol. 19, pp. 979–986, 2001.
- [6] I.K. Park, I.D. Yun, and S.U. Lee, "Color image retrieval using hybrid graph representation," *Image and Vision Computing*, vol. 17, pp. 465–474, 1999.
- [7] M.S. Kankanhalli, B.M. Mehtre, and J.K. Wu, "Cluster-based color matching for image retrieval," *Pattern Recognition*, vol. 29(4), pp. 701–708, 1996.
- [8] J.M. Fuertes, M. Lucena, N. Perez de la Blanca, and J. Chamorro-Martinez, "A scheme of colour image retrieval from databases," *Pattern Recognition Letters*, vol. 22, pp. 323–337, 2001.
- [9] C. Nastar and N. Ayache, "Frequency-based nonrigid motion analysis: application to four dimensional medical images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18(11), pp. 1067–1079, 1996.
- [10] B. V. Funt and G. D. Finlayson, "Color constant color indexing," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 17(5), pp. 522–529, 1995.
- [11] D.A. Adjehro and M.C. Lee, "On ratio-based color indexing," *IEEE Trans. on Image Processing*, vol. 10(1), pp. 36–48, 2001.
- [12] M.S. Drew, J. Wei, and Z.N. Li, "On illumination invariance in color object recognition," *Pattern Recognition*, vol. 31(8), pp. 1077–1087, 1998.
- [13] T. Gevers and A.W.M. Smeulders, "Content-based image retrieval by viewpoint-invariant color indexing," *Image and Vision Computing*, vol. 17, pp. 475–488, 1999.
- [14] T. Gevers and A.W.M. Smeulders, "Color-based object recognition," *Pattern Recognition*, vol. 32, pp. 453–464, 1999.
- [15] A.W.M. Smeulders and T. Gevers, "Pictoseek: combining color and shape invariant features for image retrieval," *IEEE Trans. on Image Processing*, vol. 9(1), pp. 102–119, 2000.
- [16] S. Di Zenzo, "A note on the gradient of a multi-image," *Computer Vision Graphics and Image Processing*, vol. 33(1), pp. 116–125, 1986.
- [17] J.Y. Kim, C.Y. Kim, Y.S. Seo, and I.S. Kweon, "Color indexing using chromatic invariant," *Pattern Recognition*, vol. 34, pp. 1189–1197, 2001.
- [18] N. Vandenbroucke, L. Macaire, and J.G. Postaire, "Unsupervised color texture features extraction and selection for soccer images segmentation," in *IEEE Int. Conf. on Image Processing*, Vancouver, 2000, vol. 2, pp. 800–803.
- [19] S.P. Engelson and D.V. McDermott, "Image signatures for place recognition and map construction," *SPIE Sensor Fusion IV*, pp. 282–293, 1991.
- [20] G. Sharma and H. J. Trussell, "Digital color imaging," *IEEE Trans. on Image Processing*, vol. 6(7), pp. 901–932, 1997.
- [21] G. Wyszecki and W. S. Stiles, *Color science : Concept and methods, quantitative data and formulaes*, John Wiley and Sons, New York, 1982.
- [22] R. C. Gonzalez and P. Wintz, *Digital Image Processing*, Addison Wesley, 1987.
- [23] C. Botte-Lecocq, A. Gillet, L. Macaire, and J.G. Postaire, "Color image segmentation based on fuzzy mathematical morphology," in *IEEE Int. Conf. on Image Processing*, Vancouver, 2000, vol. 3, pp. 340–343.

Identification of retinal vessels by color image analysis

V. Rakototmalala, L. Macaire, J. G. Postaire et M. Valette

Machine Graphics and Vision
volume 7, numéro 4
pages 725-743 - 1998

Identification of retinal vessels by color image analysis

V. Rakotomalala(*), L. Macaire(**), J-G. Postaire(**) and M. Valette(***)

(*) *Ecole des Mines de Douai,
Département Productique,
59508 DOUAI Cedex, France.*

(**) *Laboratoire d'Automatique I³D,
Université des Sciences et Technologies de Lille,
Cité Scientifique, Bât P2, 59655 Villeneuve d'Ascq Cedex, France.
Tel : (33)3-20-43-41-69.*

(***) *Service Universitaire des Maladies Infectieuses et du Voyageur,
Centre Hospitalier de Tourcoing,
135, rue du Président Coty, 59208 Tourcoing, France,
Tel : (33)3-20-69-44-24. e-mail : rv@cop6.univ-lille1.fr and Ludovic.Macaire@univ-lille1.fr*

Abstract

A recursive edge tracking of the retinal blood vessel is presented in order to achieve a bidimensional reconstruction of the retinal vasculature. First, color edge detection is applied on the original angiographic images. Second, from interactive selection of starting points near the papilla, the tracking of the two contours of a vessel is processed. The algorithm validates the next connected pixels to the current ones which satisfy the tracking hypotheses. Problems induced by the curvature of the vessel and the discontinuity of the detected edges will be dealt with. Third, during the tracking, the algorithm looks for branchings along the major vessel being tracked. Another recursive tracking procedure is then performed from these branch points until the end of the vessel is reached. Contour and the color of the vessel's body are associated in order to get more accurate extraction. The reconstructed vascular structures will be used as landmarks for the follow-up of a new retinopathy called CytoMegalovirus retinitis found on patients with acquired immune deficiency syndrome (AIDS).

Key words:

color image processing, color gradient, blood vessel reconstruction, edge tracking, color region/edge cooperation .

2

1 INTRODUCTION

1.1 The problem

The retinal lesion called CytoMegalovirus (CMV) retinitis appears on patients with acquired immune deficiency syndrome (AIDS)[9]. The patients may lose their visual capability because this ocular infection attacks the optic nerve, papilla, blood vessels, and the fovea, involving a progressive retinal destruction. The ophthalmologists must evaluate the evolution of the retinal lesion surface and color in order to determine the response of the CMV retinitis to an antiviral therapy.

Until now, the evolution of these lesions is evaluated by the ophthalmologists through a visual examination of serial fundus photographs where these lesions appear as textured yellow regions against a red background. As it is difficult to distinguish the fuzzy contours of these yellow regions, the ophthalmologists can't achieve accurate and reliable measures of these lesions. Such limitations call for an automatic system in order to process colorimetric and geometrical information of the retinitis lesions by color fundus images analysis.

Such a system will be designed in order to evaluate the evolution of a lesion by comparison of fundus images which are acquired at different stages of the illness. These images, which are captured by a numeric tri-CCD color camera, have a maximum size of 1520 x 1140 pixels. The position of the papilla in the fundus image, the illumination sources and the color camera calibration must be carefully adjusted in order to compare a series of fundus images of each patient, which are acquired at different stages of the evolution of the retinal lesion.

In order to measure the evolution of a lesion from the analysis of two different images, its spatial localisation in the two images must be automatically determined. A natural and widely used choice of such points are retinal vasculature geometric features. These points must have the following properties to be useful. First, they must be at fixed locations on patient's retina. Second, they must be present in sufficient number in all areas of the retina for effective location determination.

By example, points identifying bifurcations of the retinal vasculature generally meet the requirements of such landmarks.

1.2 Determination of landmarks by image analysis

Three approaches can be found in the literature in order to determine these landmarks.

1.2.1 Corners

In the first approach, Domingo [8] presents a method based on the detection of corners in an ocular fundus image in order to determine relevant points, which are used as landmarks for fluorescein angiograms. The criterion for the determination of these corners is founded on curvature measure. Then, in order to get robust identification of the landmarks, a feature vector is defined at each relevant point by means of an image intensity autocorrelation function. This vector is evaluated at different distances in a circular window centered at the point of interest.

1.2.2 Bifurcations

In the second approach, the pixels attached to the vessel bifurcations are detected by a specific image analysis method. Becker [2], in the context of a method devoted to the identification of an area of a patient's retina for a real time computer-assisted laser delivery system, proposes to detect bifurcations and crossing points by examining a local measure of an edge direction dispersion. This measure represents the variation of the gradient direction evaluated by means of the Sobel operator. This method is appropriate for a specific application which requires a real-time analysis, but it is very sensitive to the noise in the image.

1.2.3 Vessels

The third approach is based on the identification of the blood vessels, in order to determine the location of their centerlines and boundaries. This approach is more time-consuming than the second approach proposed by Becker, but the landmarks which are determined are the vessel bifurcations and the vessel boundaries. So, the localisation of retinal lesions should more be reliable when using such landmarks instead of Becker's approach.

Two kinds of vessel identification are reported in the literature: vessel identification by a segmentation and by a vessel tracking, as reported below.

Vessel identification by segmentation:

Akita [1] extracts the retinal vessels from color fundus images. He proposes a method which segments the vessels into subsegments. These subsegments are connected by a complex and not always efficient algorithm in order to reconstruct the vessels. The retinal vasculature is then represented by the so-reconstructed vessels binary tree.

