



HAL
open science

Étude du rôle de la couleur dans la perception visuelle des scènes naturelles

Camille Breuil

► **To cite this version:**

Camille Breuil. Étude du rôle de la couleur dans la perception visuelle des scènes naturelles. Sciences cognitives. Université Grenoble Alpes, 2018. Français. NNT : . tel-02628559v1

HAL Id: tel-02628559

<https://theses.hal.science/tel-02628559v1>

Submitted on 24 Oct 2019 (v1), last revised 26 May 2020 (v2)

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.

THÈSE

pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Sciences Cognitives, Psychologie et Neurocognition**

Arrêté ministériel : 25 mai 2016

Présentée par

Camille BREUIL

Thèse dirigée par **Nathalie GUYADER** et
codirigée par **Simon BARTHELMÉ**

Préparée au sein du

**Laboratoire Grenoble Images Parole Signal et Automatique
(GIPSA-lab)**

Dans l'École Doctorale Ingénierie pour la Santé, la Cognition
et l' Environnement (EDISCE)

Étude du rôle de la couleur dans la perception visuelle des scènes naturelles

The role of color in the visual perception of natural scenes

Thèse soutenue publiquement le **20 décembre 2018**,
devant le jury composé de :

M. Kenneth KNOBLAUCH

Directeur de Recherche, INSERM, Rapporteur

M. Peter NERI

Directeur de Recherche, CNRS, Rapporteur

M. Patrick LE CALLET

Professeur, Université de Nantes, Président du jury

M. Christian JUTTEN

Professeur, Université Grenoble Alpes, Examineur

Mme Nathalie GUYADER

Maître de conférences, Université Grenoble Alpes, Directrice de thèse

M. Simon BARTHELMÉ

Chargé de Recherche, CNRS, Encadrant de thèse



UNIVERSITÉ DE GRENOBLE ALPES
ÉCOLE DOCTORALE EDISCE
École Doctorale Ingénierie pour la Santé, la Cognition et l' Environnement

THÈSE

pour obtenir le titre de

docteur en sciences

de la Communauté Université Grenoble Alpes
**Mention : SCIENCES COGNITIVES, PSYCHOLOGIE ET
NEUROCOGNITION**

Présentée et soutenue par

Camille BREUIL

**Étude du rôle de la couleur dans la perception visuelle des
scènes naturelles**

Thèse dirigée par Nathalie GUYADER et Simon BARTHELMÉ
préparée au Laboratoire Grenoble Images Parole Signal et Automatique
(GIPSA-lab)

soutenue le 20 décembre 2018

Jury :

Rapporteurs : M. Kenneth KNOBLAUCH - Université de Lyon, France
M. Peter NERI - École Normale Supérieure, Paris, France
Directrice : Mme Nathalie GUYADER - Université Grenoble Alpes, France
Encadrant : M. Simon BARTHELMÉ - Université Grenoble Alpes, France
Président : M. Patrick LE CALLET - Université de Nantes, France
Examinateur : M. Christian JUTTEN - Université Grenoble Alpes, France

Remerciements

Je tiens à remercier, en premier lieu, Nathalie Guyader et Simon Barthelmé qui ont encadré cette thèse. Merci infiniment à tous les deux pour votre gentillesse, votre disponibilité, votre implication, vos nombreux conseils et vos précieuses relectures durant ces trois années. Merci aussi pour votre confiance et l'optimisme dont vous avez toujours fait preuve. J'ai énormément appris grâce à vous deux et votre complémentarité.

Je remercie également l'ensemble des membres du jury : Kenneth Knoblauch et Peter Neri, pour m'avoir fait l'honneur d'accepter d'être rapporteurs de ce manuscrit et pour leurs commentaires très enthousiastes sur celui-ci, Patrick Le Callet, qui a présidé le jury avec entrain et bienveillance, et Christian Jutten, qui a accepté le remplacement de dernière minute. Merci également à Valérie Goffaux, qui n'a finalement pas pu assister à la soutenance, mais dont les retours sur le manuscrit ont été précieux.

Merci à Frederick Kingdom et Ben Jennings, pour leur accueil à Montréal et leur collaboration sur une partie des travaux de cette thèse. Merci aussi à Léa Entzmann, Léo Martin et Bastien Faure-Brac qui, dans le cadre de leurs stages ou projets, ont conduit certaines des expériences présentées dans ce manuscrit. Merci à tous ceux qui nous ont conseillé sur des aspects théoriques, techniques ou expérimentaux de cette thèse, ainsi qu'à tous les participants des expériences.

Cette aventure n'aurait pas été la même sans tous ceux qui ont rendu ces trois années si agréables : la grande famille des fanfarons (Odyssound, la Farigoule, la Mort Subite, les Van Hornies,...), le BDV pour la cellule de soutien hebdomadaire (notamment toi, Corky, avec qui nous avons parcouru tant de chemin ensemble ces dernières années), ceux qui, même en étant loin, ont toujours été là (le fameux groupe,...), ceux dont la présence éclairait le quotidien (les collègues du GIPSA, les partenaires de coinche, de pause, de conférence,..) et toutes celles et ceux que je ne cite pas explicitement ici mais qui, j'espère, se reconnaîtront.

Enfin, je n'aurai pas les mots pour exprimer en quelques lignes ma gratitude envers eux, pour m'avoir toujours accompagnée de leur soutien et leur amour indéfectibles : mes sœurs, mon frère, mes parents, grands-parents et beaux-parents. Je vous aime.

Merci.

Parce qu'on ne peut pas vivre sans quelqu'un à aimer.

Romain Gary, *La vie devant soi*

Table des matières

Acronymes

Avant-propos	1
1 État de l'art	3
1.1 Rôle de la couleur dans la perception visuelle	4
1.2 La vision des couleurs	7
1.2.1 La lumière : stimulus visuel	7
1.2.2 Théories de la vision des couleurs	10
1.2.3 Traitement de l'information par le système visuel : des photorécepteurs au cortex visuel	10
1.2.4 Propriétés spatio-temporelles de la vision des couleurs	24
1.2.5 Étudier la couleur dans les scènes naturelles	30
1.3 Objectifs de la thèse et approche théorique	31
1.3.1 L'approche mécanistique	32
1.3.2 L'observateur bayésien	34
2 Quantifier l'indépendance statistique des informations de luminance et de chrominance dans les scènes naturelles	41
2.1 Analyses statistiques des propriétés spatio-chromatiques des scènes naturelles .	42
2.2 Estimation de l'information mutuelle	46
2.2.1 Méthode	46
2.2.2 Résultats	57

2.2.3	Discussion	71
2.2.4	L'information mutuelle : un indice pertinent ?	74
2.3	Conclusion	76
3	Un modèle mécanistique de la chrominance pour la perception visuelle des scènes naturelles	79
3.1	Introduction	80
3.2	Expérience 1 - Détection de couleur	84
3.2.1	Hypothèse	84
3.2.2	Design expérimental	84
3.2.3	Analyse des résultats	89
3.2.4	Résultats	96
3.3	Observateur décalé	99
3.4	Expérience 2 - Détection de luminance	103
3.4.1	Hypothèse	103
3.4.2	Design expérimental	103
3.4.3	Résultats	104
3.5	Expérience 3 - Détection de couleur (contrôle)	106
3.5.1	Hypothèse	106
3.5.2	Design expérimental	106
3.5.3	Résultats	107
3.6	Discussion	108
3.6.1	Synthèse des résultats	108
3.6.2	Cas du mécanisme logarithmique - Adaptation des photorécepteurs . . .	109
3.6.3	Perspectives	109
3.7	Conclusion	111
4	Rôle de la couleur dans la classification de contours	113

Table des matières

4.1	Introduction	114
4.2	Hypothèses	115
4.3	Expérience de classification de contours	116
4.3.1	Méthode	116
4.3.2	Résultats	121
4.4	Classifieurs	123
4.4.1	Modèle de classifieur bayésien	123
4.4.2	Propriétés des images	123
4.4.3	Analyse discriminante linéaire	127
4.4.4	Résultats	128
4.5	Comparaison des performances des observateurs humains et des classifieurs	130
4.6	Discussion	131
4.6.1	Synthèse des résultats	131
4.6.2	Perspectives	131
4.7	Conclusion	132
5	Synthèse et perspectives	133
5.1	Synthèse des principaux résultats	133
5.1.1	Analyse statistique des scènes naturelles	134
5.1.2	Description mécanistique de la vision des couleurs	134
5.1.3	Rôle de la couleur pour la classification de contours	135
5.2	Perspectives	135
5.2.1	Lien entre information mutuelle et rôle de la couleur	135
5.2.2	Un modèle mécanistique pour la perception des couleurs dans les scènes naturelles	148
	Bibliographie	167

A	Conversion de l'espace des couleurs RGB vers l'espace DKL	169
B	Propriétés des écrans d'expérimentation	173
B.1	Iiyama Vision Master Pro 513	173
B.2	Philips 201B4	174
C	Valeurs d'information mutuelle estimées avec le modèle de référence	177

Acronymes

2AFC	<i>Two-Alternative Forced Choice ou Choix Forcé à Deux Alternatives</i>
ACI	<i>Analyse en Composantes Indépendantes</i>
ACP	<i>Analyse en Composantes Principales</i>
ADL	<i>Analyse Discriminante Linéaire</i>
AIC	<i>Akaike Information Criterion ou Critère d'Information d'Akaike</i>
BY	<i>Blue-Yellow ou Bleu-Jaune</i>
Chrom	<i>Chrominance</i>
CIE	<i>Commission Internationale de l'Éclairage</i>
CRT	<i>Cathode-Ray Tube ou Tube à Rayonnement Cathodique</i>
DKL	<i>Derrington-Krauskopf-Lennie</i>
DSRP	<i>Distribution Spectrale Relative de Puissance</i>
DV	<i>Digital Value ou Valeur Digitale</i>
FA	<i>Fausse Alarme</i>
FSC	<i>Fonction de Sensibilité au Contraste</i>
IM	<i>Information Mutuelle</i>
IMN	<i>Information Mutuelle Normalisée</i>
LGN	<i>Lateral Geniculate Nucleus ou Corps Géniculé Latéral</i>
Lum	<i>Luminance</i>
RG	<i>Red-Green ou Rouge-Vert</i>
RGB	<i>Red Green Blue ou Rouge Vert Bleu</i>
TIFF	<i>Tagged Image File Format</i>
V1	<i>Cortex Visuel Primaire</i>

Avant-propos

[La couleur] est "l'endroit où notre cerveau et l'univers se rejoignent", dit [Cézanne].. Il ne s'agit donc pas de couleurs,.. simulacre des couleurs de la nature, il s'agit de la dimension, de couleur, celle qui crée d'elle-même à elle-même des identités, des différences, une texture, une matérialité, un quelque chose.

Maurice Merleau-Ponty, *L'Oeil et l'esprit*, 1964

La couleur est un des aspects irréductibles et parmi les plus fascinants de la perception visuelle. Par nature ambivalente, la couleur semble appartenir tout à la fois au monde extérieur et à notre intériorité, elle est objective et subjective, constante et fluctuante. Souvent considérée comme une propriété inhérente aux objets, elle est pourtant une sensation construite par notre cerveau. La vision des couleurs est une intrigante énigme reliant intimement et inextricablement propriétés physiques et psychologiques. Elle a interrogé au cours des siècles derniers philosophes, artistes, physiciens ou encore psychologues, et leurs découvertes ont été accompagnées de passionnantes converses. Depuis les théories les plus anciennes basées sur l'existence de rayons projetés par les yeux jusqu'aux descriptions computationnelles modernes, la couleur n'a eu de cesse de captiver.

Plus récemment, l'approche cognitive a cherché à comprendre les mécanismes permettant de passer de phénomènes physiques à des perceptions. Abordée de ce point de vue, la vision des couleurs recèle encore de nombreux mystères et constitue un microcosme intéressant où les différentes disciplines des sciences cognitives ont chacune apporté d'importantes contributions. Si la psychophysique et la neurophysiologie ont extensivement étudié la perception des couleurs à partir de stimuli simples et artificiels, le cas des scènes plus complexes a été assez peu abordé.

Dans cette thèse, nous avons exploré la question du traitement et du rôle de l'information de couleur dans un contexte écologiquement pertinent : la perception visuelle des scènes naturelles. Nous avons cherché à comprendre un peu mieux pourquoi et comment nous percevons les couleurs dans notre environnement. Nous proposons pour cela une série d'études théoriques et expérimentales basées sur de nouveaux paradigmes exploitant des images de scènes naturelles. Ce travail de thèse se positionne donc dans l'objectif général d'élaborer des modèles computationnels du traitement visuel appropriés à la perception des scènes naturelles, avec un intérêt spécifique pour le traitement de la couleur.

1

État de l'art

We are so familiar with seeing, that it takes a leap of imagination to realize that there are problems to be solved. But consider it. We are given tiny distorted upside-down images in the eyes and we see separate solid objects in surrounding space. From the patterns of stimulation on the retina we perceive the world of objects and this is nothing short of a miracle.

Richard Gregory, *Eye and Brain*, 1966

Sommaire

1.1	Rôle de la couleur dans la perception visuelle	4
1.2	La vision des couleurs	7
1.2.1	La lumière : stimulus visuel	7
1.2.2	Théories de la vision des couleurs	10
1.2.3	Traitement de l'information par le système visuel : des photorécepteurs au cortex visuel	10
1.2.4	Propriétés spatio-temporelles de la vision des couleurs	24
1.2.5	Étudier la couleur dans les scènes naturelles	30
1.3	Objectifs de la thèse et approche théorique	31
1.3.1	L'approche mécanistique	32
1.3.2	L'observateur bayésien	34

Voir peut nous apparaître en premier lieu comme quelque chose de simple et automatique. Pourtant, notre environnement visuel est particulièrement riche et complexe, et en acquérir une telle perception visuelle consciente n'a rien d'évident. Par quels traitements opérés sur la lumière pénétrant dans nos yeux aboutissons nous si rapidement, sans effort, à une représentation sensée, cohérente, tri-dimensionnelle et colorée de notre environnement ? La perception visuelle est le phénomène par lequel nous parvenons à acquérir une connaissance du monde et des objets qui nous entourent en extrayant de l'information à partir de la lumière qu'ils émettent ou réfléchissent. Bien que les sciences de la vision nous aient permis au cours des siècles derniers de faire des avancées considérables dans la compréhension de ces mécanismes, de nombreuses questions restent toujours sans réponse définitive, parmi elles : pourquoi et comment percevons-nous la couleur ? Étant donné que le système visuel humain a développé des mécanismes spécialisés dans le traitement de la couleur, on peut légitimement se demander quel rôle fonctionnel peut jouer cette information dans la perception. Avant de répondre à cette question, il semble nécessaire de comprendre un peu mieux ce qu'est la couleur, et par quel traitement notre cerveau nous rend cette information accessible.

Dans ce premier chapitre, nous commençons par présenter brièvement les travaux les plus importants concernant le rôle de la couleur dans la perception visuelle. Dans une seconde partie, nous détaillons les principales étapes du traitement visuel de la couleur. Enfin, une troisième partie décrit les objectifs de ce travail de thèse et les approches théoriques que nous avons adoptées pour celui-ci.

1.1 Rôle de la couleur dans la perception visuelle



FIGURE 1.1 – Scène extraite du film *La liste de Schindler* (S. Spielberg, 1993)

Dans le film *La liste de Schindler*, réalisé en 1993 entièrement en Noir et Blanc, Steven Spielberg n'a inséré qu'une légère touche de couleur : le manteau rouge d'une petite fille dé-

portée vers les camps de concentration (Fig. 1.1). Attirant inexorablement l'œil du spectateur lors de ses diverses apparitions, ce détail rend encore plus poignant le destin tragique du personnage. Si dans le domaine artistique, l'apport des qualités esthétiques de la couleur est indéniable et depuis longtemps exploité, son importance dans les phénomènes perceptifs qui structurent notre expérience visuelle quotidienne n'a cependant toujours pas été clairement établie. De nombreuses questions se posent et diverses hypothèses ont été formulées au sujet du rôle de la couleur lors des premières étapes du traitement visuel (niveau sensoriel) ou dans des processus plus complexes (niveau cognitif). Nous présentons ici quelques travaux parmi les plus importants sur cette question au regard de notre problématique.

Les primates sont les seuls mammifères à avoir, au cours de l'évolution, acquis une vision trichromatique, leur permettant ainsi d'extraire de leur environnement visuel plus d'information que des individus dichromates ou monochromates. En plus d'une information de luminance, ils ont également accès à une information de couleur (chrominance) (Nathans, 1999). Quels avantages écologiques présente cette aptitude ? D'un point de vue évolutif, il a été suggéré que la trichromatie était apparue, et s'était maintenue, car elle simplifiait la recherche de nourriture en facilitant la détection de fruits mûrs (Allen, 1879 ; Mollon, 1989 ; Sumner et Mollon, 2000b ; Sumner et Mollon, 2000a ; Regan et al., 2001) ou de jeunes pousses (Lucas et al., 1998 ; Dominy et Lucas, 2001 ; Sumner et Mollon, 2000b) parmi les feuillages. D'autres théories ont été formulées pour expliquer l'émergence de la vision trichromatique. Changizi et al. (2006) ont notamment suggéré que la couleur pourrait jouer un rôle crucial dans la détection des émotions (en permettant, par exemple, de mieux déceler un afflux sanguin au niveau du visage) facilitant ainsi la communication non-verbale et les interactions sociales. La détection d'objets d'intérêt (tels que de la nourriture) dans leur contexte naturel (comme des feuillages) reste cependant l'explication la plus répandue à l'évolution du système visuel vers la trichromatie, et quoi qu'il en soit, une tâche pour laquelle la vision de la couleur présente un avantage certain.

Au delà de cet intérêt adaptatif, on peut s'interroger sur l'existence d'autres bénéfices de l'information de chrominance pour la perception visuelle humaine. Il a ainsi depuis longtemps été montré que la couleur constituait un indice utile pour la segmentation et l'organisation des stimuli visuels en objets et en scènes (Koffka, 1935 ; Barlow, 1980 ; Cavanagh, 1987 ; Cavanagh, 1991 ; Fine et al., 2003). Kingdom, Beauce et al. (2004) ont également proposé que la couleur serait plus informative que la luminance pour distinguer les objets (les variations de luminance pouvant aussi résulter de la simple présence d'ombres, indépendamment d'objets). La Figure 1.2 illustre deux exemples d'utilité de la couleur dans un contexte naturel (la détection de fruits rouges et la segmentation d'objets).

Plusieurs études ont, par ailleurs, mis en évidence que la couleur pouvait améliorer la reconnaissance et la catégorisation d'objets (Wurm et al., 1993 ; Tanaka et Presnell, 1999, voir Bramão et al., 2011 pour une revue), incluant les visages, (Russell et Sinha, 2007). Les mêmes observations ont aussi été faites pour les scènes naturelles, notamment dans les cas où la couleur avait une valeur diagnostique (i.e. prédictive) élevée (par exemple, la couleur rouge est fortement diagnostique d'une tomate, le bleu de la mer,...) (Gegenfurtner et Rieger, 2000 ; Oliva et Schyns, 2000 ; Goffaux et al., 2005 ; Castelhana et Henderson, 2008). Comme



FIGURE 1.2 – Exemples d'utilité de la couleur dans un contexte naturel. En haut : La couleur facilite la détection de fruits rouges parmi les feuillages. En bas : La couleur facilite la segmentation des différents arbres présents dans l'image.

souligné par Gegenfurtner et Rieger (2000), la contribution de la couleur dans la reconnaissance des objets et des scènes naturelles se trouve à deux niveaux. D'une part au niveau sensoriel, précoce, où la couleur améliore le codage de l'information visuelle en facilitant la segmentation de l'image. D'autre part, au niveau cognitif, tardif, où la couleur contribue à la récupération d'informations provenant de la mémoire visuelle. En résumé, la couleur nous aide à reconnaître les objets plus rapidement et à mieux s'en souvenir (Wichmann et al., 2002; Spence et al., 2006). D'autres suggèrent cependant que la couleur n'interviendrait pas dans les processus de reconnaissance, que ce soit pour des objets isolés (Biederman et Gerhardstein, 1993) ou présentés dans leur contexte naturel (Delorme et al., 2000).

D'autres études, se sont intéressées à l'utilité de la couleur dans le déploiement de l'attention visuelle pour la perception des scènes naturelles. Il a, entre autres, été montré que la couleur pouvait influencer la position (Amano et Foster, 2014; Frey, Honey et al., 2008; Frey, Wirz et al., 2011) et la durée des fixations (Von Wartburg et al., 2005; Ho-Phuoc et al., 2012), orienter la recherche de cibles (Amano, Foster et al., 2012) ou encore améliorer les performances de modèles de saillance (Jost et al., 2005).

Dans cette thèse nous nous intéresserons essentiellement aux aspects sensoriels de la couleur, c'est-à-dire à l'encodage et l'utilisation de cette information à un stade précoce du traitement visuel. De nombreux travaux ont, par ailleurs, questionné son rôle fonctionnel d'un point de vue cognitif (Davidoff et al., 1997, voir Tanaka, Weiskopf et al., 2001 pour une revue)

et psychologique (Elliot, 2015 pour une revue).

Afin d'étudier le rôle de la couleur dans la perception visuelle, il est avant tout nécessaire d'en fournir une définition pertinente. Il faut pour cela comprendre quelle information sensorielle est extraite de notre environnement par le système visuel et par quels mécanismes nous parvenons à voir les couleurs.

1.2 La vision des couleurs

"Color vision is an illusion created by the interactions of billions of neurons in our brain."

—Peter Gouras, 2008

Nous présentons dans cette partie les principales connaissances actuelles concernant la vision des couleurs. Partant de la description des propriétés physiques de la lumière, nous décrivons ensuite les étapes essentielles du traitement visuel, autrement dit les interactions entre neurones, permettant d'aboutir à la perception de la couleur. Plusieurs revues synthétisent les résultats connus concernant la physiologie et l'anatomie de la vision des couleurs (Gegenfurtner et Kiper, 2003 ; Lennie et Movshon, 2005 ; Solomon et Lennie, 2007 ; Conway, 2009).

1.2.1 La lumière : stimulus visuel

En plaçant un prisme devant un rayon de lumière naturelle (blanche), Sir Isaac Newton parvient à la séparer en plusieurs composants produisant chacun une sensation colorée différente. Il est alors le premier à formuler l'hypothèse que la couleur n'est pas dans la lumière elle-même, ni créée par le prisme, mais qu'elle résulte d'un effet de la lumière sur le système visuel :

"[T]he Rays to speak properly are not coloured. In them there is nothing else than a certain Power and Disposition to strip up a Sensation of this or that Colour... So Colours in the object are nothing but a Disposition to reflect this or that sort of Rays more copiously than the rest."