Nguyen [12] detects the skeleton of coronary arteriograms thanks to a local edge detector, using density profiles of the arteries. Then, he determines the local orientations of the skeleton and determines the boundaries of the arteries from the density profiles along straight lines perpendicular to the skeleton. A sequential tracking procedure is applied to obtain segments of the arteries from

4

top-to-bottom and left-to-right scanning directions in order to obtain a structural description of the arterial tree.

Chaudhuri [4] addresses the problem of detecting blood vessels in grey levels retinal images. He proposes an operator for feature extraction based on the optical and spatial properties of the vessels. The grey levels profile of the cross section of a blood vessel is approximated by a Gaussian shaped curve. The concept of matched filter detection of signals is used to detect piecewise linear segments of blood vessels. He constructs 12 different templates that are used to search for vessel segments in a lot of directions.

Klein [10] represents the vessels as B-spline snakes in order to quantify vessel stenosis in coronary angiography. The method is based on the convolution of a bank of even and odd S-Gabor filter pairs of different orientations with the vascular images, in order to obtain an external snake energy field. Then, control points are manually selected in order to draw two B-spline snakes close to the two opposite contours of a portion of the vessel. This approach doesn't detect bifurcations and reconstruct only portions of vessels.

Vessel tracking:

Collorec [5] and Reiber [14] determine the centerline of coronary and retinal arteries respectively, by a tracking procedure which looks for the candidate points of the centerline. The centerline is composed of pixels which have the highest average brightness level computed in the corresponding search direction within a search window. Then, the two vessel boundaries are obtained by a dynamic programming method.

The method presented by Liu [11] and Sun [15], [16], is applied to a radial artery of the human hand and to coronary arteriograms. In order to determine the robustness of the algorithm, it is separately applied to two kinds of images: direct video angiograms without contrast agent and digital subtracted angiograms. Their vascular network procedure is based on a recursive tracking of the blood vessels in order to identify their centerlines and contours. The position of each centerline pixel is defined as the maximum output of the convolution of a scanline which is perpendicular to the vessel direction with a rectangular-shaped filter. In order to obtain the whole vasculature, Liu [11] looks for side-branches along both the two sides of the vessel, at a specific distance away from each left and the right sides. The vessel being tracked is then deleted from the original image to avoid interference. Once the end of the major vessel is reached, another tracking is performed from each detected initial points of side-branches, directed by a triangular-shaped filter. The problem is that there are too many parameters to be adjusted whose choice may be different from a vasculature to another.

Zhou [19] uses a similar method. In order to detect the centerline pixels, he

proposes a convolution with a gaussian-shaped filter which is more appropriate to the vessel density profile. This method is only applied to a portion of a retinal blood vessel, whose starting and ending points are interactively selected.

1.3 Our approach

We propose to determine landmarks for retinal lesion localisation by a retinal vessel identification. These vessels are reconstructed by an efficient recursive tracking.

For the analysis of CMV retinitis, ophthalmologists don't inject contrast agent into the vessels in order to enhance them. So, we propose to track these low-contrast vessels by an original vessel tracking scheme which associates both the detected vessel edges and the color body properties. The algorithm works at two step procedures. First, the vessel is reconstructed by the simultaneous tracking of its two contours. Vessel edge pixels, which form vessel contours, are detected by a color edge detection algorithm which is based on a recursive filter, known to be efficient for noisy images. This filtering procedure assures that the detected edge pixels are connected. When the edge pixels are not connected, we analyze the color body in order to fill in the edge gaps. In a second and last step, the results of the tracking scheme are validated by a measure of the vessel body color.

This specific color edge detection scheme is presented in the second section of this paper. The third section is devoted to the tracking procedure. It recursively tracks the two vessel edges, taking into account their geometry, their localisation with respect to the papilla and the measured color of the reconstructed vessel body.

As Liu and Sun [11] test their algorithm on different angiograms without contrast agent, we think that it is interesting to apply their method to color ocular fundus images. A comparative study of the vascular networks which are identified by the two approaches is presented in the last section of this paper.

2 COLOR VESSEL EDGE DETECTION

By analysis of a set of color images, Novak and Shafer [13] conclude that 10 percent of the edges extracted by a color gradient-based edge detection procedures are not identified with grey level analysis. So, we use a color gradient approach in order to detect all the low-contrast vessel edge pixels.

2.1 The algorithm of Di Zenzo

Cumani proposes an edge detection scheme which is based on second order derivative filters [6]. Such an approach is very sensitive to the presence of noise in

6

the images.

For detecting the color vessel edge pixels, we determine the color gradient magnitude and direction using the Di Zenzo algorithm [18], which computes the first-order differential operators.

The color gradient detection consists in determining, at each pixel, the direction along which the color components have the maximum rate of change. The absolute value of this maximum variation corresponds to the gradient magnitude, which is evaluated from the horizontal and vertical first derivatives of each color component.

Let denote R, G, and B the three color features of a pixel $P(x, y)$ with spatial coordinates (x,y), θ the gradient direction. The variation function is:

$$F(\theta) = p \cos^2 \theta + q \sin^2 \theta + 2t \sin \theta \cos \theta \quad (1)$$

with:

$$p = \left(\frac{\partial R(P(x, y))}{\partial x} \right)^2 + \left(\frac{\partial G(P(x, y))}{\partial x} \right)^2 + \left(\frac{\partial B(P(x, y))}{\partial x} \right)^2 \quad (2)$$

$$q = \left(\frac{\partial R(P(x, y))}{\partial y} \right)^2 + \left(\frac{\partial G(P(x, y))}{\partial y} \right)^2 + \left(\frac{\partial B(P(x, y))}{\partial y} \right)^2 \quad (3)$$

and

$$t = \frac{\partial R(P(x, y))}{\partial x} \frac{\partial R(P(x, y))}{\partial y} + \frac{\partial G(P(x, y))}{\partial x} \frac{\partial G(P(x, y))}{\partial y} + \frac{\partial B(P(x, y))}{\partial x} \frac{\partial B(P(x, y))}{\partial y} \quad (4)$$

The direction of the color gradient which maximises the function $F(\theta)$ is given by:

$$\theta_{max} = \frac{1}{2} \arctan \frac{2t}{p - q} \quad (5)$$

and the gradient magnitude is therefore equal to $\sqrt{F(\theta_{max})}$.

This maximum variation is evaluated from the first horizontal and vertical derivatives of each color feature.

For this purpose, we apply the recursive Deriche filter as a differential operator [7]. This optimal filter is known to be well-suited for edge detection in noisy images [17]. So, such a filter seems to be appropriate for vessel edge pixels detection from fundus images which are noisy and non-uniformly illuminated.

7

The behaviour of this filter is governed by a parameter α . It is adjusted according to the expected filter performance which is evaluated in terms of detection and localisation. In order to extract well-connected edge pixels, detection of all colorimetric spatial variation is more important than the accuracy of their localisation. The parameter α is then adjusted so that the trade-off between detection and localisation favours detection.

In order to extract thin edges from the so-obtained gradient image, we retain the edge pixels whose magnitude are local maxima along the gradient direction. Then, an hysteresis thresholding scheme, with a low threshold Th_l and a high threshold Th_h , provides a binary edge image. These parameters are adjusted so that a maximum of vessel edge pixels are detected. However, with such a strategy, many false edge pixels are extracted (see Figure 1). The blood vessel reconstruction algorithm (cf. section 3), allows to discriminate true vessel edge pixels from those appearing in the background of the image.

The results of Figure 1 have been obtained with the (L,a,b) color representation

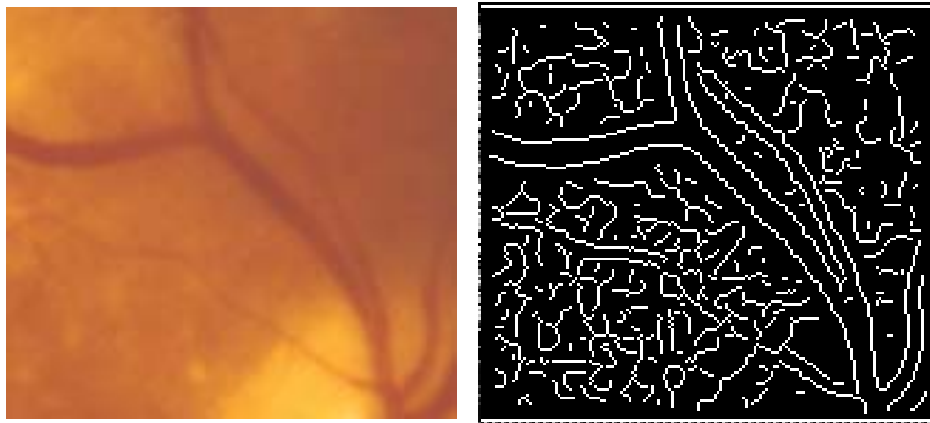


Figure 1: (a)Original image, (b)Thin detected edges

system. Since, The choice of the color system affects the performance of the vessel edge detection, it is necessary to evaluate the behaviour of the edge detection algorithm with well known systems, i.e. the (R,G,B) system, the Intensity Hue Saturation (I,H,S) system, the (I1,I2,I3) system and the uniform CIE (L,a,b) system [3]. From a statistical evaluation of the resulting binary edge image, which is not detailed in this paper, we conclude that the CIE (L,a,b) system gives the best results for vessel edge detection.

3 BLOOD VESSEL RECONSTRUCTION

A blood vessel is reconstructed by a recursive simultaneous tracking of its two contours. Our approach takes into account the following hypotheses:

- continuity properties of a blood vessel, i.e., continuity of edge pixel positions, of the curvature, diameter and cross-sectional density profile of the vessel,
- the two vessel contours are assured to be locally parallel,
- the blood vessel leads from the papilla to the external border of the fundus image.

Once the preprocessing of the color fundus image has yield the binary edge image, the color edge tracking process is divided into five steps, as presented in the following sections.

3.1 Edge tracking of a single vessel

We denote R_c (respectively L_c) the right (respectively the left) current edge pixel belonging to one of the two contours of a vessel (see Figure 2). Then,

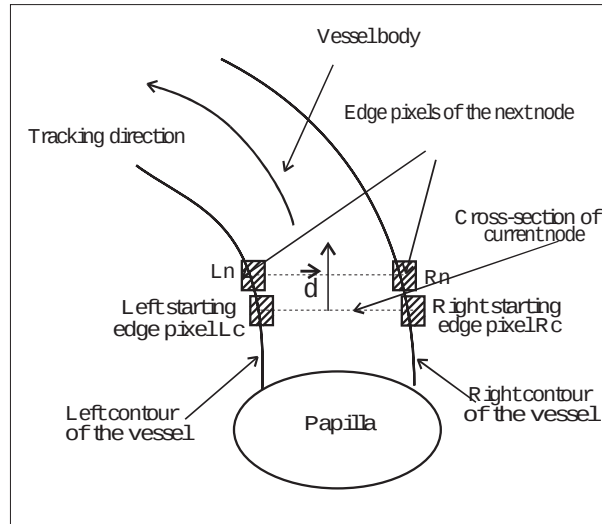


Figure 2: Vessel structure

we define the current node as the pair of pixels $R_c - L_c$. The next node is composed of the pair of the next right and left edge pixels denoted, R_n and L_n respectively. The right (or left) pixel of the next node must be connected to the

right (or left) pixel of the current node.

Finally, we call pseudo-diameter of the cross-section of the vessel, the length of the straight segment joining the right and the left pixels of a node. This segment will be called, hereafter, the cross-section of the node. We denote D_c the pseudo-diameter of the current cross-section, and D_n that of the next one.

In a first step of a vessel reconstruction, two starting edge pixels (right and left) near the papilla are interactively selected by an operator. They constitute the first current node. We evaluate the color properties of the body vessel along the current node cross-section. We also evaluate the starting direction, which leads from the papilla to the external border of the fundus, and the pseudo-diameter of the vessel at this current node.

By following the tracking direction, the algorithm looks for the next pixels R_n and L_n , which satisfy the connectivity properties with the current pixels and the other tracking hypotheses. When they are found, they form the next current node. Once the next node is validated, the current node is marked in the binary edge image to avoid interference during the subsequent tracking process.

Otherwise, the current local vessel direction, denoted $\vec{d}$ (see Figure 3), is computed along the tracking procedure.

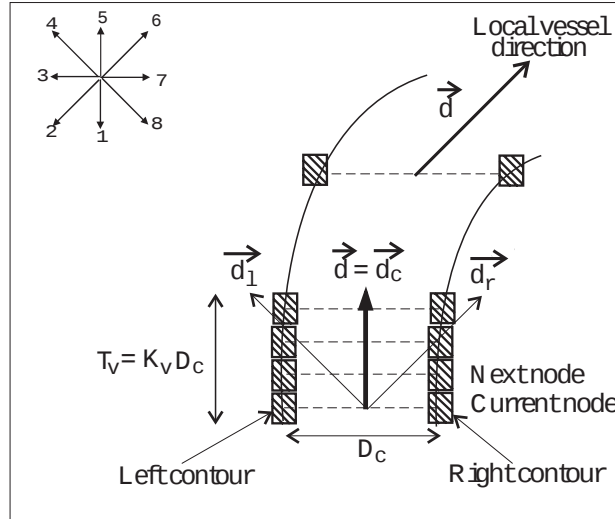


Figure 3: Determination of the local vessel direction

For this determination, we use 3 vectors, denoted $\vec{d}_c$, $\vec{d}_l$, and $\vec{d}_r$, with the following features:

10

- their origin is the center of the current cross-section,
- their direction is one among 8 possible directions, with a step of $\pi/4$,
- their length, denoted T_v , is proportional to the current pseudo-diameter, denoted D_c , as $T_v = K_v \cdot D_c$, with $K_v \geq 1$.

Then, we evaluate the mean of the red component of the pixels along each of these 3 vectors. By our experiments, we conclude that the red component seems to be a robust color attribute in order to characterise the vessel body. The current direction $\vec{d}$ is determined as the direction for which the red mean value is maximum [5]. For example, on Figure 3, $\vec{d}$ is determined according to the direction 5 for the first case and to the direction 6 for the second case.

During the tracking process, several candidate pixels connected to the current one may appear. In order to choose among these k candidates, we compute the scalar products:

$$Prod_i = \overrightarrow{L_c L_n^i} \cdot \vec{d}, \quad i = 1..k \quad (6)$$

where L_c is the left current pixel, L_n^i is the i^{th} connected pixel to L_c , and $\vec{d}$ is the local vessel direction determined as above. Then, the left next pixel is the i^{th} pixel connected to L_c which has the maximum value of $Prod_i$. The same method works for the right current pixel. This procedure has proved suitable to determine the correct vessel direction during the tracking.

3.2 Colorimetric validation

One of the tracking hypotheses relies on the continuity of each cross-sectional color properties of the vessel body.

In our approach, these color properties are identified as the mean value of the red component computed along each node cross-section.

Let define Red_c the current mean value of the red component along the current section, and Red_n the next one:

$$Red_c = \frac{1}{D_c} \sum_{P_i \in [L_c, R_c]} R(P_i(x, y)), \quad Red_n = \frac{1}{D_n} \sum_{P_i \in [L_n, R_n]} R(P_i(x, y)) \quad (7)$$

where D_c and D_n are the current and the next pseudo-diameters, respectively, and $R(P_i(x, y))$ is the value of the red component at $P_i(x, y)$.

The variation of this measure between the current and the next sections, defined as:

$$variation = \frac{Red_n - Red_c}{Red_c}, \quad (8)$$

must not exceed a threshold, denoted K_t . A next node is selected if its color is similar to that of the current one. This criterion is checked along the color edge tracking processus in order to detect the end of a vessel, or to avoid that false vessel edges are taken into account during the reconstruction.

Such a validation procedure ensures that the region which is defined by the successive cross-sections represents a true vessel body.

3.3 Analysis of vessel curvature

In order to ensure a correct colorimetric validation, the cross-section must be perpendicular to the vessel direction. So, the tracking process must take into account the curvature of the vessel.

Let us suppose that we have selected by means of connectivity analysis the right next edge pixel denoted R_n just before a curve turning right (see Figure 4).

A right turn is found when:

$$\text{dist}(R_n, L_c) < \text{dist}(R_c, L_n)$$

where $\text{dist}(\cdot)$ is the Euclidian distance between two pixels.

We are therefore looking for the nearest left edge pixel L_n for which the cross-section of the next node is perpendicular to the vessel direction. To achieve this goal, we clamp the right next edge pixel R_n and we look for the left next edge pixel along the left contour. The right contour tracking is then stopped but the left one is still running to find left next candidates.

Then, we consider that the cross-section is perpendicular to the vessel direction when:

$$\text{dist}(R_c, L_c) < \text{dist}(R_c, L_n^j)$$

where L_n^j is the j^{th} left candidate edge pixel starting from L_c that satisfies this condition.

The pixel L_n^j is then the next left pixel: $L_n = L_n^j$. For example, in Figure 4, L_n^4 is chosen as the next left edge pixel.

A similar procedure also works for a left turn.

3.4 Detection of vessel bifurcation points

During the tracking of a major vessel, the algorithm looks for possible bifurcations along each current cross-section.