—Isaac Newton, 1704

Cette découverte constitue l'un des fondements de la science de la vision des couleurs. L'apparence colorée d'un objet dépend en fait de trois éléments essentiels :

- une source lumineuse
- l'objet et ses caractéristiques de réflectance
- un observateur

1.2.1.1 Description physique de la lumière

La lumière peut être décrite de deux façons : à la fois comme une onde électromagnétique et comme un flux de particules (les photons). En fonction des phénomènes auxquels on s'intéresse, l'une ou l'autre de ces descriptions se montrera plus adaptée. Quelle que soit la description choisie, une lumière est caractérisée par sa longueur d'onde, qui correspond soit à l'énergie de l'onde, soit à celle de vibration des photons. Nos organismes ne sont sensibles qu'à une petite partie de l'ensemble du spectre électromagnétique (située environ entre 400 et 700 nm), gamme que nous appelons le spectre visible (Fig. 1.3). Les ondes radios, les rayons X ou encore la lumière infra-rouge font également partie de ce spectre mais l'humain ne dispose d'aucun récepteur sensoriel répondant à des ondes de ces longueurs, nous les rendant ainsi imperceptibles.

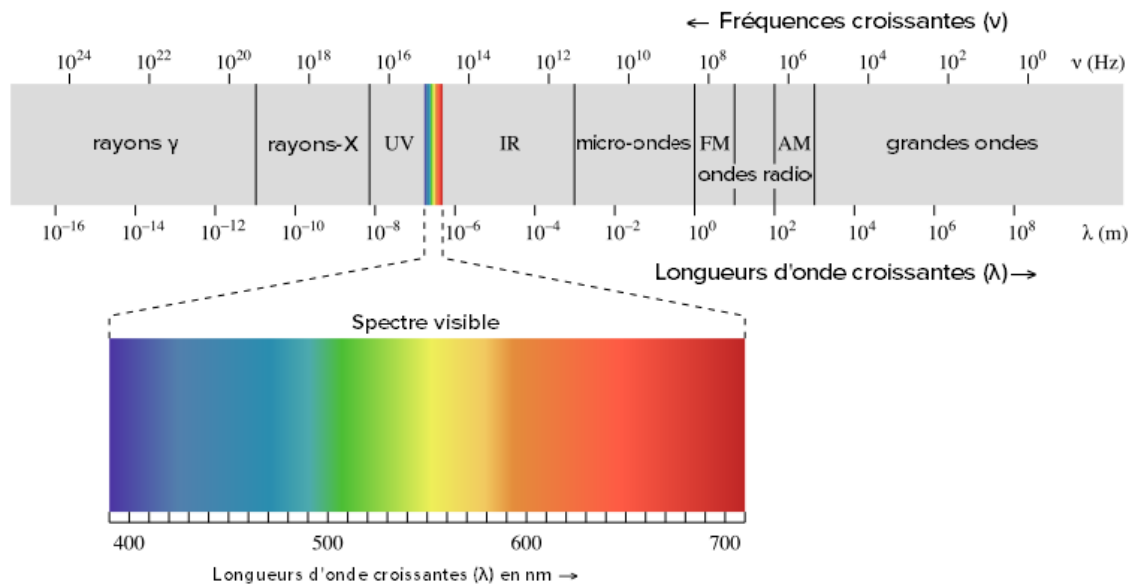


FIGURE 1.3 – Spectre électromagnétique (image issue de UC Davis ChemWiki, <https://chem.libretexts.org>)

1.2.1.2 Lumière et environnement visuel

Notre environnement visuel contient un certain nombre de sources lumineuses (le soleil, une lampe,...). La lumière émise par chacune de ces sources peut être caractérisée par son spectre électromagnétique, c'est-à-dire la répartition de son énergie en fonction de la longueur d'onde. On parle d'illuminant pour désigner la lumière qui éclaire un objet ou une scène. La Commission Internationale de l'Éclairage (CIE) définit et publie les caractéristiques spectrales des illuminants normalisés, chacun portant un nom constitué d'une lettre et éventuellement d'un chiffre. La Figure 1.4 montre les distributions spectrales relatives des illuminants A, C et D65, qui représentent respectivement la lumière moyenne d'une lampe à incandescence, une

lumière de jour moyenne et l'étalon colorimétrique de référence d'une lumière en plein jour en zone tempérée. La lumière naturelle possède un large spectre correspondant à un grand nombre de particules oscillant à différentes amplitudes, à des fréquences différentes et dans plusieurs directions de l'espace.

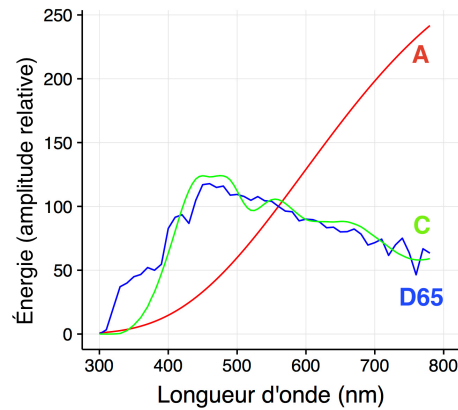


FIGURE 1.4 – Distributions Spectrales Relatives de Puissance (DSRP) des illuminants A, C et D65 des la CIE entre 300 et 800 nm

Quand les ondes émises par une source rencontrent la surface d'un objet, la majeure partie d'entre elles va être absorbée ou réfléchi. Ces phénomènes physiques sont représentés sur la Figure 1.5. Les objets ne vont pas tous réfléchir ou absorber la lumière de la même façon. On parle de spectre de réflectance pour décrire la fraction de lumière réfléchi par un objet en fonction de la longueur d'onde. Ainsi, l'apparence d'un objet va dépendre de la lumière qui parviendra finalement à l'œil de l'observateur, et donc à la fois des propriétés de l'illuminant et de celles de l'objet (spectre de réflectance). C'est cette lumière qui, après avoir pénétré dans l'œil, va suivre une série de traitements, opérée par le système visuel, qui va aboutir au percept de couleur.

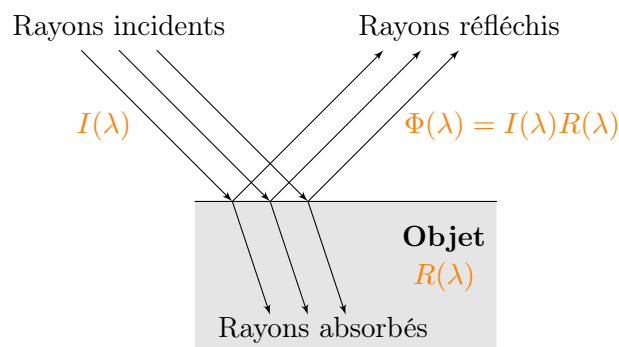


FIGURE 1.5 – Interaction de l'illuminant ($I(\lambda)$) avec un objet (de réflectance $R(\lambda)$). Une partie des rayons est absorbée par l'objet, l'autre est réfléchi. Le spectre des rayons réfléchis est le produit des spectres de l'illuminant et du réflectant : $\Phi(\lambda) = I(\lambda)R(\lambda)$.

1.2.2 Théories de la vision des couleurs

Historiquement, la première théorie de la vision des couleurs fût formulée par Sir Thomas Young dès la fin du XVIII^e siècle. Elle repose sur l'existence de trois récepteurs de sensibilités différentes permettant de produire les sensations primaires du rouge, du vert et du bleu, chacune des autres couleurs pouvant être expliquée comme une combinaison de ces trois primaires. Cette théorie fût par la suite affinée et développée par Maxwell (1855) et Von Helmholtz (1867), et est connue sous le nom de *théorie trichromatique de Young-Helmholtz*.

En dépit du succès de cette théorie, celle-ci présentait tout de même quelques inconsistances, notamment au sujet de la phénoménologie de la perception des couleurs : comment expliquer par exemple qu'une couleur ne puisse jamais apparaître comme simultanément bleue/jaune ou rouge/verte ? Plusieurs observations semblaient suggérer une sorte d'appariement des couleurs. Partant de ce constat, Ewald Hering proposa l'existence de deux couples de processus : l'un opposant le rouge au vert et l'autre le bleu au jaune. Cette théorie est connue sous le nom de théorie des oppositions de couleur (Hering, 1878).

Pendant plusieurs décennies ces deux théories se sont affrontées. Face à ce qui semblait être une impasse, c'est finalement Hurvich et Jameson (1957) qui mirent définitivement fin à cette controverse. Ils proposèrent, et apportèrent des preuves expérimentales fiables, à une élégante théorie à deux étapes, chacune des deux théories précédentes étant valable pour l'une des étapes. Aujourd'hui, nous savons que le traitement de la couleur est, en effet, initié au niveau des photorécepteurs de la rétine qui mesurent trois composantes chromatiques, puis ces composantes sont encodées sous forme de canaux d'opposition. C'est ces différentes théories qui ont, en partie, guidé les travaux de physiologie et de psychophysique sur la perception des couleurs dont nous allons présenter ici les principaux résultats.

1.2.3 Traitement de l'information par le système visuel : des photorécepteurs au cortex visuel

Dans cette partie, nous décrivons les principales étapes du traitement visuel impliquées dans la perception de la couleur. Nous présentons à la fois des résultats obtenus par des approches physiologiques et psychophysiques. Une partie des résultats électrophysiologiques présentés concerne le macaque, mais son système de vision trichromatique est très proche de celui de l'homme (De Valois, 1965 ; Jacobs, 1993).

La voie principale du traitement visuel, et donc du traitement de la couleur, comprend trois grandes étapes : la rétine, le LGN (*Lateral Geniculate Nucleus* ou Corps Géniculé Latéral) et le cortex visuel. Nous allons présenter les propriétés connues de ces différentes structures et leurs contributions respectives à l'extraction de l'information chromatique sensorielle.

La Figure 1.6 présente la voie principale du traitement visuel de la couleur, ainsi qu'une représentation anatomique schématique des trois principales structures impliquées.

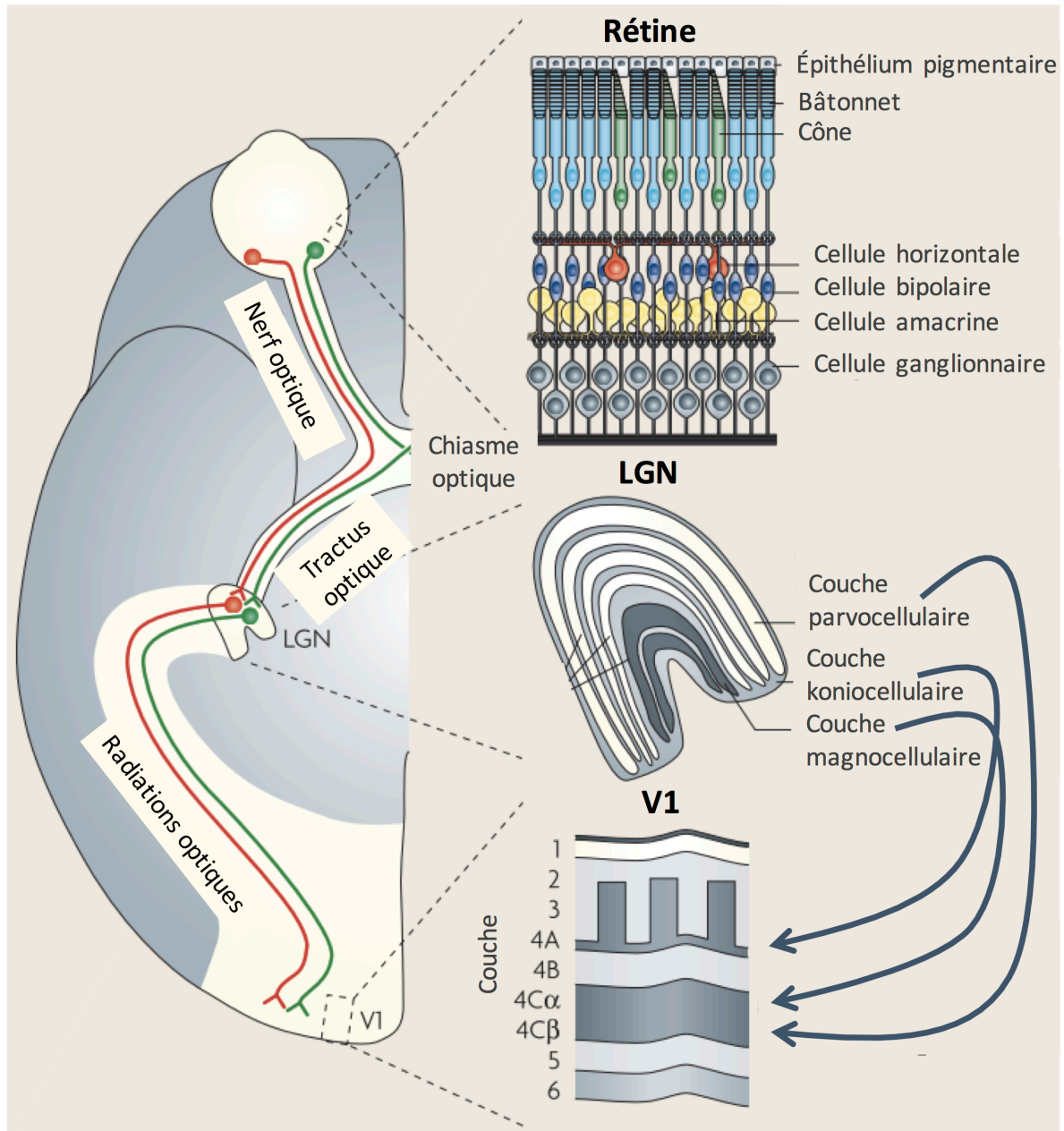


FIGURE 1.6 – Voie principale du traitement de l'information visuelle (adapté de Solomon et Lennie (2007)).

À gauche : Vue schématisée, de la rétine au cortex visuel primaire (V1), en passant par le LGN. À droite : Organisation anatomique des trois principales structures (rétine, LGN et V1). Chacune d'entre elles est constituée de différentes couches. La lumière pénétrant dans l'œil, dans un premier temps, captée par les photorécepteurs de la rétine (cônes et bâtonnets). Le signal enregistré par les photorécepteurs est ensuite majoritairement transmis au LGN, via les cellules ganglionnaires. Celles-ci sont de plusieurs types et projettent dans différentes couches du LGN. L'information est enfin transmise au cortex visuel primaire (V1), principalement au niveau de la couche 4 de celui-ci.

1.2.3.1 L'œil

L'œil est le point d'entrée du système visuel. Grâce au dispositif optique formé par la cornée, la pupille et le cristallin, l'œil focalise la lumière incidente au niveau de sa paroi interne, la rétine (Fig. 1.7).

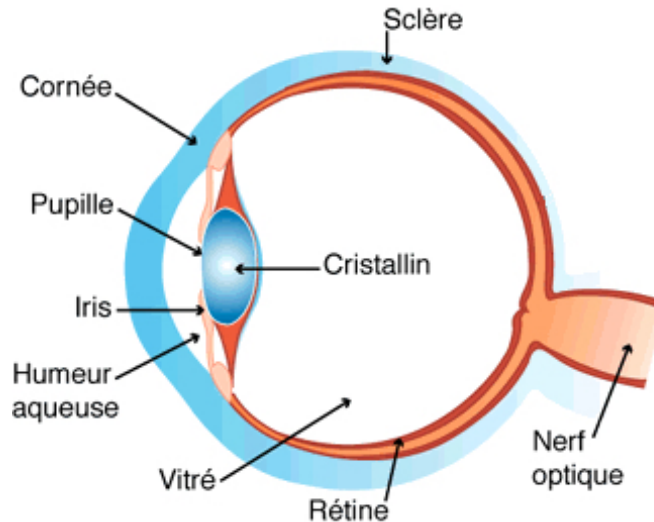


FIGURE 1.7 – Schéma d'un œil humain. Après avoir traversé les différentes structures de l'œil, la lumière est captée par les photorécepteurs de la rétine, qui vont transformer cette information en signal nerveux.

La cornée est la partie antérieure transparente du globe oculaire. Sa surface est courbée et elle est remplie d'un liquide nommé humeur aqueuse, dont l'indice de réfraction est différent de celui de l'air. Elle agit donc comme une lentille qui va focaliser la lumière vers la surface interne de l'œil.

La lumière traverse ensuite la pupille, fenêtre de taille variable contrôlée par l'iris. L'iris donne la couleur à l'œil. Il est percé en son centre par une ouverture circulaire, la pupille, qui se dilate ou se contracte selon l'intensité de la lumière, grâce à l'action des muscles lisses de l'iris.

Le cristallin, situé derrière la cornée, joue lui aussi un rôle de lentille. Il peut, sous l'action des muscles ciliaires, modifier sa courbure et permettre ainsi de former sur la rétine une image nette de l'objet regardé, quelle que soit sa distance. C'est le phénomène d'accommodation.

1.2.3.2 La rétine

Au fond de l'œil se trouve la rétine, organe sensible de la vision. La rétine peut être décrite de façon simplifiée comme trois couches (de la partie interne de l'œil vers la partie externe) : la première est constituée des cellules nerveuses, la seconde de cellules photoréceptrices et la dernière est l'épithélium pigmentaire. La lumière qui pénètre dans l'œil va traverser chacune

de ces couches dans cet ordre. La chaîne de traitement visuel démarre véritablement avec l'absorption des photons par les cellules photoréceptrices. Celles-ci vont permettre de transformer le signal électromagnétique de la lumière en signal électrochimique, qui sera ensuite transmis au système nerveux. Cette transformation se nomme *transduction* et est réalisée grâce à une chaîne de réactions photochimiques complexes (voir Jindrová, 1998 pour une revue).

Il existe deux types de photorécepteurs dans la rétine : les bâtonnets et les cônes (Fig.1.8).

Les plus nombreux sont les bâtonnets (environ entre 92 (Standring et Gray, 2008) et 100 (Hall et Guyton, 2011) millions par rétine chez l'humain). Ils sont sensibles aux faibles illuminations (conditions *scotopiques*) et deviennent vite saturés quand ils sont exposés à des niveaux d'illumination plus importants. Leur distribution n'est pas uniforme sur la rétine : leur densité est maximale à une excentricité d'environ 20° par rapport à l'axe optique et ils sont totalement absents dans la région d'environ 0.5 mm de diamètre (soit 1.7° d'angle visuel) située au centre de la rétine.

Les cônes sont moins nombreux (entre 3 et 4 millions par rétine chez l'humain) et sensibles à des intensités lumineuses plus importantes (conditions *photopiques*). Ils sont majoritairement concentrés au centre de la rétine, dans une région nommée *fovea* qui mesure environ 1.5 mm de diamètre (soit 5° d'angle visuel) (Wandell, 1995) (Figure 1.8). Il existe trois types de cônes et c'est eux qui sont responsables de la vision des couleurs. Quelques études suggèrent que les bâtonnets pourraient également jouer un rôle dans la perception des couleurs (Stabell, Definitions et al., 1998 ; Stabell et Stabell, 2002 ; Knight et al., 1998). Étant donné que les bâtonnets sont rapidement saturés quand les niveaux d'illumination augmentent (Normann et Werblin, 1974 ; DeValois et DeValois, 1990) et qu'ils sont totalement absents de la région de la rétine dédiée à la vision centrale, nous supposons que leur contribution peut être ignorée. Nous nous intéresserons donc pour la suite exclusivement à l'information visuelle fournie par les cônes.

Les cônes

Propriétés des cônes

Plus de 90% de la lumière qui pénètre notre œil n'est absorbée par aucun photorécepteur (Baylor, 1987). La probabilité qu'un photon soit absorbé par un photorécepteur dépend à la fois de son énergie (i.e. de la longueur d'onde de la lumière) et du type de photorécepteur. Il existe 3 types de cônes, chacun étant sensible à une certaine gamme de longueurs d'onde, et donc plus ou moins susceptible d'absorber un photon à une énergie donnée. Une fois le photon absorbé, la probabilité que le processus de transduction aboutisse est de 70% (Goldsmith, 1991). La seule information alors transmise au système nerveux est qu'un photon a été absorbé : un flux de photons d'intensité modérée à une longueur d'onde très favorable ou un flux plus intense à une longueur d'onde moins favorable produira donc un taux d'absorption comparable et l'information transmise par un cône sera identique. On parle de principe d'*univariance* (Rushton, 1972).

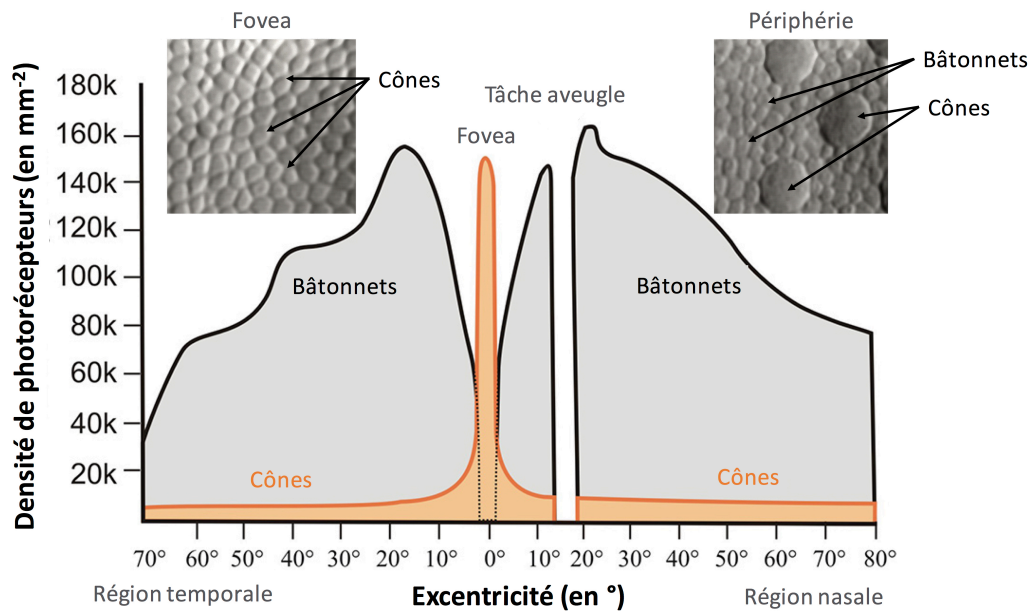


FIGURE 1.8 – Répartition des photorécepteurs dans la rétine. Les cônes sont majoritairement présents dans la *fovea*, région centrale de la rétine où les bâtonnets sont eux absents. Quand l'excentricité augmente, la densité de cônes décroît rapidement et leur taille augmente. La densité des bâtonnets est elle maximale pour une excentricité d'environ 20° par rapport à l'axe optique. La tâche aveugle correspond à l'endroit où le nerf optique et les vaisseaux sanguins quittent l'œil, les photorécepteurs y sont donc totalement absents.