12

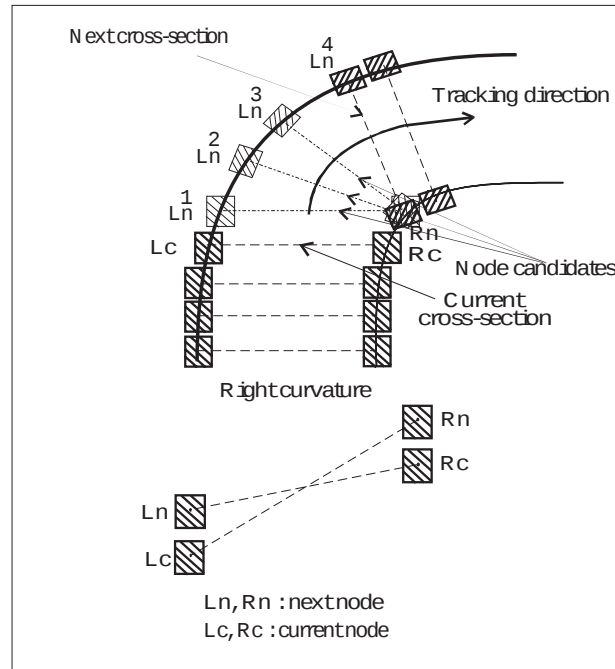


Figure 4: Tracking in a right curve

We propose to look for a pair of detected edge pixels (L_b, R_b) on each cross-section delimited by R_c and L_c which satisfies the following conditions: (see Figure 5)

- the distance between L_b and R_b must be larger than one pixel in order to ignore noise edge pixels inside the vessel body.
- the mean value of the red component along the cross-sections between (L_c, R_b) and (R_c, L_b) must be similar to the one along the previous cross-section so that the cross-sectional color properties continuity hypothesis is satisfied.

Then, the same recursive vessel tracking procedure is initiated from the new couples of starting points, namely (L_c, R_b) and (R_c, L_b), once these two pairs of points are validated.

At the end of the tracking process, the vessels are described in terms of linked successive nodes.

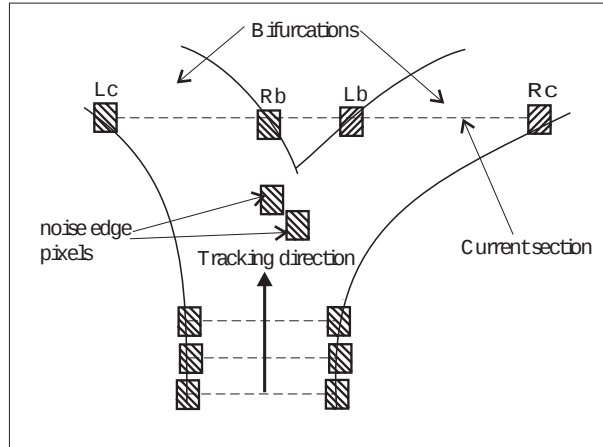


Figure 5: Detection of a bifurcation

3.5 Edge gap filling

The above described tracking procedure may fail if it is not possible to find a candidate edge pixel which satisfies the required connectivity conditions. In order to cope such situations, which correspond to contour gaps, we propose to fill in these gaps, according to the estimated current vessel direction $\vec{d}$ (see Figure 6), whose determination has been presented in section 3.1.

Two examples of vessel contour gaps are presented in Figure 6. The first edge gap to be filled is on the left vessel contour. Since the tracking scheme does not find a connected left next pixel, this gap is filled according to the local vessel direction. So, we consider that the pixel, which is connected to the left current pixel according to the direction 5, is the next left pixel. A similar approach is used for filling the second edge gap according to direction 6.

With these examples, we demonstrate that this extrapolation scheme, which associates color and edge information, seems to be efficient even when a vessel presents an high curvature. The size of such filled gaps can reach many pixels.

4 RESULTS

In order to illustrate the behaviour of our scheme, we present a first color fundus image (see Figure 7). This image contains the papilla, many retinal vessels and some C.M.V. retinitis lesions.

The edges are detected with the method presented in section 2 (see Figure 8). From this so-obtained edge image, we interactively select the position of the

14

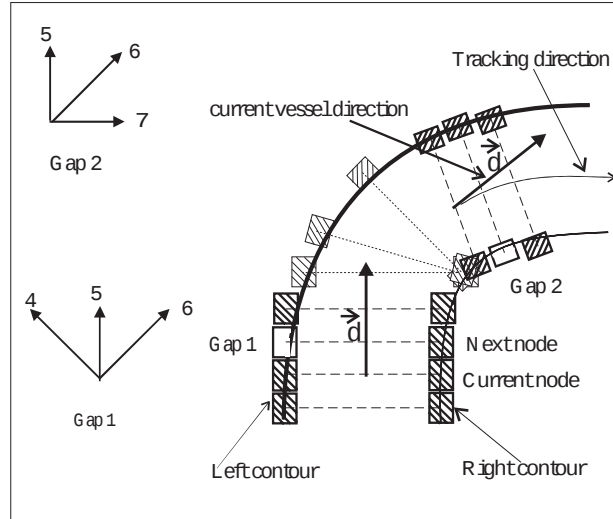


Figure 6: Edge gap filling

papilla and two vessel edge pixels, denoted left and right edge pixels, near the papilla in order to define the initial pseudo-diameter and initial direction of the tracking process.

Edge vessel discontinuities, denoted Gaps(a) and Gaps(b) are shown in Figure 8, where they are surrounded by circles and enlarged on Figure 9 and Figure 11 respectively. The figures 10 and 12 show the reconstructed vessels by their white contours. By the examination of these figures, we remark that:

- the edge gaps have been filled in correctly,
- the bifurcation of vessels are well identified.

It is interesting to compare our approach to an existing one for vessel edge identification. For this purpose, we have implemented the algorithm of Liu and Sun which is reported in [11]. We have chosen this algorithm because it seems to be well adapted to the analysis of different kinds of angiogram images. A comparative study of the vascular networks which are identified by the two approaches is presented in the last section of this paper. Their approach detects the centerline of blood vessels with matched filters whose characteristics depend on eight parameters.

We have applied the Liu and Sun's algorithm and our approach to two successive entire fundus images of the same retina, acquired at an interval of 8 days from each other after an antiviral therapy.

The Liu's algorithm needs a set of eight parameters, which is detailed in [11]. It is performed on the red colorimetric feature of the color image, with the eight following parameter values:

$$\begin{aligned} N_{ahead} = W \quad W_{main} = 2W \quad T_{sbr} = -40db \quad T_{ML} = 10, \\ W_{side} = 2W \quad N_{skip} = 2W \quad N_{off} = 3 \quad N_{width} = 6. \end{aligned}$$

The results are very sensitive to the choice of these parameters, so it's difficult to obtain by experiments a satisfying set of these parameters' values. We present in this paper, the values which seem to be the best adapted ones to our problem. (see Figure 13 and 14).

Note that, in order to define the initial direction of the Liu's tracking procedure, we have interactively selected two consecutive points near the papilla, for each vessel tracking.

Our approach is tested on the same images, with the following parameters' values (see Figure 15 and 16):

- for the color edge detection: $\alpha = 0.8$, $Th_l = 1$, $Th_h = 2$, where α is the Deriche filter parameter, Th_l and Th_h are respectively the low and high threshold values of the hysteresis filter,
- for the tracking procedure : $K_v = 1.5$, $K_t = 0.35$, where K_v is the proportional constant of the look-ahead distance for the determination of the local vessel direction $\vec{d}$, and K_t is the tolerance of the variation of the red component mean value between two consecutive nodes.

For our approach, the centerline position of the blood vessel is defined as the middle of a cross-section.

The results of the two approaches can be seen on Figure 13 and 14 for the Liu and Sun's method, and on Figure 15 and 16 for our approach, where centerlines of the reconstructed vessels are superimposed on the original color fundus image in blue color, while the edges are displayed in white.

By the analysis of these images, we have pointed the following features of our approach compared to Liu's one:

- For the vessel tracking procedure, our approach seems to be more efficient than Liu and Sun's one. Indeed, almost the whole major vessel being tracked is reconstructed (see Figure 15 and 16), while the Liu's tracking process ends up just after the identification of a little portion of the vessel body (see Figure 13 and 14). The robustness of our method compared to Liu and Sun's one is due to the good performance of Deriche filter for noisy images. Indeed, the ocular fundus images are noisy and non-uniformly illuminated, so region segmentation used by Liu and Sun is very sensitive to these artefacts.

16

- In order to compare the quality of the bifurcation detection, all starting pixels which are detected by our approach are overlaid in green in Figure 15 and Figure 16, for our approach. Detected bifurcation points by Liu and Sun's algorithm are displayed in green in Figure 13 and Figure 14. It can be observed that our method is more accurate than Liu and Sun's one for bifurcation detection. Indeed, the detected bifurcations by Liu and Sun's algorithm are almost false bifurcations, whereas the error due to our approach for bifurcation points detection is very low.

5 CONCLUSION

We have presented an automated identification of retinal blood vessels by means of a recursive color edge tracking scheme. We have proposed to use simultaneously the detected vessel edges and color measures of the vessel body in order to achieve a reliable vessel tracking procedure. Our method extracts and accurately detects the bifurcations of the vessels. These vessels will be used for automatic image registration in the next step of this work in order to compare different images of the same retina acquired at different times, in order to evaluate the evolution of the CMV retinitis lesions. This bidimensional retinal vessel reconstruction is being presently validated by ophthalmologists.

References

- [1] K. Akita and H. Kuga. Digital processing of color ocular fundus images. In *MEDINFO 80*, pages 363–372. North-Holland Publishing Company, 1980. Lindberg/Kaihara.
- [2] D. E. Becker, A. Can, J. Turner, H. Tanenbaum, and B. Roysam. Image processing algorithms for retinal montage synthesis, mapping, and real-time location determination. *IEEE Transactions on Biomedical engineering*, Vol. 45, n° 1:105–118, Jan. 1998.
- [3] M. Celenk and S. Smith. Color image segmentation by clustering and parametric-histogramming technique. pages 883–886. IEEE 8th International conference on Pattern Recognition, 1986. Paris.
- [4] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging*, Vol.8, n° 3:263–269, Sept. 1989.
- [5] R. Collorec and J.L. Coatrieux. Vectorial tracking and directed contour finder for vascular network in digital subtraction angiography. *Pattern Recognition Letters*, (Vol. 8, n° 5):353–358, 1988.
- [6] A. Cumani. Edge detection in multispectral images. *Comput. Vision Graphics Image Processing*, Vol. 53, n° 1:40–51, Jan. 1991.

- [7] R. Deriche. Fast algorithms for low level vision. *IEEE Transactions on Pattern Analysis Machine Intelligence*, Vol. 12, n° 1:78–87, Jan. 1990.
- [8] J. Domingo, G. Ayala, and P. Marco. Irregular motion recovery in fluorescein angiograms. *Pattern Recognition Letters*, Vol. 18:805–821, 1997.
- [9] M. Flores-Aguilar, Munguia, D. Bessen, P. Gangan, J. F. Arevalo, and W. Freeman. Clinical versus fundus photographic evaluation of the status of cytomegalovirus retinitis in aids patients. *Retina, the Journal of retinal and vitreous diseases*, Vol. 16, n° 5:356–372, Nov. 1996.
- [10] A. K. Klein, F. Lee, and A. A. Amini. Quantitative coronary angiography with deformable spline models. *IEEE Transactions on Medical Imaging*, Vol. 16, n° 5:468–482, Oct. 1997.
- [11] I. Liu and Y. Sun. Recursive tracking of vascular networks in angiograms based on the detection-deletion scheme. *IEEE Transactions on Medical Imaging*, Vol. 12, n° 2:334–341, Jun. 1993.
- [12] T. V. Nguyen and J. Sklansky. Computing the skeleton of coronary arteries in cineangiograms. *Computers and Biomedical research*, Vol. 19:428–444, 1986.
- [13] C.L. Novak and S.A. Shafer. Color edge detection. In *DARPA Image Understanding Workshop*, volume Vol. 1, pages 35–37, Feb. 1987. Los Angeles, CA, U.S.A.
- [14] J.H.C. Reiber, P.W. Serryus, and C. J. Slager. *Structural analysis of the coronary and retinal arterial tree*, volume Vol. 51. Pub. Martinus, 1986. Nijhorf, Dordrecht.
- [15] Y. Sun. Automated identification of vessels contours in coronary arteriograms by an adaptive tracking algorithm. *IEEE Transactions on Medical Imaging*, Vol. 8, n° 1:78–88, Mar. 1989.
- [16] Y. Sun, R. J. Lucariello, and S. A. Chiaramida. Directional low-pass filtering for improved accuracy and reproducibility of stenosis quantification in coronary arteriograms. *IEEE Transactions on Medical Imaging*, Vol. 1, n° 2:242–248, June 1995.
- [17] Y.F. Wan, J.P. Postaire, and F. Cabestaing. A new filter for edge detection. In *ICNNSP 95*, pages 443–448, 1995. Tanjin, Chine.
- [18] S. Di Zenzo. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, (n° 33):116–125, 1986.
- [19] L. Zhou, Rzerzotarski, L. Singerman, and J. Chokreff. The detection and quantification of retinopathy using digital angiograms. *IEEE Transactions on Medical Imaging*, Vol.13, n° 4:619–626, Dec. 1994.

18

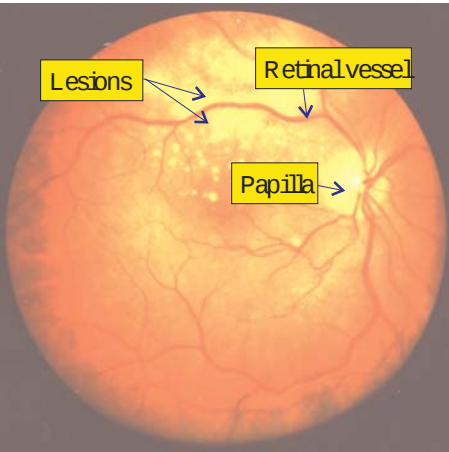


Figure 7: A color ocular fundus image

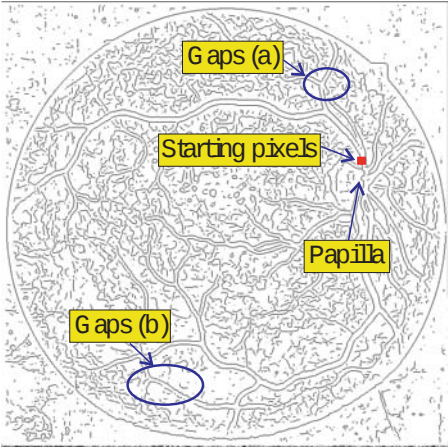


Figure 8: Edge image corresponding to Figure 7: The tracking process begins from the starting pixels near the papilla.

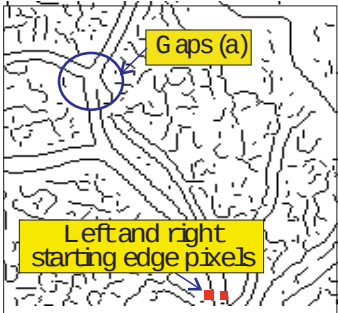


Figure 9: Gaps(a): Detected edges with gaps

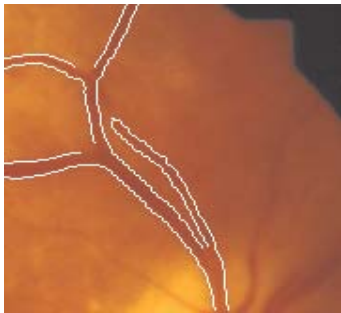


Figure 10: Gaps(a) Reconstructed blood vessels

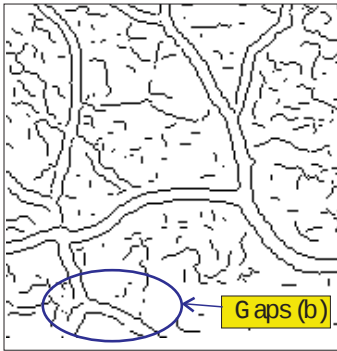


Figure 11: Gaps(b): Detected edges with gaps



Figure 12: Gaps(b): Reconstructed blood vessels

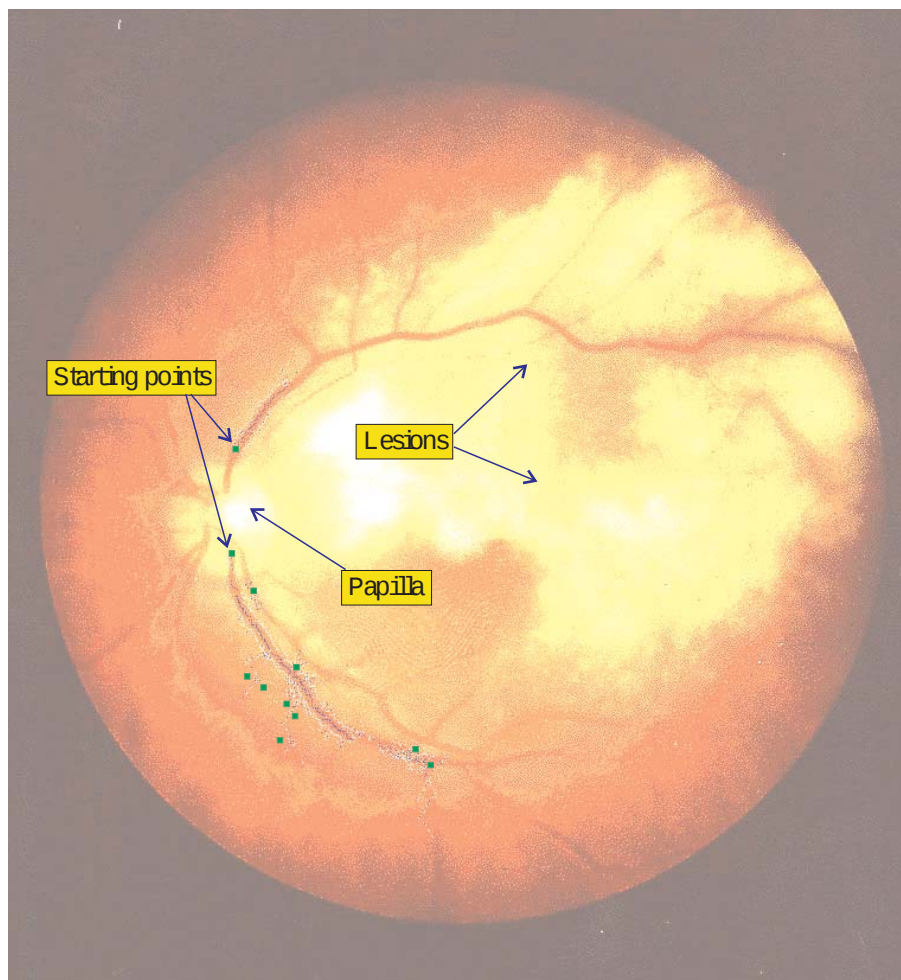


Figure 13: The result of Liu and Sun's method.