Les trois types de cônes de la rétine humaine sont nommés d'après leurs sensibilités respectives au spectre des longueurs d'ondes : les cônes L (pour *Long*) sont sensibles aux grandes longueurs d'onde (pic à environ 560 nm), les M (pour *Medium*) aux longueurs d'onde moyennes (pic à environ 530 nm) et les S (pour *Short*) aux courtes longueurs d'onde (pic à environ 440 nm) (Schnapf et al., 1987).

La taille des cônes n'est pas uniforme sur toute la rétine (de 1 à 4 μm de diamètre dans la *fovea* et 4 à 10 μm en périphérie), tout comme leur distribution spatiale. Les cônes S sont bien moins nombreux que les deux autres types et ne représentent qu'environ 5% du nombre total de cônes (Curcio et al., 1991 ; Wandell, 1995 ; Hofer, 2005 ; Roorda et Williams, 1999) . Ils ont également pour particularité d'être quasiment absents de la *foveola*, région centrale de la rétine de 0.3 mm de diamètre (Curcio et al., 1991 ; Williams et al., 1981).

Tout comme les pics de maximum de sensibilité (Valberg, 2005), le ratio de chacun des types de cônes présente de très larges variations inter-individuelles (Carroll et al., 2002 ; Hofer, 2005) et leur arrangement sur la rétine semble être aléatoire (Hofer, 2005 ; Mollon et Bowmaker, 1992) (Fig. 1.9).

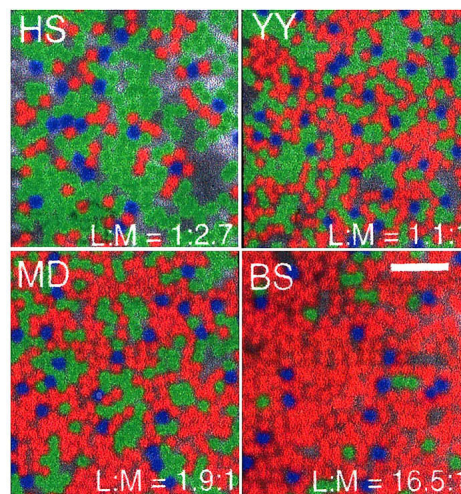


FIGURE 1.9 – Représentations obtenues par optique adaptative de la mosaïque des cônes de la rétine de 4 sujets humains (HS, YY, MD et BS). Les trois types de cônes, L (en rouge), M (en vert) et S (en bleu) sont répartis aléatoirement et avec des ratios différents. (d'après P. Gouras, <https://webvision.med.utah.edu/book/part-vii-color-vision/color-vision>).

Les différentes propriétés connues de la mosaïque rétinienne des cônes et leurs implications pour la vision des couleurs ont été recensées dans une récente revue (Brainard, 2015).

Sensibilité du système visuel

Fonction d'efficacité lumineuse

L'efficacité lumineuse spectrale est une fonction qui exprime, pour une longueur d'onde donnée, le rapport entre le flux énergétique du rayonnement électromagnétique reçu et la perception d'intensité lumineuse induite pour la vision humaine. Elle caractérise la sensibilité du système visuel humain aux différentes longueurs d'onde. Plusieurs fonctions ont été normalisées par la CIE, définissant un observateur standard dans différentes conditions (vision scotopique ou photopique, angle d'observation de 2 ou 10 ° d'angle visuel). La plus ancienne d'entre elles, concernant la vision photopique pour un champ visuel de 2 °, a été publiée par la CIE en 1924 et est connue sous le nom de $V(\lambda)$. Initialement proposée par Gibson et Tyndall (1923), cette fonction, bien qu'elle soit en pratique encore largement utilisée de nos jours, présente quelques imprécisions (Sharpe et al., 2005). Suite aux travaux de Judd (1951) puis de Vos (1978), une nouvelle fonction, $V_M(\lambda)$ fût proposée. En 2005, Sharpe et al. proposèrent également une autre variante de la fonction d'efficacité lumineuse, $V^*(\lambda)$ (Sharpe et al., 2005). Les revues de Wyszecki et Stiles (1982) ou de Boynton (1996) présentent les différentes méthodes expérimentales ayant historiquement été utilisées pour définir ces fonctions d'efficacité lumineuse.

Ces courbes ont une importance toute particulière dans la vision des couleurs car elles sont souvent utilisées pour définir la luminance, analogue psychologique de la radiance. On désigne par luminance l'intensité de la sensation visuelle associée à un rayonnement électromagnétique. Elle est généralement exprimée en candela par mètre carré ($cd.m^{-2}$). Ce n'est qu'à la fin du XX^e siècle qu'a été découvert le substrat physiologique de ces fonctions d'efficacité lumineuse. En caractérisant les réponses des cellules ganglionnaires de la rétine de primates et de leurs relais au niveau du LGN, les physiologistes observèrent un groupe de cellules dont le spectre de sensibilité était très proche de celui de ces fonctions d'efficacité lumineuse (Lee, Martin et Valberg, 1988). Ces cellules sont celles de la voie magnocellulaire, et constituent la base physiologique de la luminance. Nous les décrirons un peu plus loin. Une présentation plus complète du concept de luminance et de son support physiologique peut être trouvée dans la revue de Lennie, Pokorny et al. (1993).

Mesurer la sensibilité spectrale des cônes

La sensibilité spectrale de chacun des trois types de cônes a été mesurée, en termes de courbes d'absorption de photons, par une multitude de techniques. Historiquement, ces courbes de sensibilité spectrale ont d'abord été obtenues grâce à des méthodes psychophysiques mais des avancées technologiques ont ensuite permis de vérifier ces résultats (bien qu'avec moins de précision) en utilisant des méthodes plus directes comme la microspectrophotométrie (Dartnall et al., 1983), la densitométrie rétinienne ou encore l'électrophysiologie par *patch-clamp* (Schnapf et al., 1987)

Les différentes méthodes psychophysiques utilisées pour déterminer les courbes de sensibi-

lité des cônes partagent quelques grands principes. La principale difficulté réside dans le fait que les courbes de sensibilité des trois types de cônes (L et M notamment) sont très proches et se recouvrent partiellement, rendant ainsi impossible d'isoler sélectivement l'activité d'un type de cône. La méthode utilisée a donc consisté à faire ajuster à des observateurs deux lumières pour obtenir la même sensation, c'est-à-dire de sorte à ce que l'activité enregistrée par les cônes L, M et S soient identiques dans les deux cas. On obtient de cette manière des classes d'équivalence à partir desquelles il est possible de reconstruire les courbes de sensibilité de chacun des types de cônes.

Pour toute lumière monochromatique de référence, un individu trichromate possédant une vision des couleurs normale est capable d'ajuster une combinaison de trois lumières primaires de sorte à produire une sensation similaire. On désigne ce type d'expérience sous le nom d'expérience de correspondance des couleurs (Fig.1.10). Pour un groupe de trois primaires données (par exemple, rouge, vert et bleu), on peut ainsi répéter l'opération d'ajustement en faisant varier la longueur d'onde de la lumière monochromatique de référence sur le spectre visible et obtenir, en fonction de la longueur d'onde, l'intensité nécessaire de chacune des primaires pour qu'il y ait correspondance des couleurs. Stiles et Burch (1955) ont, par exemple, utilisé les primaires rouge (645 nm), verte (526 nm) et bleue (444 nm) à partir desquelles ils ont obtenus trois courbes, $r(\lambda)$, $g(\lambda)$ et $b(\lambda)$, nommées fonctions de correspondance des couleurs. Pour obtenir les fonctions de correspondance des couleurs, qui relient les intensités relatives de chacune des trois primaires à la longueur d'onde de la lumière monochromatique de référence à partir d'un tel dispositif expérimental, quelques conditions théoriques doivent être remplies (Wandell, 1995) :

- **La linéarité.** Si une combinaison de lumières primaires $\mathbf{e}(e_1, e_2, e_3)$ correspond à une lumière monochromatique de référence \mathbf{t} et qu'une autre combinaison $\mathbf{e}'(e'_1, e'_2, e'_3)$ correspond à une autre lumière monochromatique de référence \mathbf{t}' , alors $\mathbf{e} + \mathbf{e}'$ correspond à $\mathbf{t} + \mathbf{t}'$ (loi d'additivité de Grassmann (1854)).
- **L'indépendance.** Les trois lumières primaires utilisées doivent être indépendantes, c'est-à-dire qu'aucune ne peut correspondre visuellement à un mélange des deux autres.
- **La négativité.** Pour une lumière de référence donnée, une primaire peut prendre une valeur négative, ce qui correspond à ajouter une certaine proportion de cette primaire à la lumière de référence et faire correspondre à la lumière ainsi obtenue un mélange des deux autres primaires (formellement, ceci revient à $\mathbf{t} = e_1 + e_2 - e_3$ est identique à $\mathbf{t} + e_3 = e_1 + e_2$)

Il existe ainsi de nombreuses combinaisons de couleurs primaires permettant de produire des fonctions de correspondance des couleurs, chacune pouvant être décrite comme une combinaison linéaire des fonctions $r(\lambda)$, $g(\lambda)$ et $b(\lambda)$. En 1931, la CIE unifia celles-ci pour produire les fonctions de correspondance des couleurs standard, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ et $\bar{z}(\lambda)$ (Fig. 1.11). Ces fonctions sont adaptées pour de nombreuses applications et sont devenues des standards (pour les applications industrielles notamment) car elles présentent l'avantage d'être toujours positives et car $\bar{y}(\lambda)$ correspond à une autre courbe de référence de la CIE : la fonction d'efficacité lumineuse, $V(\lambda)$. L'inconvénient principal étant qu'il n'existe pas de groupe de primaires phy-

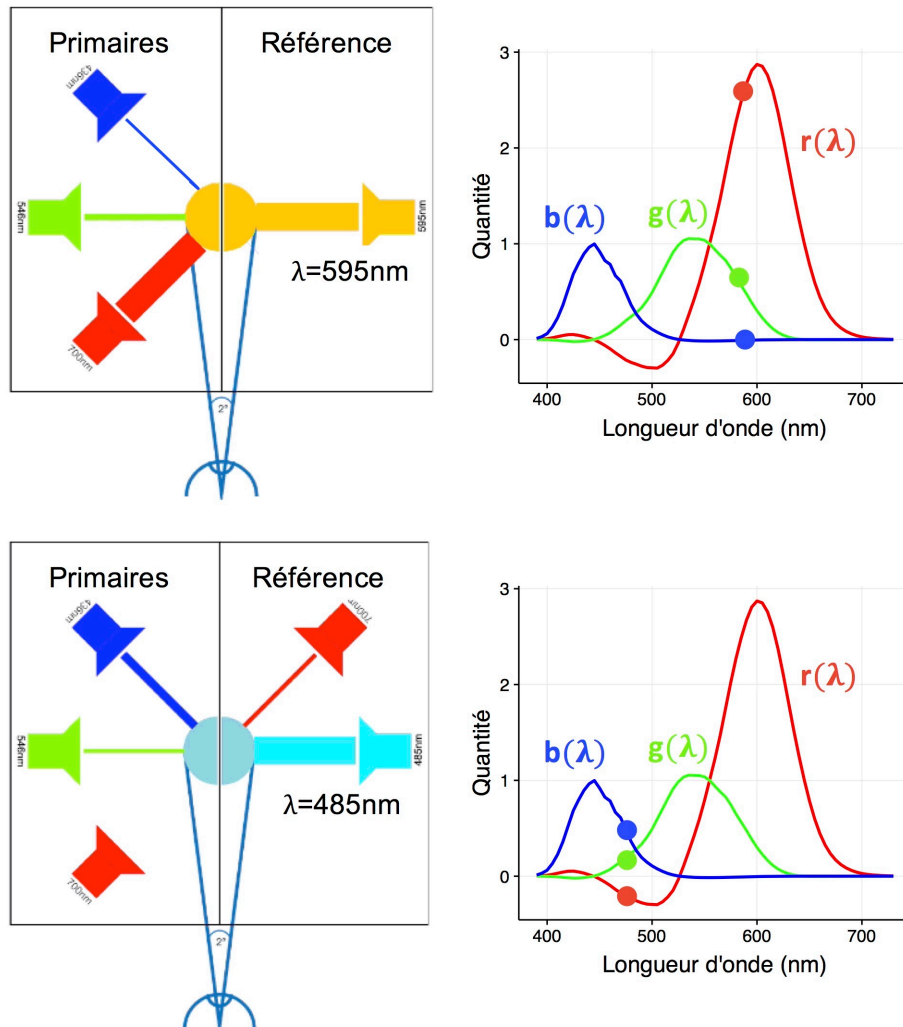


FIGURE 1.10 – Expérience de correspondance des couleurs. Trois primaires (rouge, verte et bleue) doivent être ajustées pour obtenir un résultat similaire à la lumière monochromatique de référence dont on fait varier la longueur d'onde (λ) sur l'ensemble du spectre visible. Dans le cas présenté en haut ($\lambda = 595 \text{ nm}$), un mélange des trois primaires permet d'obtenir une sensation colorée similaire à la lumière de référence. Dans le cas présenté en bas ($\lambda = 485 \text{ nm}$), il est nécessaire d'ajouter du rouge à la lumière de référence pour pouvoir obtenir une sensation identique en ajustant les primaires bleue et verte. On obtient finalement de cette façon les fonctions de correspondance des couleurs $r(\lambda)$, $g(\lambda)$ et $b(\lambda)$, qui prennent des valeurs négatives pour certaines valeurs de λ .

siquement plausibles permettant d'obtenir ces courbes, elles reposent donc sur trois primaires imaginaires.

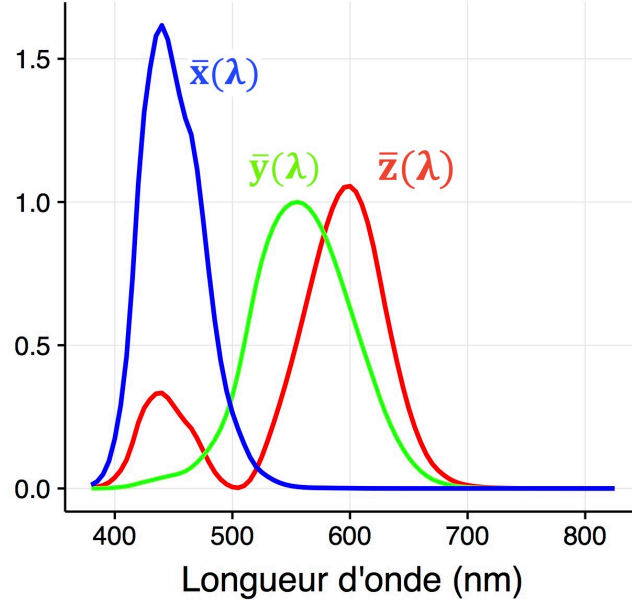


FIGURE 1.11 – Fonctions de correspondance des couleurs standard $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ et $\bar{z}(\lambda)$ proposées par la CIE en 1931.

Ces fonctions permettent cependant de représenter n'importe quel stimulus de distribution spectrale $\Phi(\lambda)$ dans un espace tri-dimensionnel XYZ où :

$$\begin{aligned} X &= k \int_{\lambda} \Phi(\lambda) \bar{x}(\lambda) d\lambda \\ Y &= k \int_{\lambda} \Phi(\lambda) \bar{y}(\lambda) d\lambda \\ Z &= k \int_{\lambda} \Phi(\lambda) \bar{z}(\lambda) d\lambda \end{aligned} \quad (1.1)$$

avec k une constante de normalisation.

Lorsqu'on s'intéresse aux mécanismes sous-tendant la perception des couleurs, il est nécessaire de pouvoir représenter les stimuli dans un espace physiologiquement pertinent, comme en termes d'excitations des cônes L , M et S plutôt que dans l'espace XYZ . Les fonctions de sensibilités spectrales des cônes ($l(\lambda)$, $m(\lambda)$ et $s(\lambda)$) peuvent être vues comme les fonctions de correspondance des couleurs associées aux trois primaires (L , M et S) qui permettraient de stimuler indépendamment et exclusivement chacun des trois types de cônes. Dans le cadre de l'expérience de correspondance des couleurs, lorsque les deux stimuli produisent une sensation identique nous avons les relations suivantes :

$$\begin{cases} l(\lambda) = l_R r(\lambda) + l_G g(\lambda) + l_B b(\lambda) \\ m(\lambda) = m_R r(\lambda) + m_G g(\lambda) + m_B b(\lambda) \\ s(\lambda) = s_R r(\lambda) + s_G g(\lambda) + s_B b(\lambda) \end{cases}$$

où l_R , l_G et l_B sont respectivement les sensibilités des cônes de type L aux primaires R, G et B et similairement pour les cônes M et S. Il existe donc une matrice $\mathbf{T} = \begin{bmatrix} l_R & l_G & l_B \\ m_R & m_G & m_B \\ s_R & s_G & s_B \end{bmatrix}$

telle que
$$\begin{bmatrix} l(\lambda) \\ m(\lambda) \\ s(\lambda) \end{bmatrix} = \mathbf{T} \begin{bmatrix} r(\lambda) \\ g(\lambda) \\ b(\lambda) \end{bmatrix}.$$

Sous certaines hypothèses, il est alors possible de résoudre cette équation et de déterminer les 9 valeurs de la matrice \mathbf{T} permettant d'obtenir les fonctions $l(\lambda)$, $m(\lambda)$ et $s(\lambda)$. Traditionnellement, on suppose que les individus dichromates (protanopes, deuteranopes ou tritanopes), c'est-à-dire, chez qui un type de cône est absent (respectivement, les types L, M ou S), ont une sensibilité identique aux individus trichromates pour les deux types de cônes dont ils disposent. On parle d'hypothèse d'Helmholtz-König (Young, 1807 ; Von Helmholtz, 1867 ; König et Dieterich, 1892). Sous cette hypothèse, et quelques autres suppositions, Smith et Pokorny (1975) ont ainsi pu déterminer psychophysiquement les fonctions fondamentales d'absorption des cônes (Fig. 1.12), depuis largement utilisée dans la littérature concernant la vision des couleurs. Plusieurs autres fonctions fondamentales ont par la suite été proposées, la CIE recommande actuellement l'utilisation des fonctions fondamentales de Stockman et Sharpe (2000) (Viénot, 2016 ; CIE, 2006). Une description plus détaillée des fonctions de sensibilité des cônes et des expériences de correspondance des couleurs ayant permis de les établir peut être trouvée dans le Chapitre 2 du livre *Color Vision : From Genes to Perception* (Stockman et Sharpe, 1999).

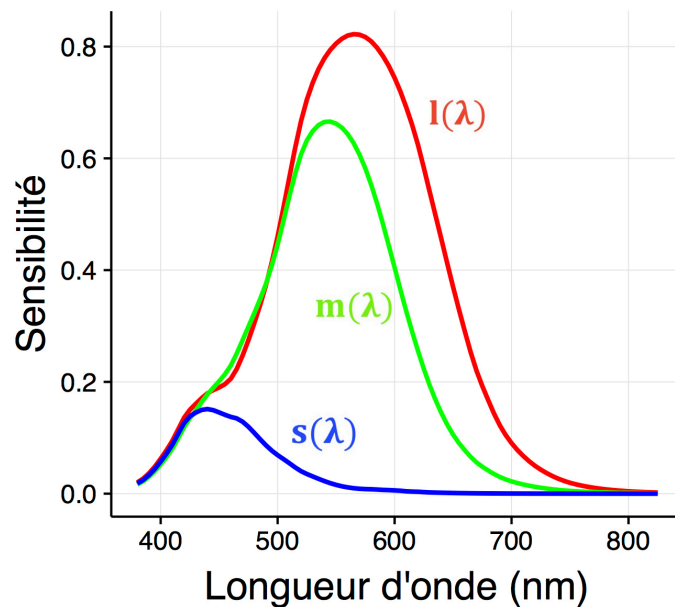


FIGURE 1.12 – Sensibilités spectrales des trois types de cônes (L, M et S) estimées par Smith et Pokorny, 1975

1.2.3.3 De la rétine vers le LGN

Les couleurs ne peuvent être perçues que grâce au traitement post-récepteurs, qui va permettre de comparer les excitations relatives de chacun des trois types de cônes. La rétine peut être vue comme une grille de récepteurs qui va ensuite transmettre l'information lumineuse, convertie en signal neuronal, vers le cortex visuel, en passant par le LGN. On distingue dans la rétine deux couches de neurones post-récepteurs, constituées chacune de deux types de cellules (Fig. 1.6). La première couche, connectée aux photorécepteurs, se compose des cellules horizontales et des cellules bipolaires. La seconde rassemble les cellules amacrines et les cellules ganglionnaires. À l'issue d'une série d'opérations réalisées par ces différentes cellules, le signal résultant est transmis via les axones des cellules ganglionnaires dans le nerf optique vers le LGN. Le détail des étapes du traitement rétinien permettant de transformer le signal initialement enregistré par les cônes en information spatio-chromatique plus complexe est présenté dans Silveira et al. (2005), Lee (2004) ou encore le chapitre 2 de l'ouvrage *Human Color Vision* (Kremers, Silveira et al., 2016).

1.2.3.4 Le LGN

En sortie de la rétine le traitement de l'information d'excitation des cônes va s'organiser en trois canaux, dont l'existence a pu être à la fois démontrée par des enregistrements en électrophysiologie (De Valois, Abramov et al., 1966 ; Lee, Martin et Valberg, 1988 ; Derrington et al., 1984), inférée d'expériences de psychophysique (Krauskopf, Williams et al., 1982) et prédite par des approches purement computationnelles, basées sur la théorie de l'information (voir Chapitre 2). Ces canaux post-récepteurs sont qualifiés d'opposition des cônes (ou opposition des couleurs) car ils sont sensibles à différentes combinaisons des réponses des cônes.

Électrophysiologiquement, cette organisation en trois canaux distincts a pu être observée dès les cellules ganglionnaires (De Monasterio et al., 1975), ainsi que pour leurs projections au niveau du LGN (Wiesel et Hubel, 1966 ; Lee, Valberg et al., 1987). Les premiers enregistrements de cellules d'opposition de cônes ont été réalisés par Russel De Valois et ses collaborateurs (De Valois, Smith et al., 1958 ; De Valois, Jacobs et al., 1964 ; De Valois, Abramov et al., 1966), suivis de près par Thorsten Wiesel et David Hubel (Wiesel et Hubel, 1966) et l'équipe de Peter Gouras (Gouras, 1968). Le revue de Lee (2014) synthétise, d'une perspective historique, les travaux d'électrophysiologie sur l'encodage de l'information chromatique par le système visuel des primates. Ces études ont permis de mettre en évidence l'existence de trois groupes de cellules ganglionnaires, correspondant à trois voies de traitement indépendantes, projetant dans des régions anatomiquement distinctes du LGN.

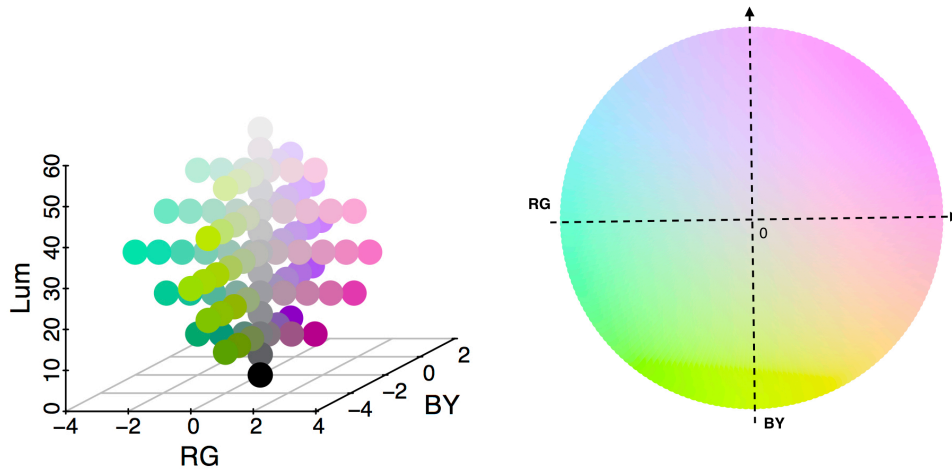
Le premier groupe somme le signal des cônes de types L et M, il s'agit du **canal de luminance** (ou **canal achromatique**), souvent noté dans la littérature $L + M$. Comme mentionné précédemment, la sensibilité des cellules de ce premier groupe correspond approximativement à la fonction d'efficacité lumineuse, $V(\lambda)$. Dans le second groupe, les signaux des cônes L et M sont soustraits les uns aux autres. Ce canal est désigné par $L - M$ et souvent, abusivement, par **canal Rouge-Vert** (ou **RG** pour *Red-Green*). Le dernier groupe de cellules compare les

réponses des cônes S à celles de la somme des réponses des cônes L et M. Il est noté $S - (L + M)$ et parfois désigné comme **canal Bleu-Jaune** (ou **BY** pour *Blue-Yellow*). Ces deux derniers canaux sont les **canaux de chrominance**. Contrairement au canal de luminance, qui porte une information sur l'intensité globale du stimulus, ces deux canaux comparent les réponses des différents types de cônes, ce qui va permettre d'accéder à l'information de couleur. Les notations RG et BY sont utilisées car elle correspondent très approximativement aux directions des réponses des neurones, mais elles sont abusives au sens où il n'y a pas lieu ici d'utiliser des noms de couleurs, la stimulation des neurones de l'un ou l'autre de ces canaux ne résultant pas nécessairement en une sensation colorée simple. Nous utiliserons cependant parfois ces notations dans la suite du manuscrit pour désigner l'information de chrominance portée par chacun de ces canaux, afin de simplifier les notations.

En psychophysique, plusieurs études ont permis de mettre en évidence, grâce à divers paradigmes, l'existence d'un canal de luminance et de deux canaux de chrominance dans le système visuel humain (Cole, Hine et al., 1993 ; Mullen et Losada, 1994 ; Bradley et al., 1988 ; Gegenfurtner et Kiper, 1992 ; Sankeralli et Mullen, 1997 ; Mullen et Sankeralli, 1999). Dans une étude particulièrement influente, Krauskopf et ses collègues (Krauskopf, Williams et al., 1982) ont montré l'existence de trois canaux distincts pour la vision des couleurs, qu'ils ont désignés comme "directions cardinales" de l'espace des couleurs. Ils ont montré qu'en habituant les participants selon l'une de ces directions cardinales, les seuils de détection dans cette même direction étaient ensuite modulés, mais qu'il n'y avait pas d'effet sur les seuils de détections mesurés dans les autres directions. Ces observations suggèrent donc l'existence de mécanismes ségrégués et opérant indépendamment selon chacune de ces directions. L'importance de ces trois canaux pour la vision a par la suite été montrée dans une série d'expériences resumées par Krauskopf, Gegenfurtner et al. (1999). Un espace tri-dimensionnel de représentation des couleurs dont les trois axes correspondent aux trois directions cardinales identifiées dans l'étude originelle de Krauskopf, Williams et al. (1982) a par la suite été défini, sous le nom d'espace DKL (pour Derrington, Lennie et Krauskopf) (Derrington et al., 1984) (Fig. 1.13). Largement utilisé dans la littérature sur la vision des couleurs, cet espace s'est révélé bien adapté pour décrire les premières étapes du traitement visuel (jusqu'au LGN) et expliquer de nombreuses données psychophysiques.

Au niveau du LGN, cette organisation selon trois canaux est maintenue, avec une structure en couches de trois types différents :

- Les couches **magnocellulaires**, qui reçoivent majoritairement les afférences du canal de luminance (Lee, Martin et Valberg, 1988 ; Kremers, Lee et al., 1992 ; Sun et al., 2006 ; Field et al., 2010).
- Les couches **parvocellulaires** où se projettent les neurones du canal de chrominance Rouge-Vert (Lee, Martin, Valberg et Kremers, 1993 ; Reid et Shapley, 1992 ; Reid et Shapley, 2002 ; Martin et al., 2001 ; Buzás et al., 2006 ; Lee, 2011).
- Les couches **koniocellulaires**, où est relayée l'information du canal de chrominance Bleu-Jaune (Dacey et Lee, 1994 ; Chichilnisky et Baylor, 1999 ; Tailby et al., 2008 ; Cheong et al., 2011).



(a) Vue tri-dimensionnelle de l'espace DKL. Les trois dimensions (Lum, RG et BY) correspondent aux trois directions cardinales de l'étude de Krauskopf, Williams et al. (1982).

(b) Représentation des couleurs dans un plan isoluminant (Lum=50) de l'espace DKL

FIGURE 1.13 – Espace tri-dimensionnel des couleurs DKL (Derrington et al., 1984)

Il y a donc trois canaux qui véhiculent l'information de couleur de la rétine vers le cortex visuel, distincts à la fois par leurs propriétés chromatiques et leurs substrats anatomiques. Bien que décrivant la ségrégation de cette information à un stade assez précoce du traitement visuel, les représentations des couleurs dans des espaces tri-dimensionnels d'opposition des cônes se sont avérées assez efficaces pour expliquer une grande partie des données psychophysiques, concernant notamment des tâches de détection et de discrimination de couleur. Il n'existe cependant pas de consensus sur la description formelle de ces canaux (pondération des différents types de cônes, intervention de divers processus d'adaptation,...) et plusieurs modèles ont été suggérés pour expliquer différentes données. Pour le canal de chrominance Rouge-Vert, par exemple, les combinaisons suivantes ont été proposées : $L - M$ (Tansley et Boynton, 1976), $0.96L - 1.28M$ (Guth et al., 1980; Ingling et Martinez-Uriegas, 1983), $L - 2M$ (Eskew et Boynton, 1987; Frome et al., 1981) ou encore $0.8L - M$ (Stromeyer, Kronauer et al., 1983). Une autre approche a été de considérer qu'en tenant compte d'une adaptation indépendante de chaque type de cône, alors la contribution des cônes L et M (exprimée, par conséquent, en termes de contraste de cône) à la chrominance Rouge-Vert était égale (Stromeyer, Cole et al., 1985). Bien que pertinentes dans le cas de stimuli simples pour lesquels un contraste peut être calculé par rapport au niveau d'adaptation d'un fond neutre, cette approche est difficilement extrapolable au cas de stimuli plus complexes. D'autres études ont, par ailleurs, fait l'hypothèse que le poids respectif des cônes L et M dans le canal Rouge-Vert dépendait directement de leur ratio dans la rétine de l'observateur (Gunther et Dobkins, 2002).

1.2.3.5 Le cortex visuel primaire (V1)

La majeure partie du signal de la rétine ayant transité via le LGN est ensuite projetée au niveau du cortex visuel primaire (V1). Le cortex visuel primaire, première étape du traitement visuel cortical, cartographie de façon complexe et largement distordue l'image rétinienne (Wandell et Winawer, 2011).

Au niveau de V1, l'encodage de l'information visuelle devient significativement plus complexe. Les neurones du LGN se combinent de diverses façons pour produire les réponses corticales. Contrairement à ce qu'il se passe au niveau du LGN, il semblerait que les réponses des neurones du cortex visuel primaire ne soient pas préférentiellement orientées selon trois directions cardinales. Ainsi, des travaux récents suggèrent qu'il n'y aurait pas vraiment de distinction entre luminance et chrominance au niveau cortical, mais plutôt un continuum de sensibilités couvrant les différentes directions de l'espace des couleurs (Kiper et al., 1997; Lennie, Krauskopf et al., 1990; Yoshioka et al., 1996; De Valois, Cottaris et al., 2000; Johnson et al., 2004; Solomon, Peirce et al., 2004; Horwitz et al., 2007). Certaines études suggèrent également que des directions autres que les directions cardinales identifiées dans le LGN seraient magnifiées dans V1 (Wachtler, Sejnowski et al., 2003).

Des études ont montré que la réponse de neurones corticaux pouvait être bien caractérisée par des combinaisons linéaires des signaux des cônes (De Valois, Cottaris et al., 2000; Lennie, Krauskopf et al., 1990; Conway, 2001; Conway, Hubel et al., 2002; Cottaris et De Valois, 1998; Conway et Livingstone, 2006). D'autres ont identifié des neurones aux propriétés de réponse plus complexes, pour lesquels le principe de linéarité semblait être violé (Hanazawa et al., 2000; Solomon, 2005; Horwitz et al., 2005). Il a cependant été suggéré que même pour ces populations, le modèle linéaire pouvait fournir une description informative sur le comportement des neurones (Solomon, 2005; Horwitz et al., 2005; Johnson et al., 2004). Il semble, malgré tout, il y avoir au niveau de V1, des neurones sensibles à la fois à des propriétés de luminance et de chrominance, et également à des propriétés spatiales (Shapley et Hawken, 2011; Rentzeperis et al., 2014) (Figures 1.15 et 1.16). Les résultats physiologiques concernant le traitement de l'information chromatique au niveau cortical sont discutés dans les revues de Schluppeck et Engel (2002) ou de Conway, Chatterjee et al. (2010).

1.2.4 Propriétés spatio-temporelles de la vision des couleurs

Aucune des dimensions de l'expérience visuelle ne peut être complètement appréhendée de façon isolée. La couleur apparente d'un stimulus n'est pas décorrélée de ses caractéristiques spatiales et temporelles et le traitement opéré par chacun des canaux post-récepteurs présente des caractéristiques spatio-temporelles spécifiques. Nous en présentons ici quelques unes.

1.2.4.1 Champs récepteurs

On parle de champ récepteur d'un neurone pour désigner la partie du champ visuel qui, lorsqu'on présente un stimulus lumineux en son sein, modifie la réponse de ce neurone. Cette définition peut être élargie au sous-ensemble des paramètres des stimulations modifiant son activité, comme c'est le cas par exemple des longueurs d'onde. Ce concept permet de mieux saisir le comportement des neurones de la rétine, comme les cellules ganglionnaires mais peut également être appliqué à des stades ultérieurs du traitement visuel.

Le champ récepteur d'une cellule ganglionnaire est de forme circulaire et possède un antagonisme centre-périphérie : le centre répond positivement à la stimulation d'un cône, tandis que la périphérie inhibe cette réponse en réagissant de façon négative à un stimulus lumineux. En plus de cette opposition spatiale, il existe aussi une opposition spectrale. La Figure 1.14 illustre le cas d'une opposition Rouge-Vert, où la partie centrale reçoit des entrées positives de cônes L et la partie périphérique des entrées inhibitrices de cônes M, et inversement.

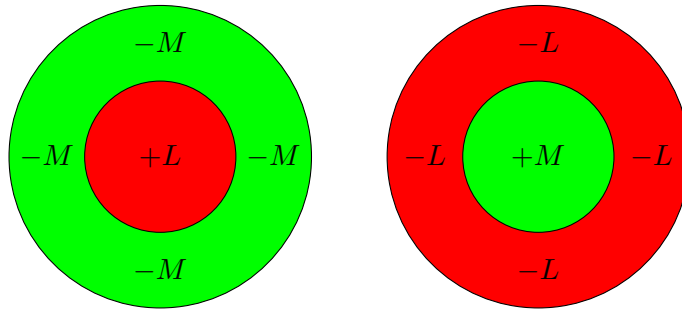


FIGURE 1.14 – Champs récepteurs de cellules ganglionnaires de la voie parvocellulaire, présentant une opposition spatiale et spectrale de type centre-périphérie pour les cônes de types L et M.

Des cellules présentant de tels champs récepteurs ont également été observées physiologiquement chez le macaque au niveau du LGN. Les champs récepteurs des cellules de la voie magnocellulaire sont plus larges que ceux de la voie parvocellulaire.

Au niveau du cortex visuel primaire, on sait depuis les travaux pionniers d'Hubel et Wiesel (Hubel et Wiesel, 1959; Hubel et Wiesel, 1962; Hubel et Wiesel, 1968a) que les neurones répondent fortement à des contours ou des barres de contraste. Les réponses des neurones du LGN semblent se combiner de différentes façons pour former des neurones corticaux de champs récepteurs variés, dont certains présentant une orientation spatiale (Fig. 1.15). Des études plus récentes, ont montré l'existence de neurones présentant des champs récepteurs de type double-opposition, sensibles à des contours de couleurs (Johnson et al., 2008) (Fig. 1.16).

1.2.4.2 Sensibilité au contraste

Notre capacité à discriminer les objets et les détails plus fins présents dans notre environnement repose sur les contrastes, c'est-à-dire les variations de luminance ou de chrominance.

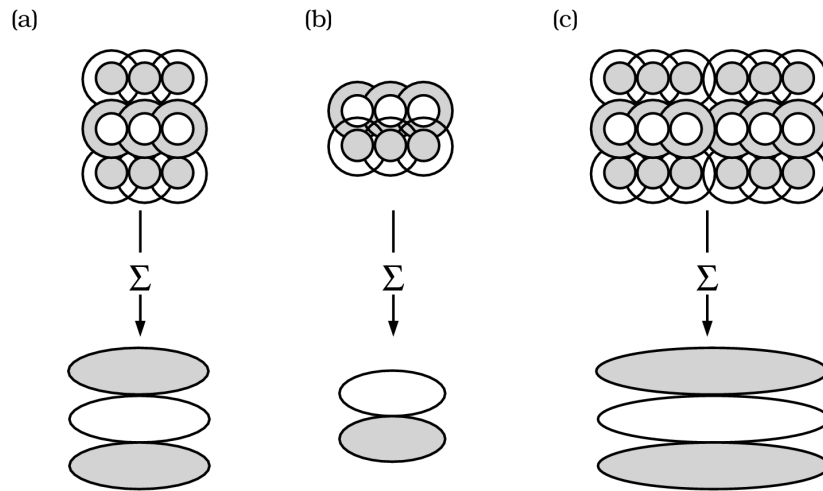


FIGURE 1.15 – Champs récepteurs orientés de V1 obtenus à partir de combinaisons des réponses de neurones de champs récepteurs circulaires trouvés dans la rétine et dans le LGN (d'après Wandell, 1995).

La sensibilité du système visuel aux fréquences spatiales de luminance ou de chrominance n'est pas uniforme et est caractérisée par les fonctions de sensibilité au contraste (FSC). Celles-ci ont été mesurées chez l'humain pour chacun des trois canaux de façon indépendante (Mullen, 1985). La méthode utilisée consiste à présenter des réseaux spatiaux (i.e. une alternance de bandes) selon l'une des directions cardinales de l'espace des couleurs, et à mesurer les seuils de détection en fonction de la fréquence. Les stimuli utilisés, ainsi que l'allure des courbes de sensibilité obtenues sont présentés sur la Figure 1.17. Dans le domaine des fréquences spatiales, les canaux de chrominance sont plutôt de type passe-bas tandis que le canal de luminance a une réponse passe-bande.

1.2.4.3 Adaptation

L'adaptation désigne la capacité d'un organisme à modifier sa sensibilité en fonction des conditions de stimulation. Ce concept général s'applique à tous les domaines de la perception et peut prendre de nombreuses formes. Dans le cadre de la perception visuelle, la gamme des stimuli rencontrés varie considérablement, à la fois dans le temps et dans l'espace. On dénombre en conséquence une multitude de mécanismes d'adaptation, allant de réactions presque réflexes (dilatation de la pupille,..) à des processus purement cognitifs (qui dépendent, par exemple, de connaissances préalables sur les objets et les illuminations dans leur environnement (Fairchild, 1992 ; Fairchild, 1993)). Ainsi, des mécanismes d'adaptation peuvent intervenir sous différentes formes et à différents stades du traitement visuel (voir la revue de Webster (2015) pour une discussion sur l'adaptation visuelle en générale et celle de Fairchild (2013) pour l'adaptation chromatique en particulier).

Concernant la vision des couleurs, ces phénomènes d'adaptation jouent un rôle particulièrement important. Comme mentionné précédemment, la seule information à laquelle nous avons

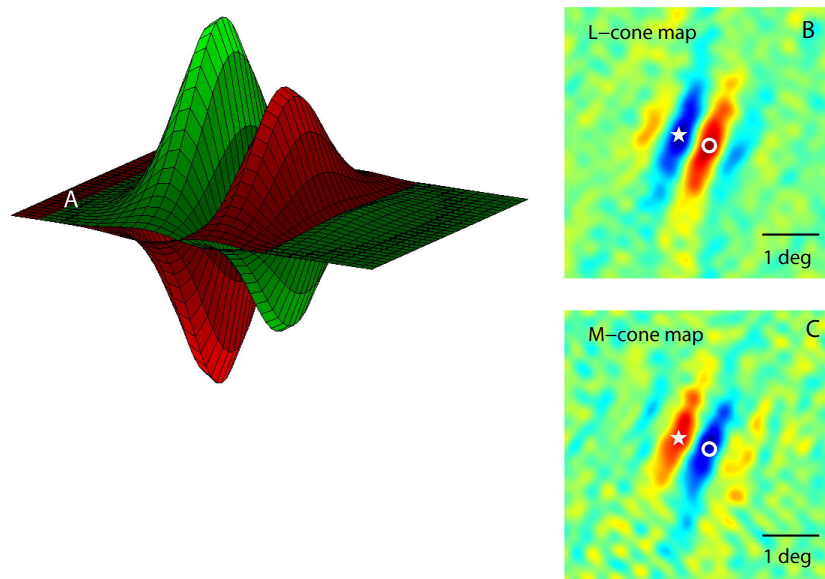
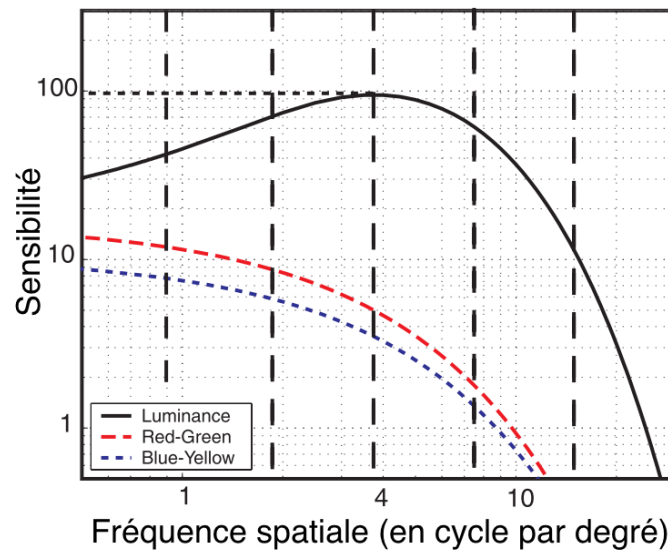


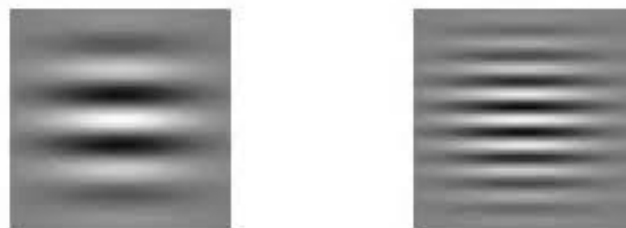
FIGURE 1.16 – Champ récepteur de type double-opposition observé pour des neurones de V1 (d’après Johnson et al., 2008). A : Représentation schématique d’un champ récepteur présentant un antagonisme spatial de type bord à bord avec des pondérations de cônes opposées. La pondération au-dessus du plan horizontal est positive et un accroissement de la lumière provoquera une augmentation de la réponse, alors que la pondération au-dessous du plan horizontal est négative et entraînera une diminution de la réponse. B : carte spatiale bi-dimensionnelle de sensibilité à l’excitation de cônes de type L pour un tel neurone de V1. C : Sensibilité à l’excitation des cônes de type M. À l’emplacement étoilé en B, la carte du cône L est inhibitrice, alors qu’en C, au même emplacement, la carte du cône M est excitatrice, et vice versa pour l’emplacement marqué par les cercles ouverts. Le schéma A est une représentation tri-dimensionnelle de la superposition des deux cartes des cônes L et M pour donner un profil général.



(a) Fonctions de sensibilité au contraste (FSC) des canaux de luminance et de chrominance



(b) Allure des stimuli (réseaux spatiaux) utilisés pour déterminer les FSC de chacun des canaux (respectivement, de gauche à droite : pour la canal de luminance, le canal Rouge-Vert et le canal Bleu-Jaune.).



(c) À gauche : Un réseau en basses fréquences spatiales. À droite : Un réseau en hautes fréquences spatiales

FIGURE 1.17 – Fonctions de sensibilité au contraste des canaux post-récepteurs et allure des stimuli utilisées pour la construction des courbes

accès pour déterminer l'apparence colorée d'un objet est le spectre résultant de l'interaction d'un illuminant avec cet objet. Afin d'apprécier la couleur d'un objet, nous devons donc estimer à partir du spectre lumineux qui nous parvient, les propriétés respectives de l'illuminant et du réfléchissant. Le système visuel humain va s'adapter aux changements d'illuminants afin de préserver une apparence stable de l'objet en dépit des variations d'éclairage. Une fraise par exemple, nous apparaîtra toujours rouge, qu'elle soit observée sous une lumière bleutée produite par des tubes fluorescents ou rougeoyante au moment d'un coucher de soleil (Fig. 1.18). On nomme ce phénomène *constance des couleurs*. Par le processus d'adaptation chromatique, le système visuel semble donc décompter l'effet de l'éclairage, pour pouvoir considérer la couleur comme un attribut stable de l'objet. Mais les traitements par lesquels cette opération est réalisée ne sont pas clairs et plusieurs modèles d'adaptation chromatique ont été développés.