20

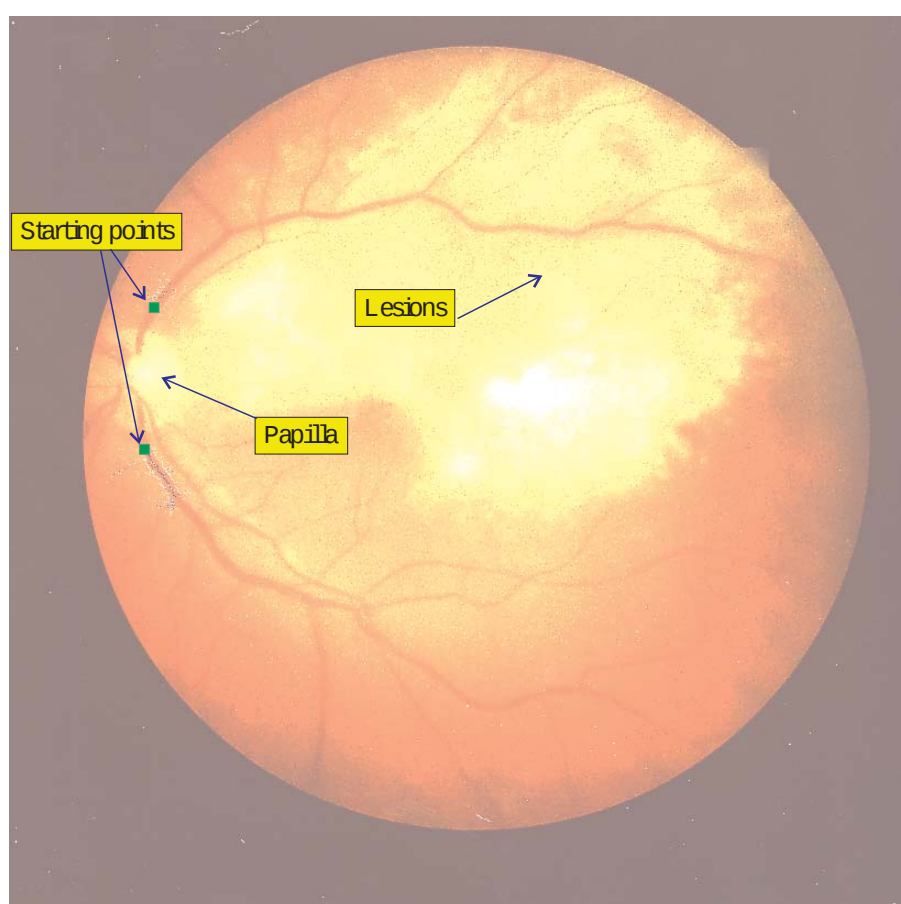


Figure 14: The result of Liu and Sun's method on the same retina, acquired after eight days of antiviral therapy.

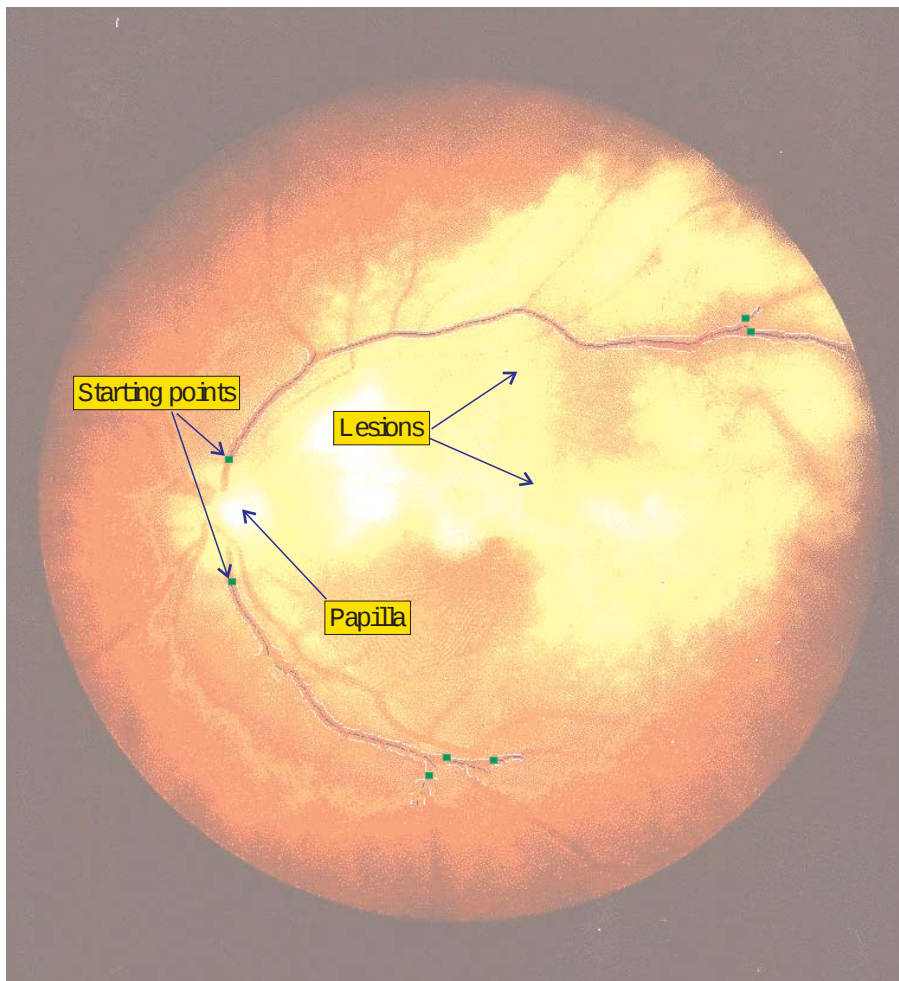


Figure 15: The result of our approach.

22

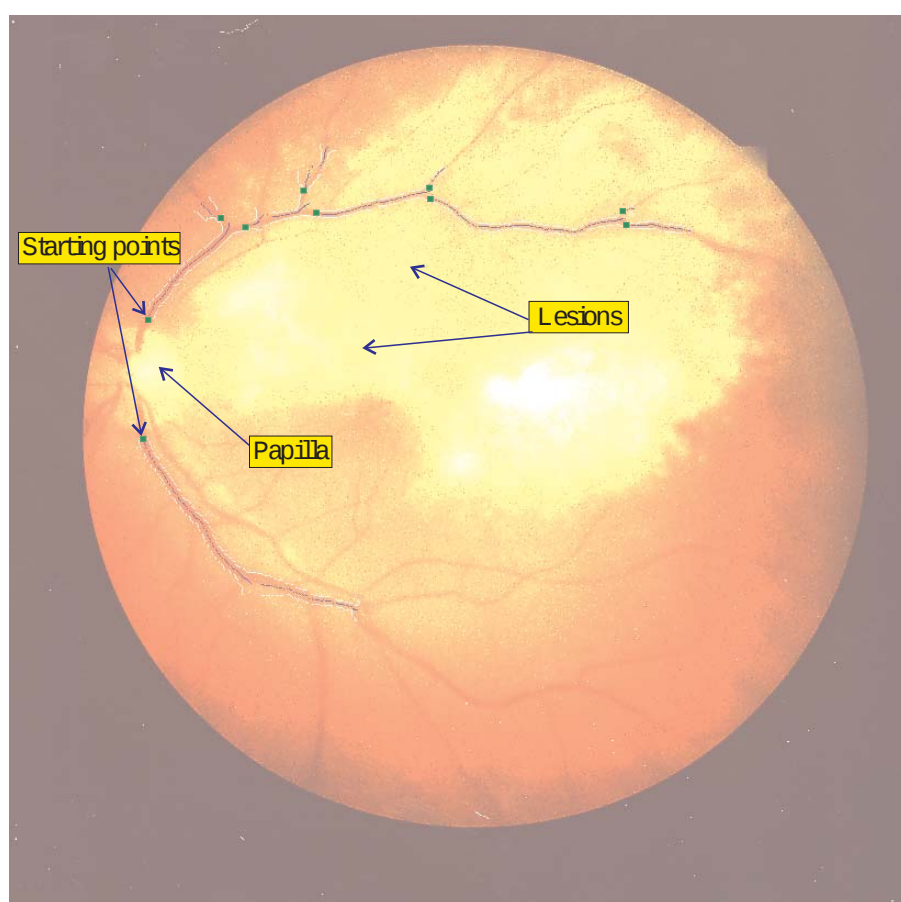


Figure 16: The result of our approach on the same retina, acquired after eight days of antiviral therapy.