FIGURE 1.18 – Illustration du phénomène de constance des couleurs : ces fraises paraissent rouges, alors qu'en réalité les pixels qui les constituent sont gris. Nous sommes si habitués à percevoir les fraises rouges que face à cette image nous supposons que la scène est éclairée par un illuminant bleuté et soustrayons inconsciemment celui-ci à chacun des pixels pour finalement percevoir les fraises telles que nous nous y attendons : rouges. (D'après Akiyoshi Kitaoka)

Le premier modèle d'adaptation chromatique fût proposé par Von Kries en 1902. Il s'agit d'un modèle théorique très simple qui, selon ses mots, peut se résumer de la façon suivante :

"this can be conceived in the sense that the individual components present in the organ of vision are completely independent of one another and each is fatigued or adapted exclusively according to its own function."

—Von Kries, 1902

Le modèle d'adaptation de Von Kries repose sur l'hypothèse que la constance des couleurs est obtenue en adaptant individuellement les gain des trois réponses des cônes, ces gains dépen-

dant du contexte sensoriel, c'est-à-dire de l'environnement et de l'historique. Formellement, ce modèle peut être exprimé par les équations suivantes :

$$\begin{aligned} L_a &= k_L L \\ M_a &= k_M M \\ S_a &= k_S S \end{aligned} \tag{1.2}$$

où L , M et S sont les réponses initiales des cônes, k_L , k_M et k_S les facteurs de gain respectifs de chacun des cônes et L_a , M_a et S_a les réponses post-adaptation.

Un aspect essentiel de ce modèle réside dans le choix des valeurs des gains k_L , k_M et k_S . En pratique, ces coefficients sont souvent fixés comme étant égaux à l'inverse des réponses respectives des cônes L, M et S pour la présentation d'un stimulus considéré comme neutre ou maximal. Ce choix implique donc de pouvoir définir un stimulus neutre ou maximal de référence pour estimer les réponses post-adaptation aux stimuli d'intérêt.

Bien que très simple, ce modèle s'est révélé adapté pour décrire certains phénomènes perceptifs, même si le fonctionnement de ces mécanismes semble en réalité plus complexe, comme Von Kries l'anticipait déjà lui même :

"But if the real physiological equipment is considered, on which the processes are based, it is permissible to doubt whether things are so simple."

—Von Kries, 1902

De nombreuses théories et modèles d'adaptation chromatique sensorielle ont été proposés à la suite des travaux de Von Kries, mais la question de savoir sous quelles formes et à quels stades du traitement de l'information visuelle interviennent précisément ces processus n'est toujours pas tranchée.

Nous avons, dans cette partie, présenté les résultats les plus importants concernant le traitement de la couleur par le système visuel humain. Certaines de ces propriétés ont pu être bien caractérisées par l'apport conjoint de travaux de psychophysique et de physiologie au cours des siècles derniers. Le traitement de la couleur par le système visuel demeure cependant en partie nébuleux, difficile à saisir dans sa globalité et d'une complexité qui semble ne cesser de croître au cours des différentes étapes. Malgré toutes les avancées dans la compréhension de la vision des couleurs qu'ont permises ces études, une question importante se pose : dans quelle mesure ces résultats, souvent obtenus à partir de stimuli artificiels, nous renseignent-ils sur le comportement du système visuel dans des conditions naturelles ?

1.2.5 Étudier la couleur dans les scènes naturelles

La plupart des études présentées jusqu'ici ont été motivées par une conception modulaire du traitement visuel, c'est-à-dire l'idée que les différents attributs de notre environnement visuel (couleur, forme, mouvement) étaient traités par des canaux indépendants (Livingstone

et Hubel, 1987; Zeki, 1978). Beaucoup des connaissances sur le fonctionnement du système visuel et la perception des couleurs ont été acquises grâce à l'utilisation de stimuli simples, tels que des *patches* colorés ou des réseaux spatiaux, permettant ainsi de modifier sélectivement l'activité d'un type de cône ou d'un canal. Mais ces stimuli ne sont pas représentatifs de notre environnement visuel naturel : les scènes naturelles sont bien plus complexes et présentent certaines spécificités que ne permettent pas d'appréhender les stimuli simples. Les gammes de contrastes présentes dans les scènes naturelles, par exemple, sont bien moindres que celles des stimuli utilisés en psychophysique ou en électrophysiologie. En conséquence, les réponses de neurones à la stimulation créée par des scènes complexes ne peuvent pas directement être prédites à partir de résultats obtenus avec des stimuli simples (Brady et Field, 2000; Tadmor et Tolhurst, 2000). La question de l'extension des propriétés connues du traitement de la couleur à la perception de stimuli plus complexes, tels que les scènes naturelles, a été peu abordée et est loin d'être triviale.

Depuis quelques années, la recherche sur la vision des couleurs s'est donc en partie orientée vers l'étude du fonctionnement des canaux de luminance et chrominance dans le cas de stimuli plus complexes, questionnant notamment leur interaction entre eux et avec d'autres mécanismes, tels que le traitement des formes. Impossible, en effet, de dissocier ces différentes propriétés dans le cas de stimuli naturels qui présentent simultanément des contrastes de luminance, de chrominance et une structure spatiale. Plusieurs études soutiennent que le traitement de la couleur et de la forme sont inextricablement liés dans le cortex (Shapley et Hawken, 2011; Moutoussis, 2015).

L'accent a donc été mis sur l'analyse du traitement de l'information de luminance et de chrominance à partir de stimuli moins artificiels et plus représentatifs de notre environnement visuel naturel (Shevell et Kingdom, 2008). Des études se sont ainsi intéressées aux propriétés spatio-chromatiques des scènes naturelles (Yoonessi et al., 2008; Hansen et Gegenfurtner, 2009; Provenzi et al., 2016), à la contribution de l'information de chrominance dans des processus tels que la détection de contours dans des scènes naturelles (Hansen et Gegenfurtner, 2017) ou encore à la sensibilité à la distortion ou l'estompation des contenus respectifs de luminance ou de chrominance dans des scènes naturelles (Jennings, 2015; Jennings et Kingdom, 2017).

Des récentes revues sur la vision des couleurs, soulignent également la nécessité de s'orienter vers l'étude des mécanismes de traitement de la couleur dans des conditions plus naturelles afin de pouvoir pleinement comprendre les phénomènes perceptifs (Witzel et Gegenfurtner, 2018; Foster, 2018).

1.3 Objectifs de la thèse et approche théorique

La question principale qui a guidé ces travaux de thèse est la suivante : à quoi nous sert-il de voir les couleurs ? Plus précisément, nous nous sommes attachés à explorer, à partir de scènes naturelles, quelle information pouvait apporter la chrominance qui n'était pas déjà présente dans la luminance, et comment cette information pouvait être exploitée par le système visuel.

Nous présentons ici le cadre théorique général selon lequel nous avons abordé cette question.

L'approche métathéorique que nous avons adoptée est centrée sur le traitement de l'information, c'est-à-dire que nous considérons la perception visuelle comme une chaîne de traitements permettant à partir de propriétés physiques, de parvenir à une perception. Cette approche est largement dominante dans les sciences de la vision actuellement, et a notamment été formalisée dans un livre influent de David Marr paru en 1982 (Marr, 1982) dans lequel il distingue trois niveaux de description nécessaires pour une compréhension complète d'un système complexe de traitement d'information : computationnel, algorithmique et implémentatif. Au delà de la perception visuelle, l'approche computationnelle de l'esprit humain a profondément impacté les sciences cognitives en général : le paradigme du traitement de l'information a été utilisé avec succès pour l'étude de phénomènes aussi variés que la perception auditive, la mémoire ou encore la résolution de problèmes.

1.3.1 L'approche mécanistique

Comme nous avons pu le souligner dans la section précédente, le traitement réalisé par le système neuronal de la vision des couleurs est très complexe, et chercher à saisir le comportement individuel de chacun des neurones impliqués dans la perception des couleurs est une tâche laborieuse et quelque peu utopique. Des études suggèrent même l'existence d'un certain nombre de processus aléatoires et stochastiques dans les réponses et les combinaisons neuronales de la voie de traitement de la couleur (Dacey, 2000 ; Rowe, 2002). Il paraît donc sensé d'approcher la question de la perception non pas en considérant la structure physiologique du système visuel, mais par une analyse macroscopique des phénomènes psychophysiques. Pour cette thèse nous adoptons donc une approche dite *mécanistique*. L'approche mécanistique s'intéresse à la nature des différentes étapes nécessaires pour passer d'une propriété physique à une perception. Dans le cadre de la vision des couleurs, il s'agit de décrire les étapes du traitement visuel en termes d'une série d'opérations sur les réponses initiales des cônes (Eskew, 2009 ; Stockman et Brainard, 2010).

Les premières étapes du traitement (rétine et LGN), semblent pouvoir être assez bien caractérisées par des mécanismes simples basés sur des combinaisons linéaires des réponses des cônes, comme celui schématisé sur la Figure 1.19, et de tels modèles ont pu expliquer efficacement des performances de détection et de discrimination pour des stimuli simples (Lennie et D'Zmura, 1988). Fournir une description mécanistique corticale pour la vision des couleurs semble cependant plus ardu. Il n'y a actuellement pas de consensus sur le nombre et la nature des mécanismes corticaux de la vision des couleurs (Hansen et Gegenfurtner, 2013 ; Gegenfurtner, 2003 ; Eskew, 2009 ; Stockman et Brainard, 2010).

Les principales questions auxquelles il est nécessaire de répondre pour proposer un modèle mécanistique de la vision des couleurs sont les suivantes :

- Combien d'étapes de traitement sont nécessaires ?
- Comment pondérer les différents types de cônes dans chacun des mécanismes ?
- Les mécanismes de luminance et de chrominance peuvent-ils être décrits par de simples

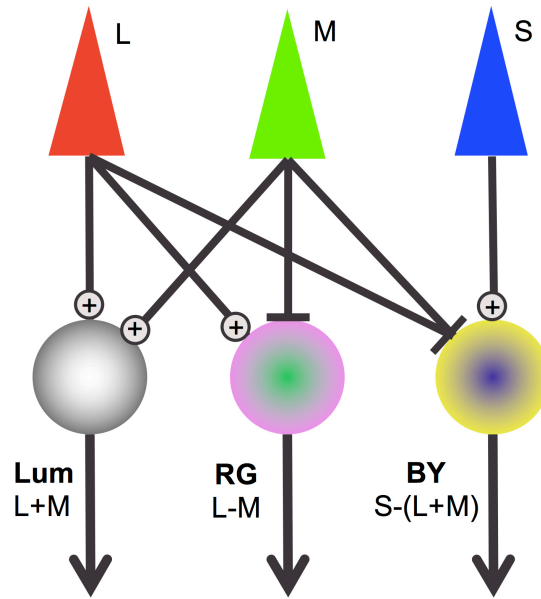


FIGURE 1.19 – Modèle mécanistique simple des premières étapes (rétine et LGN) du traitement de la couleur par le système visuel

combinaisons linéaires des réponses des cônes ou est-il nécessaire de considérer des opérations plus complexes ?

- Doit-on intégrer des processus d'adaptation ? Si oui, sous quelle forme et à quelle(s) étape(s) ?
- Existe-t-il des interactions entre les différents mécanismes et les différentes étapes du traitement ?

L'ambition de l'approche mécanistique est de parvenir à fournir, à partir de mesures comportementales, un modèle parsimonieux du traitement de l'information chromatique par le système visuel. Dans le cas théorique idéal où le modèle mécanistique décrirait parfaitement le comportement individuel de chacun des neurones des voies visuelles, ce modèle pourrait prédire parfaitement les performances d'observateurs. En pratique, étant donnée la complexité intrinsèque du système neuronal, il s'agit plutôt ici de fournir une description statistique correspondant à la réponse préférentielle des groupes de neurones constituant les différents canaux de traitement à chacune des étapes. Nous supposons qu'il est malgré tout possible par cette approche de formuler une description simple des informations sensorielles de luminance et de chrominance extraites par le système visuel, justifiant de données psychophysiques acquises à partir de scènes naturelles.

Nous nous attacherons donc dans ces travaux à utiliser un tel modèle. La question de l'extension de la validité des modèles proposés sur la base de stimuli simples à des tâches de détection ou de discrimination dans des scènes naturelles a été très peu abordée (Párraga, Breilstaff et al., 1998 ; Lovell et al., 2006 ; To et al., 2010). Afin d'obtenir des descriptions des informations sensorielles de luminance et de chrominance appropriées nous testerons donc notamment l'efficacité de modèles mécanistiques simples, basés sur des propriétés connues

du système visuel, à expliquer des données psychophysiques de détection dans des scènes naturelles.

1.3.2 L'observateur bayésien

Lorsque nous percevons un stimulus, celui-ci est dans un premier temps encodé par le système sensoriel, puis, dans un second temps, décodé par notre cerveau afin d'interpréter le signal neuronal sous forme de percept. La perception peut donc être modélisée par un processus d'encodage-décodage. Ces deux étapes ont souvent été étudiées séparément et, au cours des dernières années, deux théories dominantes ont émergées pour modéliser chacune d'entre elles : le codage efficace et le décodage bayésien. Ces deux théories ont récemment été combinées pour définir le modèle d'un observateur bayésien (Wei et Stocker, 2015). Ce cadre théorique est particulièrement adapté pour étudier comment le système visuel humain parvient à s'en sortir avec la complexité et l'ambiguïté intrinsèque à son environnement, et comprendre la connexion entre les statistiques des scènes naturelles et les réponses neuronales.

1.3.2.1 L'approche écologique et le codage efficace de l'information

"Ask not what's inside your head, but what your head's inside of."

—Williams Mace, 1977

Cette citation de Williams Mace capture succinctement l'approche de la perception formulée par James J. Gibson (Gibson, 1979), fondateur de la théorie écologique. L'idée selon laquelle les statistiques des stimuli auxquels ils sont exposés dans notre environnement ont façonné l'évolution de nos systèmes sensoriels et jouent un rôle important dans la perception et la cognition a commencé à apparaître dès le milieu du XIX^e siècle, pour depuis ne cesser de prendre de l'ampleur (Barlow, 2001). Le développement de la théorie de l'information par Shannon (Shannon et Weaver, 1949) donna un essor particulier à ce paradigme. En introduisant des concepts tels que *la capacité d'un canal*, *l'information* ou *la redondance*, il influença de nombreux chercheurs étudiant la vision qui ont depuis proposé diverses théories basées sur le codage efficace de l'information présente dans notre environnement.

Attneave (1954) formula l'idée que la finalité de la perception visuelle était de représenter de façon efficace les signaux entrants. Barlow (1961), dans un contexte neurobiologique, fit lui l'hypothèse que le rôle des neurones des premières étapes des traitements sensoriels était de réduire la redondance statistique des signaux entrants. Ces deux théories, bien que formulées dans des contextes différents soutiennent la même idée : les stimuli physiques de notre environnement présentent une certaine redondance dont nos systèmes sensoriels ont su tirer profit pour encoder et transformer ces signaux de la façon la plus efficace possible. Ils suggèrent ainsi de considérer le cadre de la théorie de l'information pour relier les statistiques de l'environnement aux réponses neuronales, grâce au concept de *codage efficace*. De nombreuses variantes de cette hypothèse du codage efficace ont depuis été proposées par différents auteurs (voir Simoncelli et Olshausen (2001) pour une revue).

Une approche méthodologique basée sur cette hypothèse du codage efficace consiste à examiner les propriétés statistiques de l'environnement et montrer que les transformations prédites par optimisation statistique des quelques critères fournissent une bonne description des propriétés de réponse d'un groupe de neurones sensoriels. Comme souligné par Geisler (2007), l'analyse des statistiques des images naturelles constitue donc, à ce titre, un outil pertinent pour avancer dans la compréhension de la perception visuelle.

Dans ces travaux nous nous intéresserons aux statistiques des images de scènes naturelles afin de chercher à quantifier la redondance des informations de luminance et de chrominance. Pour comprendre le rôle que peut jouer la couleur dans la perception visuelle, nous nous interrogerons notamment sur l'information supplémentaire qu'elle apporte par rapport à la luminance dans un contexte naturel.

1.3.2.2 Décodage bayésien

Bien que la théorie écologique proposée par Gibson ait constitué une révolution dans l'approche de la perception et préfiguré une grande partie des travaux computationnels modernes, l'un de ses aspects a été particulièrement controversé : l'idée que la perception était un processus direct. Gibson supposait en effet que toute l'information visuelle de notre environnement était capturée par la rétine, déterminant ainsi entièrement le percept en résultant.

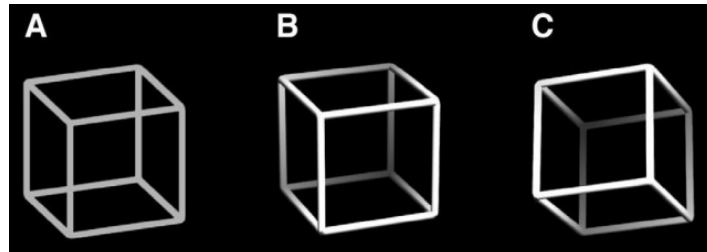
Pourtant, l'une des théories les plus influentes de l'histoire des sciences de la vision et, peut-être parmi les plus difficiles à saisir pleinement, est que la perception est un processus *actif* par lequel nous *construisons* une représentation de notre environnement. Le système visuel ne permet, en effet, d'extraire qu'une partie de l'information contenue dans notre environnement. Les entrées sensorielles sont, par définition, ambiguës : différents stimuli se projettent de la même façon sur la rétine (Fig. 1.20). Le système visuel doit donc résoudre un problème inverse qui consiste à sélectionner parmi une infinité de solutions possibles celle qu'il estime la plus probable d'avoir donné lieu à la réponse sensorielle. On parle d'*inférence* pour désigner ce processus de choix.

C'est Helmholtz qui a posé les bases de cette théorie, en suggérant dès 1867 (Von Helmholtz, 1867) que la vision résultait d'inférences inconscientes. C'est également lui qui formula l'hypothèse que ces inférences étaient faites en choisissant parmi toutes les possibilités celle qui était la plus probable, hypothèse connue sous le nom de principe de *vraisemblance*.

"The psychic activities that lead us to infer that there in front of us at a certain place there is a certain object of a certain character, are generally not conscious activities, but unconscious ones. In their result they are equivalent to a conclusion [...]. But what seems to differentiate them from a conclusion, in the ordinary sense of that word, is that a conclusion is an act of conscious thought."

—Helmholtz, 1867

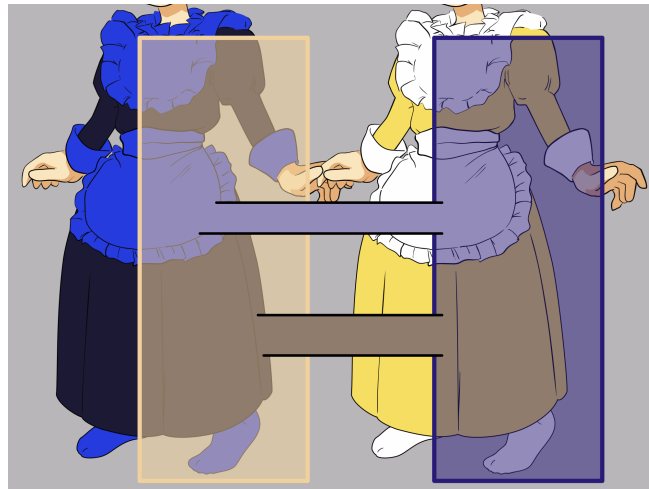
Dans le cadre de la perception visuelle nous pouvons formuler le modèle général suivant :



(a) Cube de Necker. Le stimulus A est ambigu et peut être interprété comme un cube orienté de deux façons différentes (B ou C)



(b) La Robe : photographie devenue virale sur Internet en 2015, lorsque les observateurs se trouvèrent en désaccord quant à savoir si les couleurs des rayures de la robe étaient noires et bleues ou blanches et or.



(c) L'information fournie par le contexte de la photographie n'est pas suffisante pour déterminer s'il s'agit d'une robe bleue et noire avec un illuminant jaune ou d'une robe blanche et or avec un illuminant bleu. Ce stimulus est par conséquent ambigu.

FIGURE 1.20 – Exemples de stimuli ambigus

- Les scènes naturelles contiennent un certain nombre de propriétés d'intérêt, que nous noterons \mathbf{S} .
- Nous disposons d'un modèle interne de la structure des scènes, qui définit, *a priori*, la probabilité des propriétés \mathbf{S} dans le monde extérieur, $p(\mathbf{S})$. Ce modèle peut, entre autres, résulter de l'accumulation de connaissances au cours des apprentissages.
- Le signal contenant les propriétés du monde extérieur (\mathbf{S}) est encodé par le système visuel. Nous modélisons cette transformation par une fonction de transformation interne, φ . Nous ajoutons au résultat de cette transformation un bruit tardif, que nous notons ϵ . La représentation finale des propriétés \mathbf{S} dans l'espace sensoriel de l'observateur est donc $I = \varphi(\mathbf{S}) + \epsilon$.

La perception consiste pour un observateur à estimer, compte tenu de l'information sensorielle I à laquelle il a accès, quelles sont les propriétés de la scène, \mathbf{S} . Autrement dit, il doit calculer la probabilité des propriétés \mathbf{S} sachant I , $p(\mathbf{S}|I)$.

Cette probabilité peut être calculée en appliquant la règle de Bayes :

$$p(\mathbf{S}|I) = \frac{p(I|\mathbf{S})p(\mathbf{S})}{p(I)} \quad (1.3)$$

Selon le formalisme bayésien :

- $p(\mathbf{S}|I)$ est appelée distribution (ou densité, ou probabilité) *a posteriori* et représente la connaissance que nous avons des propriétés \mathbf{S} avec l'observation de l'information sensorielle I
- $p(I|\mathbf{S})$ est la fonction de vraisemblance, elle quantifie à quel point certaines valeurs de \mathbf{S} sont compatibles avec l'observation I
- $p(\mathbf{S})$ est l'*a priori* et représente les connaissances préalables que nous avons des propriétés \mathbf{S}
- $p(I)$ est la probabilité marginale de I

Sur la base de la distribution *a posteriori* l'observateur peut alors générer une estimation, $\hat{\mathbf{S}}(I)$, qui correspond au percept associé au stimulus présentant les propriétés \mathbf{S} .

Ce modèle très général fournit un cadre théorique qui peut être appliqué à l'étude de nombreux phénomènes perceptifs (Knill et Richards, 1996 ; Kersten et al., 2004 ; Schrater et Kersten, 2002). Il a entre autres été utilisé pour modéliser la vision des couleurs (Brainard, 1997 ; Brainard et Gazzaniga, 2009), le processus de groupement de contours dans des scènes naturelles (Geisler et al., 2001) ou encore l'incertitude visuelle (Barthelmé et Mamassian, 2011) et le jugement de similarité (Van Den Berg et al., 2012). Des tutoriels d'implémentation de modèles bayésiens dans le cadre de la perception sont présentés par Vincent (2015) et Mamassian et al. (2002) et une discussion plus générale sur l'apport de ces modèles à l'étude de la cognition est proposée par Jones et Love (2011).

Le modèle de l'observateur bayésien est donc un cadre théorique d'observateur idéal qui encoderait le plus efficacement possible l'information disponible et prendrait, compte tenu de l'information sensorielle accessible et de ses connaissances préalables, la décision la plus

vraisemblable. Ce modèle est synthétisé sur la Figure 1.21. Pour nos travaux, nous nous référons à ce cadre théorique pour modéliser les étapes d'encodage de l'information et de prise de décision dans des tâches de détection ou de classification.

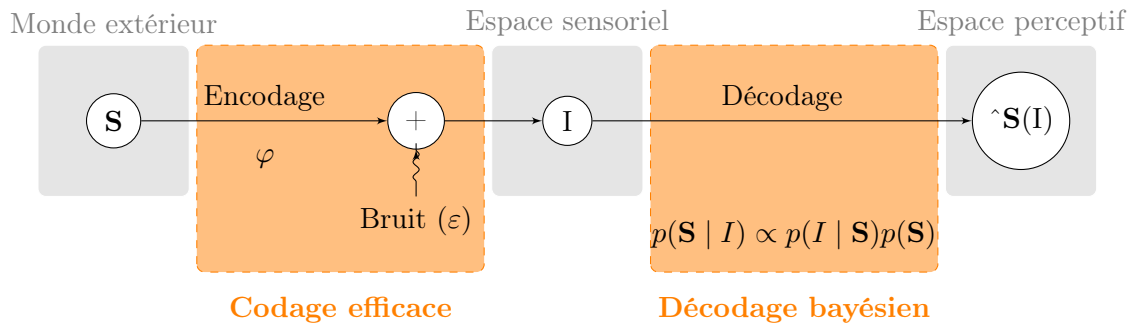


FIGURE 1.21 – Représentation schématique du modèle d'observateur bayésien. L'information visuelle du monde extérieur est dans un premier temps encodée dans l'espace sensoriel de l'observateur selon le principe de codage efficace. Dans un second temps, cette information est décodée par un processus de décodage bayésien pour aboutir à une perception.

La suite de ce manuscrit présente les travaux réalisés durant cette thèse, et l'organisation des chapitres retrace le cheminement de notre démarche.

Nous avons, dans un premier temps, cherché à mesurer l'indépendance de l'information de luminance et de l'information de chrominance dans des images de scènes naturelles (**Chapitre 2**). L'hypothèse qui a guidé cette analyse statistique était que plus l'information de chrominance serait différente de l'information de luminance, plus elle serait susceptible d'être exploitée par le système visuel. Nous avons donc établi une chaîne de traitements permettant, pour chaque image individuelle, de calculer un indice quantifiant l'indépendance de ces deux informations.

Cette première étude nous a fait apparaître la nécessité de fournir une description précise des informations de luminance et de chrominance sensorielles, adaptée au cas des scènes naturelles. Cette question ayant été très peu abordée dans la littérature, nous avons mené une série d'expériences psychophysiques de détection de couleur et de luminance dans des scènes naturelles et considéré le modèle théorique d'un observateur bayésien pour confronter l'efficacité de modèles mécanistiques simples à expliquer nos données expérimentales (**Chapitre 3**).

Nous avons ensuite cherché à évaluer le rôle de l'information de chrominance lors d'une étape spécifique de la perception des scènes naturelles : la classification des contours comme résultant de changements d'illuminant ou de variations des propriétés matérielles (**Chapitre 4**). À nouveau, nous avons proposé un paradigme expérimental original, reposant sur des stimuli extraits de scènes naturelles, et développé un modèle de classifieur, basé sur un observateur bayésien, pour justifier nos données empiriques. Nous avons ainsi pu mettre en évidence une contribution de l'information de chrominance pour cette tâche spécifique.

Enfin, dans un dernier chapitre nous synthétisons les résultats de ces travaux et présen-

tons les perspectives ainsi entrouvertent pour la compréhension du rôle de la couleur dans la perception visuelle des scènes naturelles (**Chapitre 5**).

2

Quantifier l'indépendance statistique des informations de luminance et de chrominance dans les scènes naturelles

How much of beauty — of color, as well as form— on which our eyes daily rest goes unperceived by us!

David Henry Thoreau, *Journal*, 1860

Sommaire

2.1	Analyses statistiques des propriétés spatio-chromatiques des scènes naturelles	42
2.2	Estimation de l'information mutuelle	46
2.2.1	Méthode	46
2.2.2	Résultats	57
2.2.3	Discussion	71
2.2.4	L'information mutuelle : un indice pertinent ?	74
2.3	Conclusion	76

Afin d'étudier le rôle de la couleur dans des processus cognitifs complexes, tels que la perception de scènes, nous formulons l'hypothèse suivante : plus l'information de chrominance sera indépendante de l'information de luminance, plus elle sera susceptible d'être exploitée par le système visuel. Nous proposons une méthode basée sur la théorie de l'information

permettant de quantifier, à l'échelle d'une image, la dépendance statistique entre information de luminance et information de chrominance.

Dans ce chapitre, nous présentons d'abord brièvement les principaux travaux existants sur l'analyse statistique des scènes naturelles dans le domaine de la vision des couleurs. Nous détaillons ensuite la chaîne de traitement, inspirée de l'article de Hansen et Gegenfurtner (2009), que nous avons utilisée pour mesurer la dépendance entre canaux de luminance et de chrominance à différents niveaux du traitement visuel. Enfin, nous évaluons la fiabilité et la robustesse de notre chaîne de traitement en testant l'influence des différents paramètres sur les mesures d'indépendance.

2.1 Analyses statistiques des propriétés spatio-chromatiques des scènes naturelles

L'approche écologique de la perception visuelle a conduit de nombreuses études à analyser les propriétés statistiques des scènes naturelles, afin de comprendre comment le système visuel pouvait encoder efficacement l'information présente dans notre environnement (Simoncelli et Olshausen, 2001). Ce cadre théorique a notamment été utilisé avec succès dans le domaine de la vision des couleurs, où il a permis de modéliser et d'analyser l'information qui pouvait être optimalement extraite de notre environnement grâce à la trichromatie et aux mécanismes post-récepteurs.

Une première approche a consisté, connaissant les propriétés d'absorption des trois types de cônes et l'existence des canaux post-récepteurs, à déterminer, à partir des statistiques des scènes naturelles, quels signaux pouvaient être encodés par ces canaux afin de réduire la redondance et maximiser l'information transmise. Une étude pionnière et particulièrement influente fût publiée par Buchsbaum et Gottschalk (1983). S'interrogeant sur l'intérêt et la nature des canaux d'opposition des cônes, ils réalisèrent une Analyse en Composantes Principales (ACP) sur les sensibilités spectrales des trois types de cônes afin de déterminer les transformations linéaires permettant de décorréler au mieux les signaux. Ils montrèrent ainsi que trois axes orthogonaux, correspondant à des combinaisons linéaires des réponses des cônes, permettaient de décorréler l'information initiale, et donc d'optimiser l'efficacité de sa transmission au système visuel. Plusieurs études ont par la suite étendu l'analyse proposée par Buchsbaum et Gottschalk à l'étude des statistiques des réponses des cônes à des images de scènes naturelles (Burton et Moorhead, 1987; Webster et Mollon, 1997; Ruderman et al., 1998) et corroboré leurs résultats.

Au delà des corrélations entre canaux à l'échelle du pixel, une autre information particulièrement intéressante est la façon dont ces valeurs évoluent d'un pixel à l'autre dans une image. Contrairement à des images de bruit aléatoires, les scènes naturelles présentent d'importantes redondances spatiales : deux pixels situés à proximité sont susceptibles de contenir une information assez similaire. L'information contenue par un pixel va donc en partie dépendre de celle de ses voisins, on parle alors de redondance de second-ordre. Ces relations existant entre

pixels distants vont définir des structures telles que des contours, des formes ou des textures, caractéristiques des scènes naturelles.

Un certain nombre d'études se sont attachées à définir les composantes spatio-chromatiques principales permettant de décorréler au mieux les signaux de réponse des cônes à des images de scènes naturelles. La méthode utilisée consistait à encoder les images sous forme de *patches* de dimensions $n \times n$ pixels puis à effectuer une ACP (Ruderman et al., 1998) ou une Analyse en Composantes Indépendantes (ACI) (Lee, Wachtler et al., 2002; Wachtler, Lee et al., 2001; Doi et al., 2003) pour déterminer les composantes spatio-chromatiques permettant de décorréler le plus efficacement les signaux. Toutes ces études confirmèrent la séparation de l'information chromatique en 3 canaux combinant linéairement les réponses des cônes dans 3 dimensions isolées de l'espace des couleurs : un canal de luminance dans la direction noir-blanc et deux canaux de chrominance dans les directions approximatives rose-vert et bleu/violet-jaune/chartreuse. Les structures spatiales des composantes obtenues diffèrent légèrement en fonction de la méthode utilisée. L'exemple des 27 composantes spatio-chromatiques principales obtenues par Ruderman et al. (1998) pour des *patches* de dimensions 3×3 pixels est présenté sur la Figure 2.1.

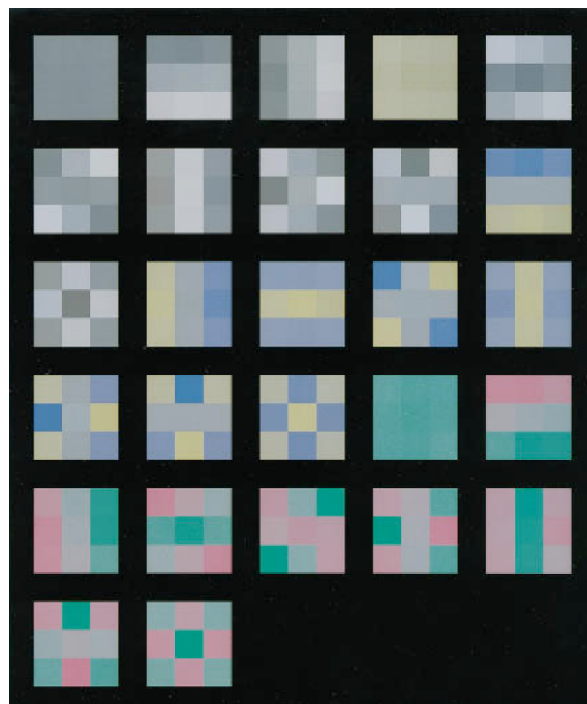


FIGURE 2.1 – Composantes spatio-chromatiques principales obtenues par Ruderman et al. (1998) selon l'ordre décroissant de leur pourcentage d'explication de la variance totale, de gauche à droite et de haut en bas.

Plus récemment, des analyses des propriétés statistiques des scènes naturelles ont été réalisées en vue de comparer l'information contenue dans les différents canaux post-récepteurs (le canal de luminance, Lum, et les deux canaux de chrominance, Rouge-Vert et Bleu-Jaune). Plusieurs études se sont notamment intéressées à la comparaison de l'information spatiale

contenue par chacun des canaux d'opposition des cônes. Párraga, Brelstaff et al. (1998) ont ainsi employé quatre définitions différentes des canaux de luminance et de chrominance Rouge-Vert, basées sur divers modèles, et analysé un set de 29 images de scènes naturelles en termes de fréquences spatiales. Ils n'ont pas observé de différence entre le contenu fréquentiel de leurs canaux. Ce résultat peut à première vue sembler en contradiction avec les prédictions de l'approche écologique : des données psychophysiques ont, en effet, montré que le système visuel humain était plus sensible aux basses fréquences en chrominance qu'en luminance (Mullen, 1985), on pourrait donc s'attendre à observer plus de basses fréquences dans le contenu de chrominance des scènes naturelles que dans le contenu de luminance. Une autre explication est que l'information contenue par les canaux de chrominance en hautes fréquences serait très redondante avec celle de luminance, et que le système visuel aurait donc intérêt à privilégier le traitement des basses fréquences spatiales de chrominance pour optimiser l'extraction de l'information.

Les auteurs expliquent, eux, en partie leur résultat par l'absence de considération de tâche spécifique pour l'analyse. Il s'agit effectivement ici d'une des limites des études sur les statistiques des scènes naturelles : peut-on totalement décorrélérer l'étude du fonctionnement du système visuel de la considération de tâches spécifiques et écologiquement pertinentes ? En supposant que le système visuel a évolué de façon à extraire optimalement l'information de son environnement, la finalité de ce traitement est la survie. Ainsi, il a été montré que les processus attentionnels ou la capacité à effectuer des tâches de détection ou de discrimination avaient façonné le traitement visuel. En réalisant des analyses sur les images de scènes naturelles, sans considérer de tâche spécifique, nous sommes donc limités à l'étude de corrélations à l'échelle des pixels, et ce contexte neutre nous fournit donc nécessairement une information plus restreinte et moins pertinente que lorsqu'on tient compte d'une tâche. Dans une autre étude reprenant une méthodologie comparable mais une base d'images différentes (Párraga, Troscianko et al., 2002), les auteurs ont par la suite montré que les propriétés spatio-chromatiques du canal Rouge-Vert semblaient optimisées pour certaines tâches particulières, telles que la détection d'objets rouges parmi des feuillages.

Johnson, Kingdom et al. (2005) s'intéressent également aux propriétés spatiales des scènes naturelles pour leurs contenus en luminance et en chrominance. Ils calculent les corrélations entre canaux en termes de pixels puis pour des ordres plus élevés, obtenus en appliquant différents filtres spatiaux sur chacun des canaux. Ils montrent que pour les données brutes (à l'échelle du pixel) la corrélation entre canaux est quasi-nulle mais pas pour les ordres plus élevés, pour lesquels ils observent des corrélations qui varient en fonction des fréquences spatiales. Leurs résultats suggèrent donc que la structure spatiale des scènes naturelles dans les différents canaux présente une certaine dépendance statistique. Ces corrélations peuvent en partie s'expliquer par le fait que, dans les scènes naturelles, les frontières des objets correspondent souvent à des changements simultanés de luminance et de chrominance. En regardant les cartes de contours d'une image pour chacun des trois canaux post-récepteurs, les contours de la plupart des objets devraient donc apparaître dans les trois.

L'extraction des contours constitue une étape critique pour la perception : c'est en effet à ces variations que notre système visuel est principalement sensible (Hubel et Wiesel, 1968b)

et cette information est indispensable pour le traitement des scènes complexes (Marr, 1982). Pourtant, peu d'études se sont intéressées à la dépendance statistique entre contours de luminance et contours de chrominance dans les scènes naturelles. L'une des raisons en est que la majorité des études portant sur les statistiques des scènes naturelles se sont uniquement intéressées aux propriétés spatiales achromatiques (Field, 1987; Field, 1994; Tolhurst et al., 1992) et qu'il est souvent considéré que la perception des formes repose majoritairement sur le traitement de l'information de luminance, l'information de chrominance étant souvent redondante. C'est dans ce sens que vont les résultats de Zhou et Mel (2008) qui ont estimé les contours pour chacun des canaux d'opposition des cônes sur un grand nombre d'images de scènes naturelles et montré que les contours des canaux de chrominance étaient souvent associés à des contours de luminance. Fine et al. (2003) ont analysé sur 12 images hyperspectrales les valeurs de luminance et de chrominance à différentes positions spatiales en fonction de la distance entre les pixels et observé une forte association, dont ils ont conclu que les contours de luminance et de chrominance n'étaient pas indépendants.

A l'inverse, Hansen et Gegenfurtner (2009) ont analysé les distributions jointes des contours de luminance et de chrominance sur une large base d'images de scènes naturelles et ont conclu à leur indépendance. Leur étude se base sur des calculs d'information mutuelle et fait apparaître que, dans les scènes naturelles, les contours purement chromatiques ne sont pas plus rares que les contours achromatiques et que la plupart des contours combinent à la fois des contrastes de luminance et de chrominance mais que la magnitude et le signe de ces changements sont indépendants entre les trois canaux post-récepteurs. Ils en concluent que les contours chromatiques constituent une source d'information indépendante des contours de luminance, qui peut donc tout aussi bien être exploitée par le système visuel pour la perception des scènes naturelles. Dans une autre étude, Peyvandi et al. (2013) ont également proposé un cadre théorique pour évaluer l'information contenue dans les contours de chacun des trois canaux post-récepteurs pour des images de scènes naturelles, basé sur des calculs d'entropie et d'information mutuelle. Cette approche semble donc pertinente pour mesurer l'information portée par la chrominance, et son indépendance avec l'information portée par la luminance.

Beaucoup d'études ont comparé les performances de participants dans des tâches visuelles pour des images couleur et des images en niveaux de gris, et des résultats parfois contradictoires ont été obtenus concernant le rôle de la couleur. Nous formulons l'hypothèse que si la chrominance apporte peu d'information, alors il n'est pas surprenant que lorsqu'elle soit retirée les performances des observateurs ne soient pas ou peu modifiées. D'une image à l'autre, l'information de chrominance varie et l'effet de la suppression de celle-ci sera donc susceptible d'être différent, et d'autant plus important que l'information de chrominance apportait une information indépendante de celle de luminance. Nous cherchons donc un moyen de mesurer cette information de chrominance pour une image donnée, et notamment l'indépendance de celle-ci avec l'information de luminance. La méthode proposée par Hansen et Gegenfurtner (2009) nous semble un bon point de départ et nous allons donc nous en inspirer pour construire une chaîne de traitement permettant de quantifier la redondance statistique entre l'information de luminance et l'information de chrominance dans des images de scènes naturelles. Notre objectif est de pouvoir ainsi sélectionner des images pour lesquelles l'information de chrominance est la plus indépendante possible de l'information de luminance, et serait donc d'autant

plus susceptible de modifier les performances d'observateurs lors de tâches spécifiques si elle est retirée des images. Nous avons ainsi pu établir un classement d'images en fonction de ce critère. Nous avons également identifié un certain nombre de défauts dans le traitement proposé que nous avons cherché à corriger.

2.2 Estimation de l'information mutuelle

Nous avons construit une chaîne de traitement imitant les différentes étapes du traitement visuel et mesuré à chacune des étapes la redondance de l'information portée par les différents canaux, via une estimation de l'information mutuelle. Ce processus est répété indépendamment sur un grand nombre d'images de scènes naturelles, nous permettant ainsi d'obtenir un classement des images en fonction de l'indépendance de leur contenu de chrominance avec celui de luminance.

2.2.1 Méthode

2.2.1.1 Base d'images

Nous avons sélectionné pour cette étude 1338 images issues de la base d'images couleur McGill (<http://tabby.vision.mcgill.ca>, Olmos et Kingdom, 2004a). Ces images, dites *calibrées*, ont été acquises avec des appareils dont les gammas et les sensibilités spectrales de chacun des capteurs R, G et B ont été précisément mesurés. Ces informations permettent notamment d'effectuer une correction gamma sur les images ou de les convertir du format RGB au format LMS correspondant à l'espace d'absorption des cônes L, M et S de la rétine humaine. Les images originales sont enregistrées au format .TIFF et ont 3 composantes RGB. Elles sont de dimensions 768×576 pixels et se répartissent en neuf catégories : Animaux, Fleurs, Forêts, Fruits, Paysages, Scènes artificielles, Ombres, Neige et Textures. La Figure 2.2 présente des exemples d'images appartenant à chacune des neuf catégories.

2.2.1.2 Chaîne de traitement de l'information visuelle

Étape 1 : Conversion dans l'espace d'absorption des cônes (rétine)

Les images RGB ont dans un premier temps été converties dans l'espace LMS, correspondant aux réponses respectives de chacun des types de cône L, M et S de la rétine d'un observateur humain. Afin d'effectuer cette conversion, il est nécessaire de tenir compte de deux types de paramètres :

- Les paramètres concernant le dispositif sur lequel est affichée ou a été enregistrée l'image (écran ou appareil photo).
- Les paramètres du système récepteur, c'est-à-dire, ici, les cônes (cellules sensorielles de notre système de perception visuelle).

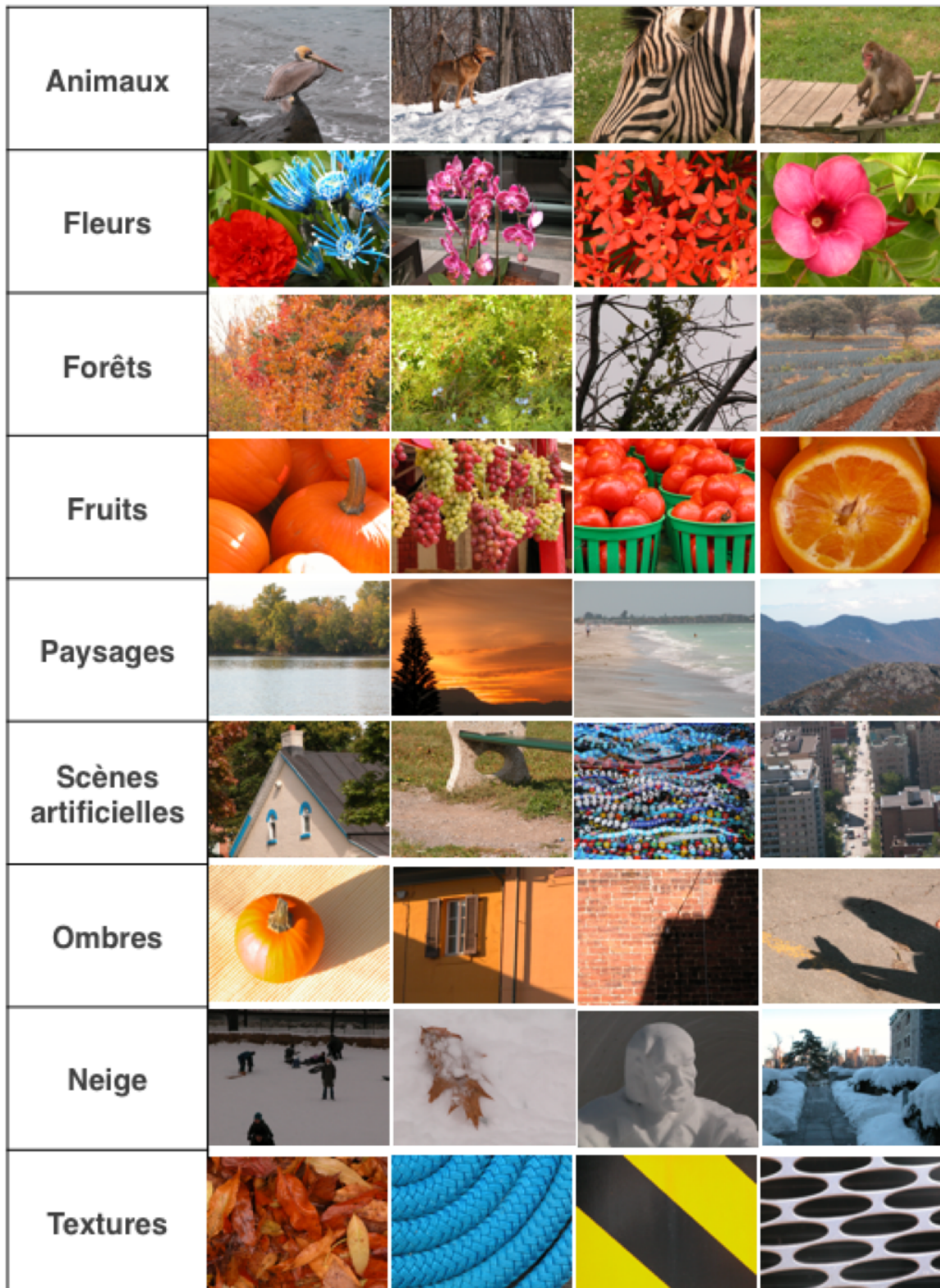


FIGURE 2.2 – Exemples d'images des neuf catégories (Animaux, Fleurs, Forêts, Fruits, Paysages, Scènes artificielles, Ombres, Neige et Textures) de la base calibrée McGill (<http://tabby.vision.mcgill.ca>).