Determination of compatibility coefficients for color edge detection by relaxation

V. Ultré, L. Macaire et J. G. Postaire

Proc. of the IEEE Int. Conf. on Image Processing (ICIP 96)
volume 3 - Lausanne (Suisse)
pages 1045-1048 - 1996

DETERMINATION OF COMPATIBILITY COEFFICIENTS FOR COLOR EDGE DETECTION BY RELAXATION

Ludovic Macaire, Vincent Ultré and Jack-Gerard Postaire

Centre d'Automatique de Lille
 Université des Sciences et Technologies de Lille
 Bâtiment P2- Cité Scientifique
 59655 Villeneuve d'Ascq cedex - FRANCE
 email: Ludovic.Macaire@univ-lille1.fr

ABSTRACT

In this paper, we propose a new algorithm which detects boundaries of objects from a color image. The result of this method is a binary image where the boundary pixels are only represented. This approach uses a co-operating relaxation process in order to enhance edge probabilities. The relaxation algorithm processes probabilities resulted from a gradient output of the different features of the color images. So, this algorithm is able to detect edges from variations of red, green and blue images.

We present in the first part of this paper the relaxation scheme. In the second part, we propose two kinds of compatibility coefficients evaluations. These coefficients are compared in the third part of this paper using a rather complex image.

keywords: color images, edge detection, relaxation

1. INTRODUCTION

Contour detection is a major issue in image processing. In fact, the information contained in contours is very useful and often indispensable for image segmentation. In color images, the notion of contour is considerably richer than in grey level ones. In a grey level image, only a difference of lightness between adjacent homogeneous regions is considered as a contour. Therefore, the contour between two equally bright neighbouring objects of different hue cannot be detected. To overcome this problem, the chrominance information present in color images must be taken into account. A number of algorithms have been developed to compute the gradient in color images. The available techniques for edge detection are, in fact, adaptations of the following three paradigms.

The first method, defined by Claxton and Kwok [1] makes use of a particular metric in the color space to

define the notion of colorimetric distance. This method reduces the problem to grey level contour detection. For this reason, Luong wonders if any improvement is made with respect to grey level methods [2]. The second method computes the gradient on each of the three image components, and then combines the results, as the sum of their absolute values or as their maximum absolute values [3]. The third approach proposed by Di Zenzo combines the three components to compute the gradient. The tensor-based method computes the gradient as a vector. The direction of the gradient in the image is that in which the module is the highest one [4]. The method of Lee-Cok also uses the first derivatives of image components, seeking the eigenvalues and eigenvectors of a matrix defined in terms of those first derivatives. The gradient module is the square root of the highest eigenvalue [5]. Another approach is based on the second directional derivative [6].

Once the colour gradient is obtained, the image has to be binarised to extract the contours. This involves the definition of a suitable global threshold in such a way as to ensure contour connectivity while limiting the number of false points. It is well known that, it is often difficult to determine automatically a global threshold for the binarization of the gradient image.

This problem leads us to develop a local thresholding algorithm using a co-operating relaxation process to enhance edge probabilities [7]. The labeling relaxation algorithm processes these probabilities which are evaluated by computing the gradient of the color image. So, this algorithm is able to detect edges from a variation of intensity, saturation or hue.

We present in the first part of this paper the relaxation scheme. In the second part, we propose two kinds of compatibility coefficients evaluations. These coefficients are compared in the third part of this paper using a rather complex image.

2. THE LABELING RELAXATION ALGORITHM

The relaxation algorithm takes into account four classes of pixels. Three classes represent a gradient of the three colour image features R,G,B. The classes of the red, green and blue edges are denoted as λ_R , λ_G and λ_B respectively. At each pixel and for each color feature, the gradient is evaluated using the well known Sobel filter. The higher the feature response of this filter, the higher is the probability of the pixel to belong to the corresponding edge class. The last class, denoted λ_{NE} or no-edge class, represents the no edge pixels. The probabilities to belong to these four classes are computed from the three former probabilities [8]. Since, for each pixel, the sum of the probabilities of the four classes must be equal to 1.

Let $P_{(x,y)}^k(\lambda_n)$ the probability that the pixel (x,y) belongs to the class λ_n at the iteration k of the relaxation process. The relaxation algorithm works on the pixels of the three R,G,B images and on a set of four classes $\lambda_R, \lambda_G, \lambda_B, \lambda_{NE}$. It iteratively computes the probability that a pixel (x,y) belongs to a class, by taking into account its compatibility with its neighbours.

We initially assign to each pixel (x,y) the probability $P_{(x,y)}^0(\lambda_n)$ that it belongs to a class λ_n , $n=R,G,B,NE$. Then we examine the probabilities at neighbouring pixels, increasing $P_{(x,y)}^k(\lambda_n)$ if supporting evidence is found for it (i.e., if there exist high probabilities that the neighbours of (x,y) have compatible classifications with the assignment of this point to the class λ_n), or decrease it if contradictory evidence is found. The compatibility that the pixel (x,y) belongs to the class λ_n and a neighbour (x+i,y+j) belongs to the class λ_m , which is denoted $r_{ij}(\lambda_n, \lambda_m)$, is processed with the compatibility coefficients and the probabilities of these neighbours. The influence of the neighbouring on the probability $P_{(x,y)}^k(\lambda_n)$ is called the contribution and is denoted $q_{(x,y)}^k(\lambda_n)$. It is computed following equation (1). We consider the eight neighbours of the pixel (x,y).

$$q_{(x,y)}^k(\lambda_n) = \sum_{i=-1}^1 \sum_{j=-1}^1 \sum_{\alpha=R,G,B,NE} r_{ij}(\lambda_n, \alpha) P_{(x,y)}^k(\alpha) \quad (1)$$

The probability at the $k+1^{th}$ iteration is given by the following equation:

$$P_{(x,y)}^{k+1}(\lambda_n) = \frac{P_{(x,y)}^k(\lambda_n) \left((1 + q_{(x,y)}^k(\lambda_n)) \right)}{\sum_{\alpha=R,G,B,NE} P_{(x,y)}^k(\alpha) \left((1 + q_{(x,y)}^k(\alpha)) \right)} \quad (2)$$

All the probabilities $P_{(x,y)}^k(\lambda_n)$ are updated in par-

allel on each pixel (x,y) and for each class. The process is iterated until the probability of each pixel to belong to a class remains constant. At the end of the process, the pixels probabilities to belong to the no-edge class tend to 0 or to 1, so it is very easy to threshold the image, by examining the probabilities of the pixels to belong to this class. We consider that a pixel with a low probability to belong to no edge class represents an edge pixel.

3. THE COMPATIBILITY COEFFICIENTS

Let $r_{ij}(\lambda_n, \lambda_m)$ be the compatibility coefficient that a pixel (x,y) belongs to the class λ_n and that a neighbour (x+i,y+j) belongs to the class λ_m . The function $r_{ij}(\lambda_n, \lambda_m)$ must be bounded by -1 and 1. If $r_{ij}(\lambda_n, \lambda_m)$ goes to 1, the assignement of the pixel (x,y) to the class λ_n is compatible with the assignement of the pixel (x+i,y+j) to λ_m . Otherwise, if $r_{ij}(\lambda_n, \lambda_m)$ tends to -1, the assignement is incompatible. The efficiency of the relaxation algorithm depends on the choice of the compatibility coefficients $r_{ij}(\lambda_n, \lambda_m)$. Experimental determination of these coefficients seems to give good results for edge detection in grey level images. In our case, it is too difficult to experimentally adjust the compatibility coefficients for 4 classes [9].

Two methods are proposed. First, we consider the correlation between two classes. Secondly, we evaluate the mutual information between two classes [10].

3.1. The correlation between two classes

The correlation seems to be the most appealing method. Indeed, if two classes are highly correlated, we can say that the assignement to these two classes is compatible. The estimation of $r_{ij}(\lambda_n, \lambda_m)$ from the correlation of two classes λ_n and λ_m is given by the following equation:

$$A(x, y) = (P_{(x,y)}^0(\lambda_n) - \bar{P}^0(\lambda_n)) \cdot (P_{(x,y)}^0(\lambda_m) - \bar{P}^0(\lambda_m))$$

$$r_{ij}(\lambda_n, \lambda_m) = N^2 \cdot \frac{\sum_{(x,y)} A(x, y)}{\sigma(\lambda_n) \cdot \sigma(\lambda_m)} \quad (3)$$

N represents the number of pixels in the image, $\bar{P}^0(\lambda_n)$ and $\sigma(\lambda_n)$ the mean and the variance of the initial probability of the pixels to belong to the class λ_n , respectively.