Il existe ainsi deux façons de calculer des coordonnées LMS d'une image :

- On peut, grâce aux informations concernant les conditions d'enregistrement (durée d'exposition) et les dispositifs utilisés (gammas, sensibilités spectrales des trois primaires de l'appareil,...), obtenir des coordonnées LMS qui correspondraient à la réponse enregistrée par les cônes d'un observateur se trouvant dans les conditions d'observation réelles de la scène. Les auteurs de la base McGill fournissent une fonction Matlab *rgb2lms.m* qui permet d'effectuer directement cette conversion sur chacune des images.
- On peut calculer les coordonnées LMS associées à l'affichage des images RGB sur un écran dont on connaît les propriétés (gammas, sensibilités spectrales des trois primaires de l'appareil,...), c'est-à-dire la réponse des cônes d'une personne qui observerait les images présentées sur cet écran particulier.

Nous avons ici retenu la seconde approche, car la finalité de l'analyse que nous présentons est de pouvoir quantifier la redondance de l'information de chrominance et de l'information de luminance sur un groupe d'images que nous souhaitons par la suite utiliser dans le cadre d'expériences où elles seront présentées à des observateurs. Nous avons utilisé une procédure de conversion spécifiquement adaptée à notre écran d'expérimentation (Iiyama Vision Master Pro 513) qui tient compte à la fois des propriétés de non-linéarité de ses trois primaires (R, G et B) et de sa luminance résiduelle. Ces caractéristiques ont été mesurées précisément à l'aide d'un spectroradiomètre (CS 2000, Konica-Minolta, Inc). L'excitation de chaque type de cône résultant de la présentation de l'image sur l'écran est ensuite calculée grâce aux fonctions d'absorption des cônes de Smith et Pokorny (1975). Les mesures de calibration de notre écran, ainsi que les détails des étapes de cette conversion sont fournis en annexe (Annexes A et B).

Étape 2 : Conversion dans l'espace d'opposition des cônes (LGN)

Les images sont ensuite projetées dans l'espace d'opposition des cônes post-récepteurs constitué de 3 canaux : un canal de luminance (Lum) et deux canaux de chrominance (que nous noterons RG et BY). Dans un premier temps, nous utilisons les définitions des canaux proposées par Párraga, Brelstaff et al. (1998) et utilisées dans plusieurs études sur les statistiques des images naturelles (Johnson, Kingdom et al., 2005 ; Hansen et Gegenfurtner, 2009 ; Párraga, Troscianko et al., 2002). La spécificité de ces définitions est de proposer une adaptation locale de l'information de chrominance en fonction de la luminance. Ces définitions se présentent sous forme de transformations non linéaires des signaux issus des cônes, incluant une normalisation des canaux de chrominance par des termes dépendants du signal de luminance.

Les transformations qui ont été effectuées à partir des coordonnées LMS sont les suivantes :

$$\begin{aligned}
 Lum &= L + M + \varepsilon \\
 RG &= \frac{L - \alpha M}{Lum} \\
 BY &= \frac{2S - \beta Lum}{2S + \beta Lum}
 \end{aligned} \tag{2.1}$$

avec $\varepsilon = 10^{-52}$ pour éviter des divisions par zero dans les cas où $L + M = 0$ et (α, β) un

couple de paramètres définis par les propriétés de l'écran d'affichage (Annexe A).

Étape 3 : Détection des contours (V1)

Avant d'effectuer une détection de contours, nous avons appliqué un filtre gaussien d'écart type 1 pixel sur chacun des 3 canaux afin de réduire le bruit potentiellement présent dans nos images. Pour l'étape de détection des contours nous avons ensuite utilisé des filtres de Sobel, notés ici S_V pour le filtre vertical et S_H pour le filtre horizontal :

$$S_V = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \text{ et } S_H = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Les contours dans les deux directions sont estimés en appliquant chacun des filtres, puis ces résultats sont moyennés pour obtenir les cartes de contours finales. En notant I_X la matrice correspondant à l'ensemble des pixels d'une image dans l'une des trois dimensions de l'espace d'opposition des cônes ($X \in \{Lum, RG, BY\}$) et E_X la matrice de contours obtenue, l'opération réalisée est donc formellement :

$$E_X = \frac{I_X * S_V + I_X * S_H}{2} \quad (2.2)$$

où $*$ est l'opérateur du produit de convolution.

La Figure 2.3 synthétise les différentes étapes du traitement.

2.2.1.3 L'information mutuelle

Afin de quantifier la redondance de l'information entre les trois canaux aux différentes étapes du traitement visuel, nous avons estimé les valeurs d'information mutuelle. L'information mutuelle de deux variables aléatoires X et Y est une quantité représentant leur dépendance statistique. Pour un couple (X, Y) , l'information mutuelle est nulle si, et seulement si, les deux variables sont indépendantes et elle croît lorsque la dépendance augmente.

L'information mutuelle d'un couple de variables (X, Y) est définie dans le cas continu de la façon suivante :

$$I(X, Y) = \int_{\mathbb{R}} \int_{\mathbb{R}} p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right) dx dy \quad (2.3)$$

avec $p(x, y)$ la densité de la loi de probabilité de (X, Y) et $p(x)$ (resp. $p(y)$) la densité marginale de la loi de X (resp. de Y).

Cette mesure fournit une information complète sur la dépendance de deux variables aléatoires et permet de révéler des interactions plus complexes qu'une simple corrélation linéaire

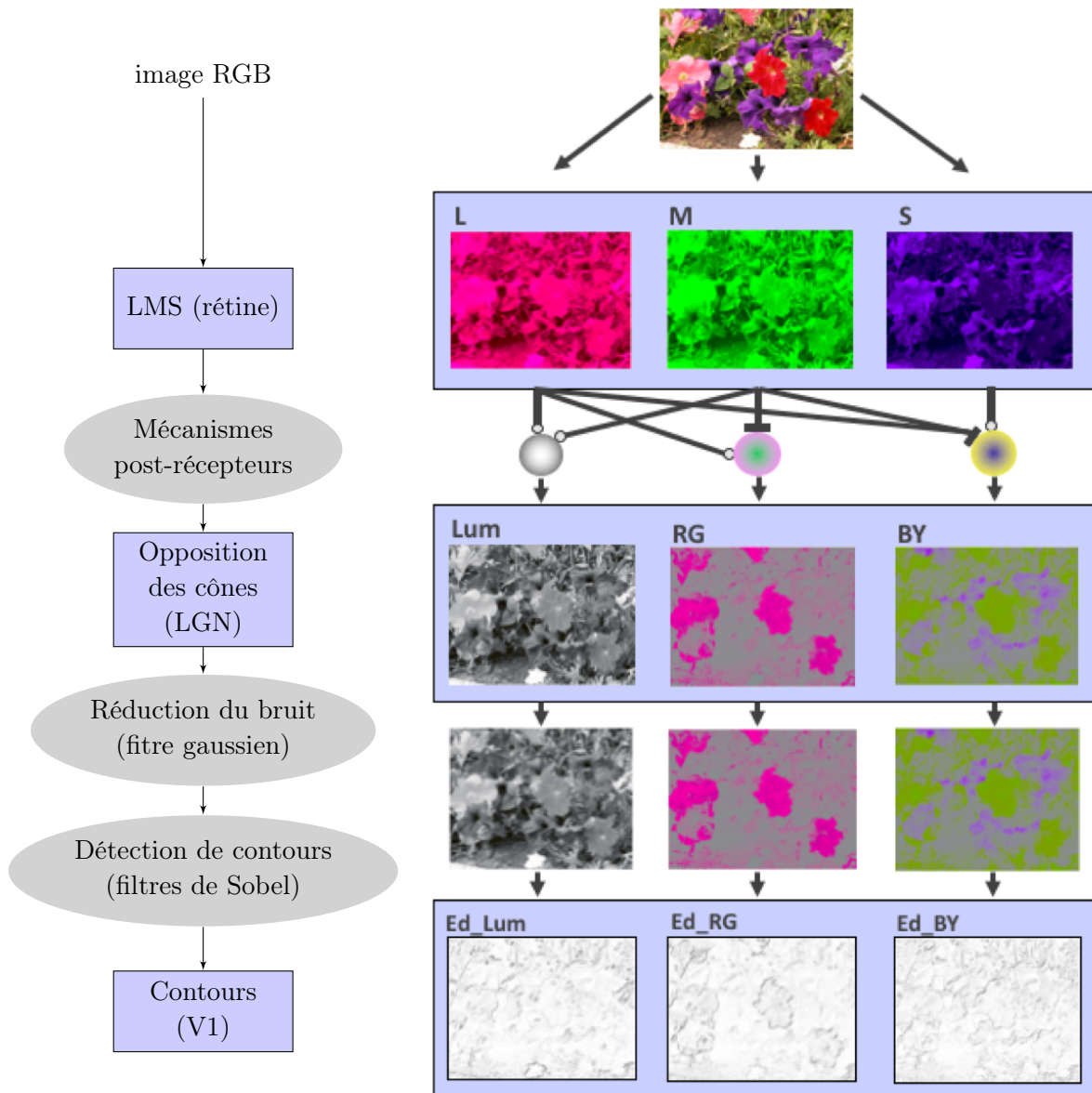


FIGURE 2.3 – Les étapes de la chaîne de traitement. L'image est dans un premier temps convertie dans l'espace d'excitation des cônes L, M et S (rétine), puis dans l'espace d'opposition des cônes (LGN). Nous effectuons ensuite un lissage destiné à réduire le bruit, puis une détection de contours pour obtenir le résultat final, imitant le résultat de l'encodage de l'image dans le cortex visuel primaire (V1) par le système visuel.

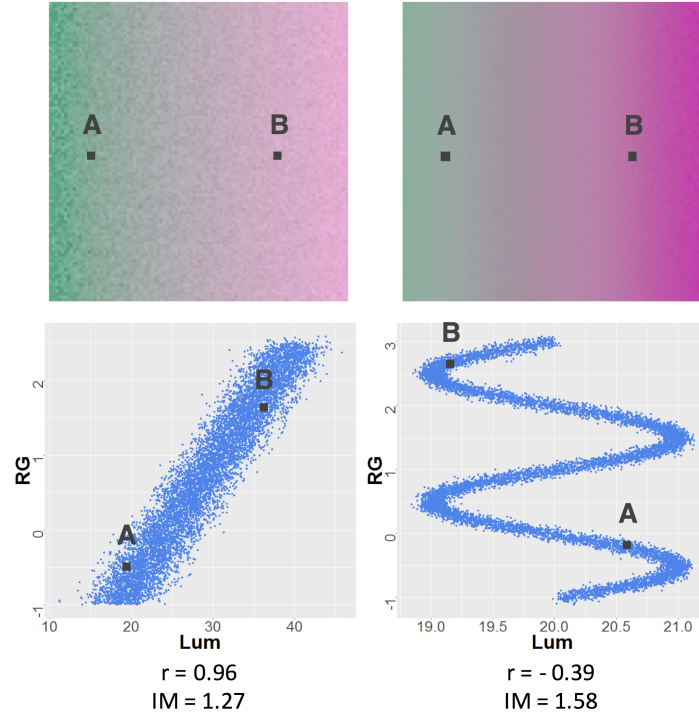


FIGURE 2.4 – Valeurs d'information mutuelle (IM) et de corrélation (r) entre luminance (Lum) et chrominance Rouge-Vert (RG) de deux images artificielles. Comme illustré par les distributions des pixels dans les repères RG/Lum, il existe pour l'image de droite une forte interaction entre ces deux valeurs, bien capturée par l'information mutuelle ($IM=1.58$) mais pas par la corrélation ($r=-0.39$) puisqu'il ne s'agit pas d'une relation linéaire. L'information mutuelle nous fournit donc une estimation plus précise des interactions entre variables, non limitée à des dépendances linéaires.

(Figure 2.4). Cependant, le passage de la définition théorique à son estimation effective présente quelques difficultés. Il n'est, en effet, pas toujours facile d'estimer concrètement la densité de probabilité jointe de deux processus aléatoires. Plusieurs algorithmes ont ainsi été proposés pour obtenir une estimation de l'information mutuelle sans passer par un calcul de densité de probabilité. Nous avons utilisé pour notre étude un estimateur basé sur la méthode des KNN (*k-Nearest Neighbors* ou méthode des plus proches voisins) proposé par Kraskov et al. (2004). Par rapport à d'autres méthodes, cet estimateur a l'avantage de présenter un très faible biais et de pouvoir être utilisé même avec peu de données. L'article de Walters-Williams et Li (2009) répertorie les différents estimateurs existants et les études comparant leurs performances.

L'équation de l'estimateur de l'information mutuelle que nous utiliserons est la suivante :

$$\hat{I}(X, Y) = \psi(N) + \psi(k) + \frac{1}{N} \sum_{i=1}^N \mathbb{E}[\psi(n_x(i) + 1) + \psi(n_y(i) + 1)] \quad (2.4)$$

avec $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ la fonction digamma, N le nombre de points mesurés, $n_x(i)$ et $n_y(i)$ les

nombre de points contenus entre le point i et la distance à son k -ième voisin selon les deux directions.

Cette équation est obtenue à partir d'un estimateur de l'entropie et en utilisant la relation $I(X, Y) = H(X) + H(Y) - H(X, Y)$, où $H(X)$ et $H(Y)$ sont les entropies des distributions X et Y et $H(X, Y)$ est l'entropie conjointe de X et Y . La figure 2.5 présente de façon simplifiée le principe de l'estimateur : étant donné le paramètre k ($k=3$ dans l'exemple), pour chaque point i de coordonnées (x_i, y_i) , on trace la boule centrée en (x_i, y_i) incluant les k plus proches voisins du point. On projette ensuite la boule sur chacun des demi-espaces et on relève le nombre de plus proches voisins dans cette projection, bords exclus, ce qui correspond aux valeurs $n_x(i)$ et $n_y(i)$ (sur l'exemple, $n_x(i) = 5$ et $n_y(i) = 4$).

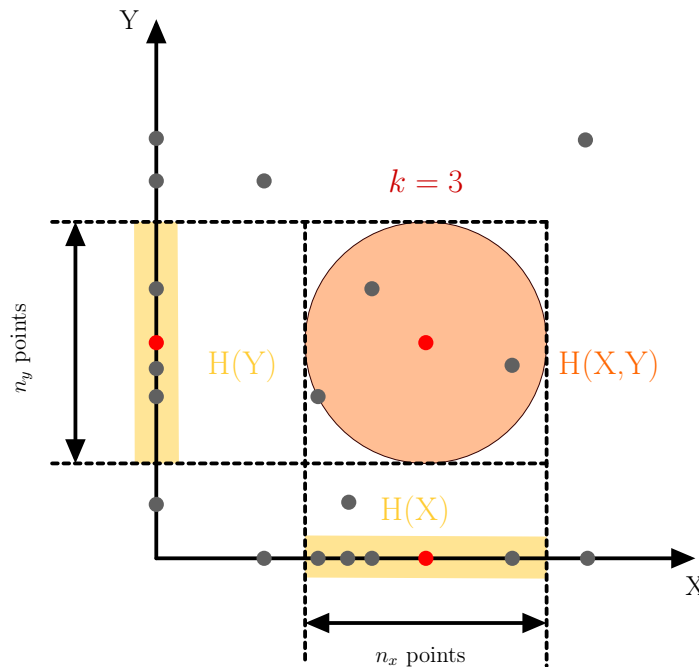


FIGURE 2.5 – Illustration de la méthode KNN. On trace la boule centrée sur le point de coordonnées (x_i, y_i) incluant ses k plus proches voisins ($k=3$). On peut ensuite estimer les entropies $H(X)$ et $H(Y)$ à partir de la projection de la boule sur chacun des demi-espaces X et Y en relevant le nombre de plus proches voisins ($n_x(i)$ et $n_y(i)$) contenus dans cette projection. Une estimation de l'information mutuelle peut ensuite être obtenue à partir de ces estimateurs d'entropie.

L'un des principaux inconvénients de cet estimateur est que le résultat obtenu dépend en partie de la valeur du paramètre k utilisée pour la recherche des plus proches voisins. Il n'existe pas de méthode permettant de choisir optimalement la valeur de k . Nous avons donc choisi dans un premier temps de fixer empiriquement la valeur de k à 25.

Pour les calculs d'information mutuelle, nous avons sélectionné aléatoirement 10 000 pixels

sur chacune des images. Ce nombre a été choisi de façon à minimiser les temps de calcul, tout en obtenant des résultats robustes au tirage aléatoire des pixels. Nous avons ensuite divisé chaque canal par son écart type. Cette étape, dite de réduction, permet notamment d'obtenir des grandeurs sans unité. Nous avons enfin utilisé l'estimateur de Kraskov, avec $k = 25$, pour calculer les valeurs d'information mutuelle.

2.2.1.4 Les paramètres

Dans la chaîne de traitement présentée précédemment (Fig 2.3), à chacune des étapes, plusieurs paramètres ont été fixés de manière empirique ou en reprenant la méthode proposée par Hansen et Gegenfurtner (2009). Afin d'évaluer la fiabilité de la méthode proposée et des valeurs d'information mutuelle obtenues, nous avons cherché à mesurer l'influence de certains de ces paramètres sur nos résultats.

Paramètres de la chaîne de traitement

Étape 1 : Conversion dans l'espace d'absorption des cônes (rétine)

Notre modèle de référence se base sur les fonctions fondamentales d'absorption des cônes de Smith et Pokorny (1975). Ces courbes correspondent à la sensibilité spectrale de chacun des trois types de cônes et ont été mesurées à l'aide de divers protocoles expérimentaux, conduisant à des résultats sensiblement différents (voir Chapitre 1). La CIE recommande actuellement l'utilisation d'un autre triplet de fonctions : celles proposées par Stockman et Sharpe (2000).

Une différence importante entre ces deux estimations des fonctions fondamentales d'absorption des cônes est que les fonctions de Smith et Pokorny (1975) ont, par définition, pour propriété que la somme des valeurs de L et M correspond à la fonction d'efficacité lumineuse modifiée par Judd-Vos, $V_M(\lambda)$ (Judd, 1951 ; Vos, 1978) tandis que, dans le cas des fonctions de Stockman et Sharpe (2000), une combinaison linéaire de L et M permet d'obtenir la courbe $V^*(\lambda)$ (Sharpe et al., 2011). Les courbes $V_M(\lambda)$ et $V^*(\lambda)$ représentent la sensibilité spectrale de l'oeil humain et correspondent ici aux définitions utilisées pour la luminance.

Dans le cas de l'utilisation des fonctions fondamentales de Smith et Pokorny (1975), nous avons :

$$Lum = V_M = L + M + \varepsilon \quad (2.5)$$

En utilisant les fonctions fondamentales de Stockman et Sharpe (2000) la définition de la luminance est la suivante :

$$Lum = V^* = 0.69 L + 0.35 M + \varepsilon \quad (2.6)$$

Pour la définition de l'information de chrominance portée par les canaux RG et BY, nous suivons la même procédure que précédemment, consistant à déterminer le couple (α, β) adapté à l'écran considéré qui permet de vérifier la contrainte $RG = BY = 0$ pour l'ensemble des pixels achromatiques. Dans le cas de l'utilisation des fonctions fondamentales de Stockman et Sharpe (2000) et pour notre écran (Iiyama Vision Master Pro 513) nous avons $(\alpha, \beta) = (1.16, 0.45)$. Nous désignerons par la suite par **CIE06** le modèle basé sur ces définitions.

Étape 2 : Conversion dans l'espace d'opposition des cônes (LGN)

Bien qu'il soit unanimement admis que le système visuel convertit les signaux de sortie des cônes dans un espace d'opposition afin de réduire la redondance d'information entre canaux, la façon dont ces mécanismes post-récepteurs combinent les réponses des cônes n'est pas claire et plusieurs modèles ont été proposés (Ingling et Tsou, 1988; Buchsbaum et Gottschalk, 1983; Ruderman et al., 1998). Nous avons, pour notre modèle de référence, utilisé les définitions proposées par Párraga, Brelstaff et al. (1998) et intégrant une normalisation divisive des réponses chromatiques par la luminance. Afin de contrôler l'effet du choix de ce modèle sur nos résultats, nous les avons comparés à ceux obtenus en considérant deux autres modèles couramment utilisés pour la conversion vers l'espace d'opposition des cônes :

- **Modèle linéaire simple (Lin).** Il s'agit du modèle le plus simple, qui revient à considérer que les mécanismes post-récepteurs peuvent être décrits comme de simples combinaisons linéaires des réponses des cônes, sans intégrer de processus d'adaptation. Les définitions correspondantes que nous utiliserons ici sont les suivantes :

$$\begin{aligned}Lum &= L + M \\RG &= L - \alpha M \\BY &= 2S - \beta(L + M)\end{aligned}\tag{2.7}$$

où (α, β) est un couple de paramètres adapté à l'écran de présentation.