3.2. The mutual information between two classes

PELEG proposed to compute these coefficients with the initial probabilities of the pixels to belong to the

four classes, by the evaluation of the neighbouring mutual information between two classes [10]. The compatibility coefficients definition is based on the mutual information of the classes to which are assigned to neighbouring pixels. So $r_{ij}(\lambda_n, \lambda_m)$ is computed by the equation (4), where N is the number of pixels of in the image.

$$A(x, y) = P_{(x,y)}^0(\lambda_n)$$

$$B(x, y) = P_{(x+i,y+j)}^0(\lambda_m)$$

$$C(x, y) = \sum_{(x,y)} (P_{(x,y)}^0(\lambda_n))$$

$$D(x, y) = \sum_{(x,y)} P_{(x+i,y+j)}^0(\lambda_m)$$

$$r_{ij}(\lambda_n, \lambda_m) = \log \frac{N \cdot \sum_{(x,y)} (A(x, y) \cdot B(x, y))}{C(x, y) \cdot D(x, y)} \quad (4)$$

This method satisfies our idea about the nature of the coefficients. If two neighbouring pixels are assigned to two classes having a high positive correlation, we can expect them to have an high mutual information, and vice versa. This mutual information is easily computed; it is the log of the ratio of conditional probabilities to unconditional probabilities.

4. RESULTS

It is very interesting to compare theses two approches for compatibility coefficients evaluation. We choose a complex image of a morrocan mosaic, which is difficult to segment because it contains a lot of objects (see fig. 1). We compute the relaxation algorithm which works on the four classes $\lambda_R, \lambda_G, \lambda_B, \lambda_{NE}$. The probabilities whose values are between 0 and 1, are coded with one byte, so that their values vary between 0 and 255. After four iterations, the probabilities tend to 0 or 1. We threshold the probabilities for the pixels to belong to the no-edge class, in order to extract the boundaries.

Figure 2 contains the results of the relaxation using the compatibility coefficients derived from the correlation between two classes.

Figure 3 shows the results using the mutual information.

In both cases, the threshold has been adjusted at 0,5, i.e. in the middle of the range of the probabilities.

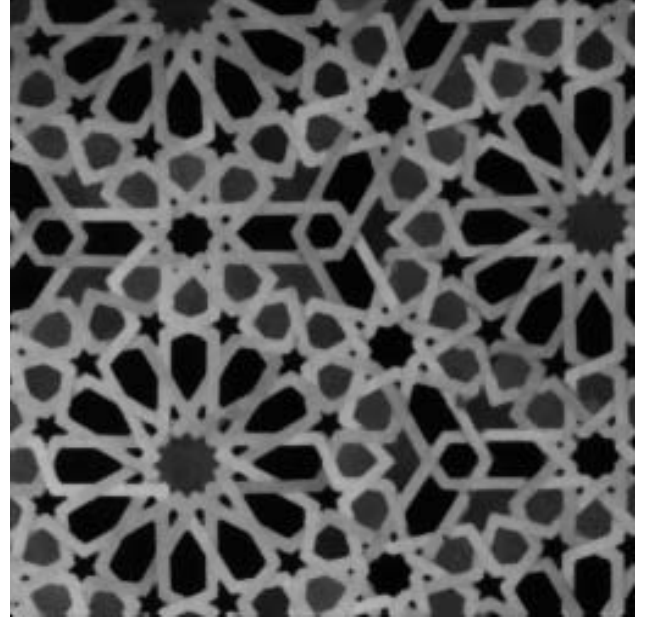


Figure 1: color image

These edge images show that the relaxation process using compatibility coefficients derived from the correlation yield the edges which are sometimes broken, while the edges resulting from the relaxation using mutual information are well closed, but are sometimes merged together.

5. CONCLUSION

Our approach differs from classical colour contour detection methods in the following aspect. The classical methods result in a grey level contour image from which it is generally difficult to extract the edges. Our method yields a quasi binary edge image extracted from a colour image. To achieve this goal, we used a relaxation-based algorithm. Our method is very interesting, because it takes into account simultaneously the three color features in order to detect boundaries. So a grey level variation in each feature can give rise a boundary. Beside the intensity variation, our method reveals contours resulting from variation in color characteristics, such as hue and saturation.

6. REFERENCES

- [1] P.R Claxton and E.Kwok, "The use of colour to segment and label images", in *Alwey Vision Conf.*, 1987, pp. 295–302.

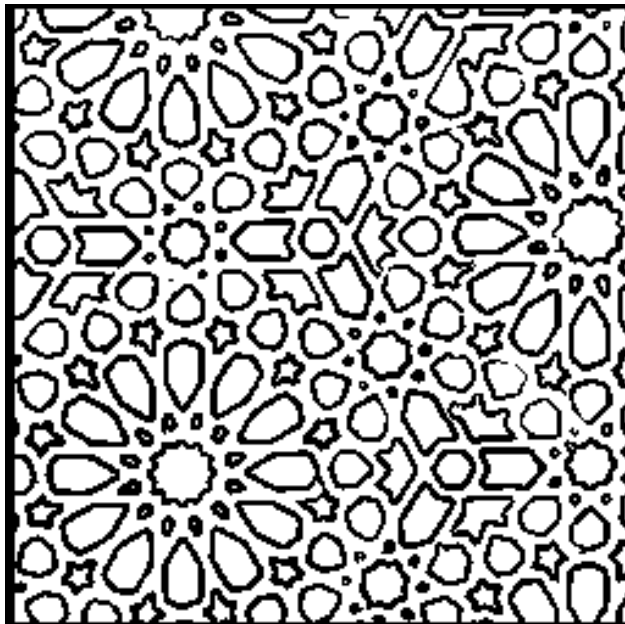


Figure 2: result of relaxation using correlation

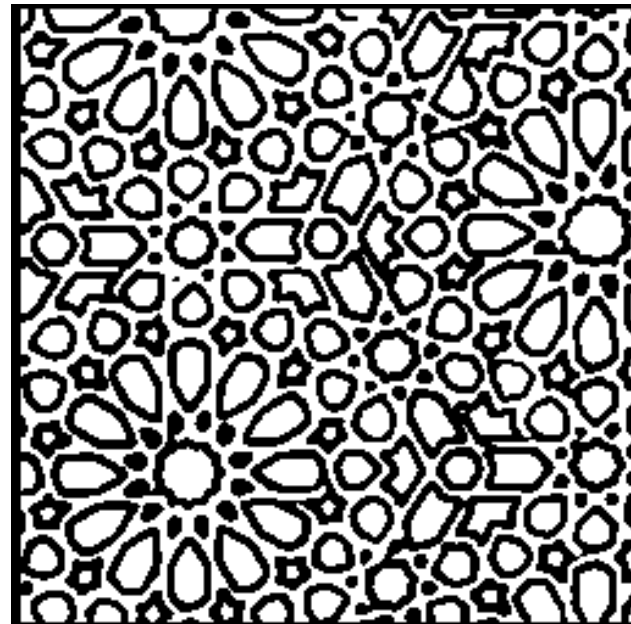


Figure 3: result of relaxation using mutual information

- [2] Q.T. Luong, "La couleur en vision par ordinateur: une revue", *Traitement du signal*, vol. 8, pp. 548–555, 1991.
- [3] T. Kanade, "Image understanding research at cmu", in *Alvey Vision Conf. ARPA Image Understanding Workshop*, 1987, pp. 32–40.
- [4] S.Di Zenzo, "A note on the gradient of multi-image", *Computer Vision, Graphics, and Image Processing*, vol. 33, pp. 116–125, 1986.
- [5] H.C Lee and D.R. Cok, "Detecting boundaries in a vector field", *IEEE Transactions on Signal Processing*, vol. 39, pp. 3–34, 1991.
- [6] A. Cumani, "A note on the gradient of multi-image edge detection in multispectral images", *Computer Vision, Graphics, and Image Processing*, vol. 53, pp. 40–51, 1991.
- [7] A.Danker and A.Rosenfeld, "Blob detection by relaxation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 702–710, 1981.
- [8] V.Ultre and L.Macaire, "Multispectral edge detection by relaxation algorithm", in *IST SPIES Symposium on Electronic Imaging*, 1995, vol. 4, pp. 252–258.
- [9] V.Ultre and H.Deleplanque, "Une nouvelle technique d'étiquetage probabiliste itératif pour l'extraction des contours", in *AGI' 94, Poitiers*, 1994, pp. 151–160.
- [10] S.Peleg and A.Rosenfeld, "Determining compatibility coefficients for curve enhancement relaxation processes", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 8, pp. 548–555, 1978.