- **Modèle non-linéaire logarithmique (Log).** Ces définitions ont été proposées par Ruderman et al. (1998) dans leur étude visant à répliquer les résultats de Buchsbaum et Gottschalk (1983) sur la décorrélation des réponses des cônes grâce aux mécanismes d'opposition. Avant de pratiquer leur ACP, ils ont converti les signaux de réponse dans un espace logarithmique puis soustrait à chaque signal sa valeur moyenne sur l'ensemble de l'image. Cette première transformation est justifiée par la volonté d'obtenir un espace permettant d'améliorer la représentation des données mais également par des résultats psychophysiques suggérant une non-linéarité de la réponse des photorécepteurs en fonction de l'intensité du stimulus lumineux présenté (loi de Weber-Fechner). La soustraction par la moyenne équivaut à une procédure d'adaptation comparable à celle proposée par Von Kries (Kries, 1902). Les transformations correspondantes effectuées sur les coordon-

nées LMS de nos images sont :

$$\begin{aligned} L_{log} &= \log(L) - \overline{\log(L)} \\ M_{log} &= \log(M) - \overline{\log(M)} \\ S_{log} &= \log(S) - \overline{\log(S)} \end{aligned} \quad (2.8)$$

où $\overline{\log(L)}$ (resp. $\overline{\log(M)}$, $\overline{\log(S)}$) est la valeur moyenne de $\log(L)$ (resp. $\log(M)$, $\log(S)$) sur l'ensemble des pixels de l'image.

Finalement, les trois canaux post-récepteurs obtenus dans le cadre de ce modèle sont les suivants :

$$\begin{aligned} Lum &= L_{log} + M_{log} \\ RG &= L_{log} - M_{log} \\ BY &= 2S_{log} - (L_{log} + M_{log}) \end{aligned} \quad (2.9)$$

Il s'agit des 3 axes expliquant la plus grande variance obtenus par Ruderman et al. (1998) dans leur étude, à ceci près que nous avons supprimé la contribution des cônes S dans le calcul de la luminance et négligé les facteurs qu'ils utilisaient pour normaliser leurs axes à l'unité. Des définitions identiques sont utilisées dans l'article de Hansen et Gegenfurtner (2009), en comparaison des définitions du modèle de référence.

Nos définitions des canaux dans l'espace d'opposition des cônes dépendent également du couple de paramètres (α, β) de notre écran. Pour contrôle, nous testerons donc le cas d'une conversion utilisant les mêmes définitions des canaux d'opposition que le modèle de référence, mais en considérant un autre écran. Les définitions des canaux d'opposition des cônes seront donc celles du modèle de référence avec (α', β') un nouveau couple de paramètres spécifique à un autre écran (Philips 201B4) dont nous connaissons également les caractéristiques.

Paramètres de l'estimation de l'information mutuelle

Nous avons réduit les variables (c'est-à-dire, divisé la valeur de chaque pixel par l'écart-type de la distribution sur l'ensemble de l'image) avant d'estimer l'information mutuelle. Cette opération permet notamment d'obtenir des grandeurs sans unité et indépendantes de l'échelle pour tous les canaux. Elle a cependant pour inconvénient d'affecter la dispersion des variables, ce qui peut avoir des conséquences sur les valeurs d'information mutuelle calculées, comme nous en discuterons plus loin.

Par ailleurs, dans la formule de l'estimateur d'information mutuelle utilisée, nous avons empiriquement fixé le paramètre k à la valeur 25. Il est parfois préconisé de choisir $k = \sqrt{N}$ où N est le nombre total d'observations. Pour comparaison nous calculerons donc les valeurs d'information mutuelle dans le cas où $k=100$.

2.2.1.5 Métriques

Le modèle proposé décompose une image couleur selon trois canaux d'information, en trois étapes de traitement. Nous calculons pour chaque image une valeur d'information mutuelle pour chaque paire de canaux, à chacune des trois étapes du traitement : l'espace LMS d'absorption des cônes (rétine), l'espace d'opposition des cônes (LGN) et l'espace des contours (V1).

Afin de comparer les différents modèles nous utilisons deux métriques que nous présentons ici.

Évolution des valeurs moyennes d'information mutuelle le long du traitement visuel

Cette métrique consiste simplement à calculer les valeurs moyennes d'information mutuelle sur l'ensemble des images à chacune des étapes de traitement. Notons C_1 , C_2 et C_3 les trois canaux d'information, avec :

- $\{C_1, C_2, C_3\} = \{L, M, S\}$ dans l'espace LMS d'absorption des cônes,
- $\{C_1, C_2, C_3\} = \{Lum, RG, BY\}$ dans l'espace d'opposition des cônes et
- $\{C_1, C_2, C_3\} = \{Ed_Lum, Ed_RG, Ed_BY\}$ dans l'espace des contours.

Nous nous intéresserons aux valeurs d'information mutuelle moyennes pour chaque paire de canaux indépendamment ($\overline{IM}(C_i, C_j)$, $\{i, j\} \in \{1, 2, 3\}$) ainsi que sur l'ensemble des trois canaux ($\overline{IM}(C_1, C_2, C_3) = \sum_{i,j} \frac{\overline{IM}(C_i, C_j)}{3}$). Nous nous restrindrions ici à l'analyse des valeurs moyennes d'information mutuelle afin de mesurer l'évolution de la redondance d'information le long du traitement visuel et de comparer les différents modèles, ainsi qu'aux mesures concernant les canaux de luminance et de chrominance Rouge-Vert. Nous laissons de côté le canal Bleu-Jaune car il intervient moins dans la vision centrale, en raison notamment de l'absence de cônes S dans la *foveola*, et car plusieurs études soulignent un rôle prépondérant de l'information du canal Rouge-Vert (Mullen et Kingdom, 2002 ; Frey, Wirz et al., 2011).

Corrélation de rang

Notre objectif étant d'obtenir un critère qui nous permette de quantifier la dépendance statistique entre information de luminance et information de chrominance pour chacune des images de notre base, nous nous intéressons également au classement relatif des images obtenu selon ce critère pour les différents modèles. Nous calculons pour cela la corrélation de Spearman (corrélation de rang), ρ , entre les valeurs d'information mutuelle obtenues à partir de deux modèles différents. Le coefficient de Spearman entre deux variables X et Y est calculé avec la formule suivante :

$$\rho(X, Y) = 1 - \frac{6 \sum_{i=1}^n [r(X_i) - r(Y_i)]^2}{N^3 - N} \quad (2.10)$$

avec N le nombre d'observations pour chacune des variables, $r(X_i)$ le rang de X_i dans la distribution $X_1 \dots X_N$ et $r(Y_i)$ le rang de Y_i dans la distribution $Y_1 \dots Y_N$.

On obtient ainsi une mesure de la corrélation, non pas entre les valeurs prises par les deux variables, mais entre les rangs de ces valeurs au sein de leurs distributions respectives.

2.2.2 Résultats

Dans cette section nous présentons dans un premier temps les estimations d'information mutuelle obtenues avec notre modèle de référence (**Ref**). Nous caractérisons ensuite un biais potentiellement présent dans de précédentes études. Nous comparons enfin les résultats obtenus avec les différents modèles afin de mesurer la sensibilité de notre critère d'évaluation d'indépendance statistique aux paramètres de la chaîne de traitement.

2.2.2.1 Résultats du modèle de référence

Nous avons, pour chacune des images, calculé les valeurs d'information mutuelle entre chaque paire de canaux et à chacune des trois étapes de traitement dans le cadre du modèle de référence. Les résultats obtenus sont présentés sur la figure 2.6.

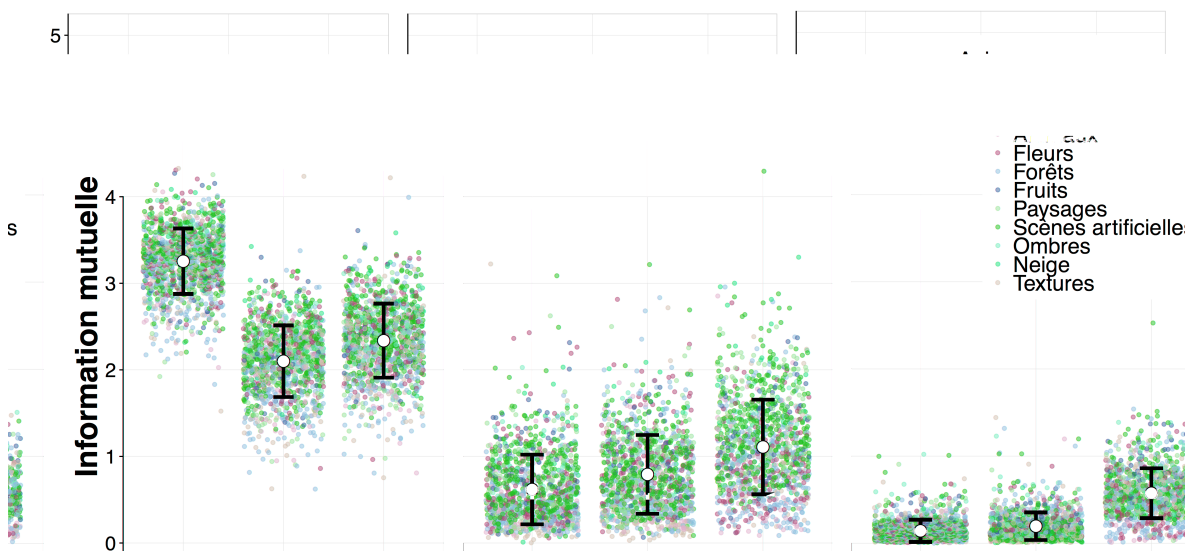


FIGURE 2.6 – Valeurs d'information mutuelle entre chaque paire de canaux, à chaque étape du traitement visuel (LMS, Opposition des cônes et Contours), obtenues avec le modèle de référence (**Ref**).

Nous constatons que les valeurs d'information mutuelle obtenues :

1. Décroissent en moyenne le long du traitement visuel.
2. Sont très variables d'une image à l'autre

Ce premier résultat est en accord avec ceux obtenus par Hansen et Gegenfurtner (2009) et l'hypothèse de codage efficace de l'information : si les systèmes sensoriels encodent l'information de sorte à réduire la redondance entre canaux et optimiser sa transmission, on s'attend effectivement à observer cette décroissance d'information mutuelle au cours des étapes du traitement.

Le second résultat semble mettre en évidence un artefact dans la méthode proposée par Hansen et Gegenfurtner (2009) : dans cette étude, les auteurs avaient, en effet, effectué pour chaque paire de canaux à chacun des niveaux du traitement visuel un unique calcul d'information mutuelle pour l'ensemble des images de leur base. La figure 2.7 illustre le biais produit par cette méthode : les valeurs d'information mutuelle obtenues varient beaucoup d'une image à l'autre (Figure 2.6), en comparant directement les distributions sur l'ensemble des images, les spécificités individuelles sont perdues. Il s'agit d'un exemple de paradoxe de Simpson (Simpson, 1951), où l'effet observé pour une image peut être modifié lorsqu'on considère plusieurs images ensemble (Fig. 2.7). Effectuer un calcul d'information mutuelle indépendamment pour chacune des images permet donc de caractériser plus précisément la redondance entre information de luminance et information de chrominance. Notre modèle de référence nous fournit donc un indicateur permettant d'évaluer pour chaque image la redondance statistiques entre l'information contenue par chacun des canaux et d'effectuer un classement de nos images en fonction de ce critère.

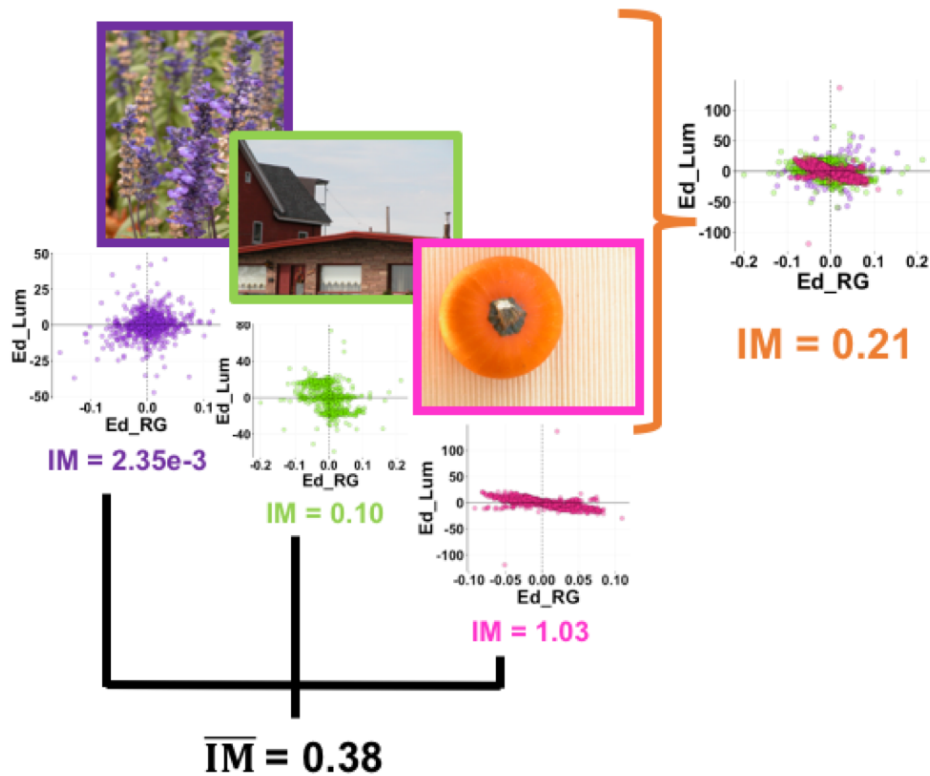
Analyse par catégorie

On peut supposer que les images des différentes catégories ne possèdent pas les mêmes propriétés spatio-chromatiques : par exemple, les images de la catégorie Fleurs contiennent certainement plus de contours chromatiques avec des valeurs élevées que celles de la catégorie Neige. Ces différences pourraient conduire à des valeurs d'information mutuelle sensiblement différentes d'une catégorie à l'autre. Afin de vérifier cette hypothèse, nous avons calculé les valeurs d'information mutuelle entre les 3 canaux de contours pour chacune des catégories indépendamment. Nous présentons également les histogrammes de densités jointes des contours de luminance et de chrominance Rouge-Vert des images de chacune des catégories (Fig.2.8). Nous n'observons pas de différence notable entre les images des différentes catégories pour ces deux indicateurs.

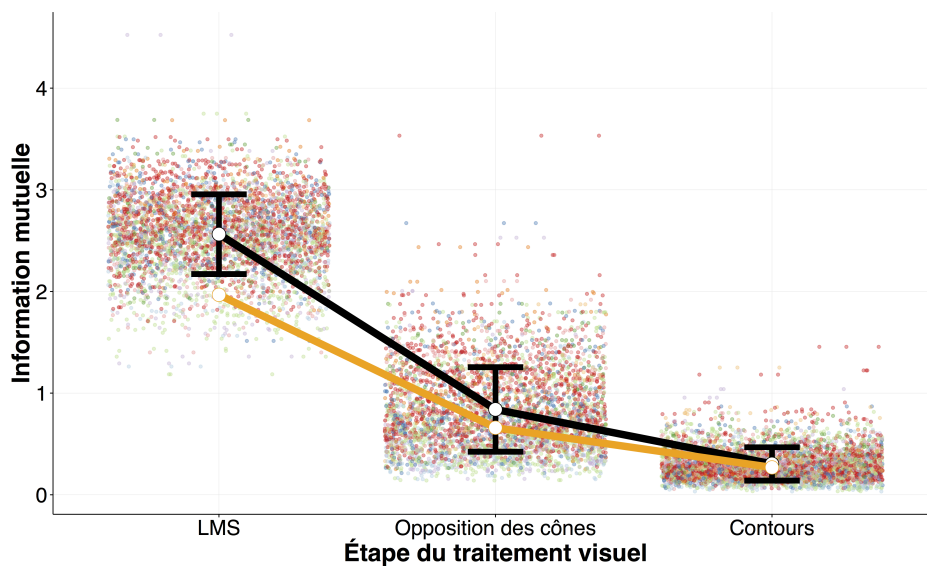
Analyse du classement des images

Afin de comprendre un peu mieux comment l'information mutuelle était reliée aux propriétés spatio-chromatiques des images, nous nous sommes intéressés aux 20 images ayant obtenu les valeurs d'information mutuelle entre contours de luminance et contours de chrominance Rouge-Vert les plus faibles (resp. les plus élevées) avec notre modèle de référence. Ces images sont présentées sur la Figure 2.9 (resp. Figure 2.10).

Visuellement, les images de ces deux groupes semblent présenter quelques spécificités : teintes vert-gris pour les images à faible information mutuelle, orangé-noir pour les images à forte information mutuelle, ... Afin de comparer ces deux groupes nous avons analysé les valeurs

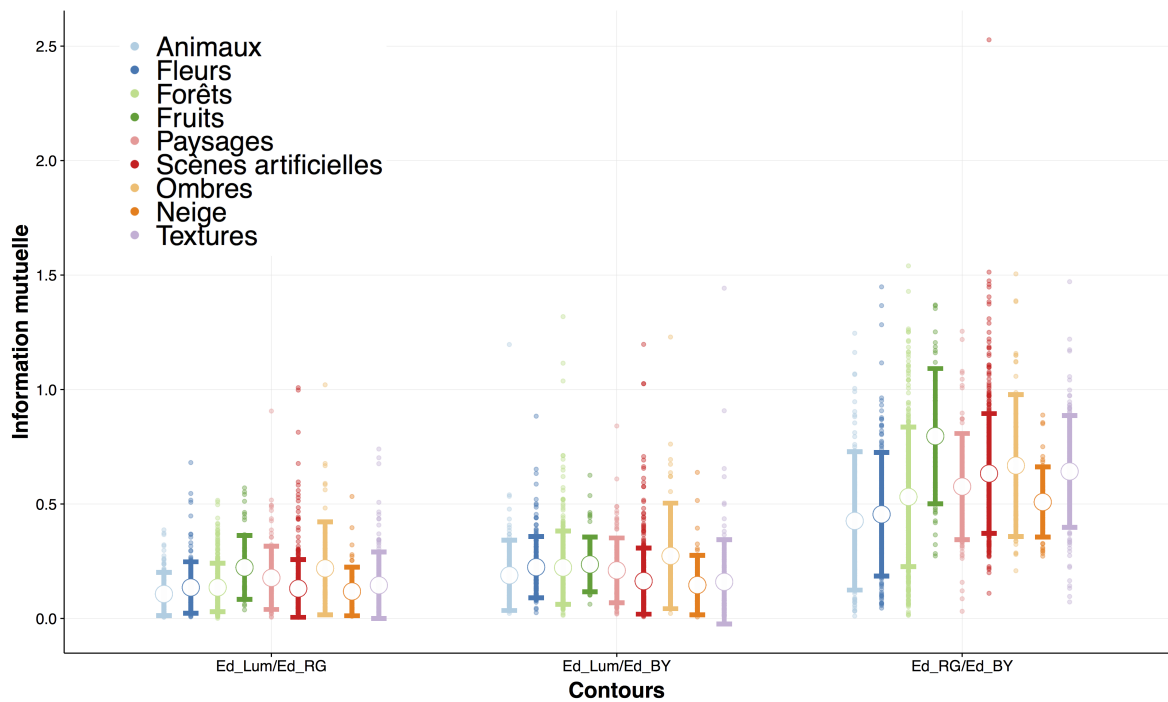


(a) Illustration du biais provoqué par la fusion des images pour le calcul d'information mutuelle. Les trois images présentées sur l'exemple possèdent des distributions de valeurs de contours de luminance et de chrominance Rouge-Vert très différentes, ce qui se traduit également par des valeurs d'information mutuelle différentes. En fusionnant les trois images on obtient un résultat qui n'est plus représentatif des distributions individuelles de chacune des images.

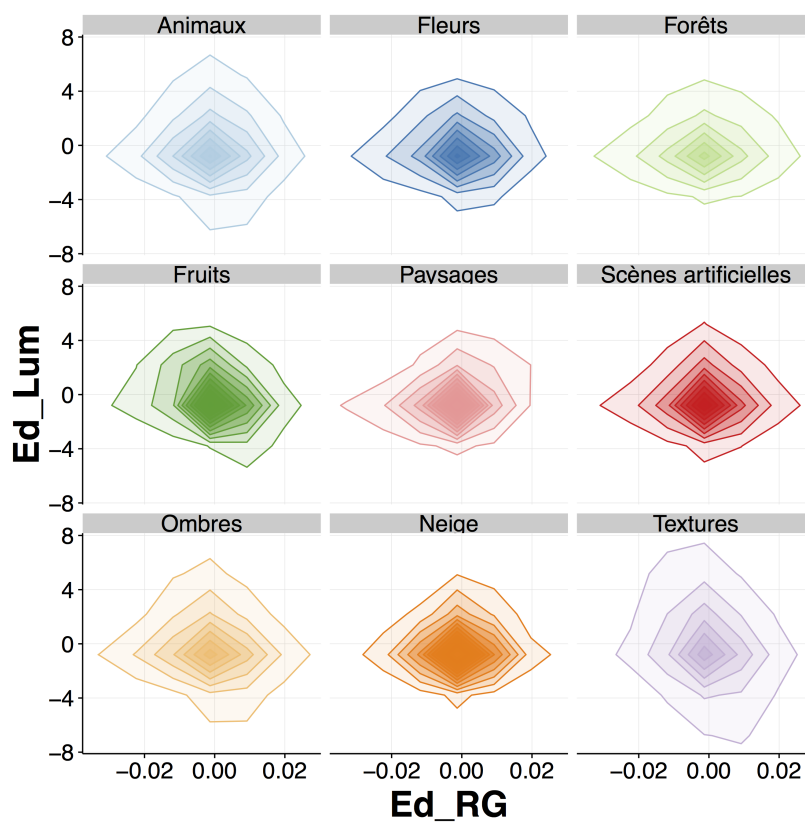


(b) Information mutuelle moyenne entre canaux le long du traitement visuel obtenue en fusionnant les images (en orange) ou en calculant d'abord la valeur de chaque image individuellement (en noir). La fusion des images tend à diminuer les valeurs d'information mutuelle.

FIGURE 2.7 – Artéfact lié à la fusion des images pour le calcul d'information mutuelle.



(a) Information mutuelle entre chaque paire de canaux de contours pour chacune des catégories d'images.



(b) Histogrammes de densités jointes des contours d'information de chrominance Rouge-Vert (Ed_RG) et des contours d'information de luminance (Ed_Lum) pour chacune des catégories d'images.

FIGURE 2.8 – Analyse par catégorie.

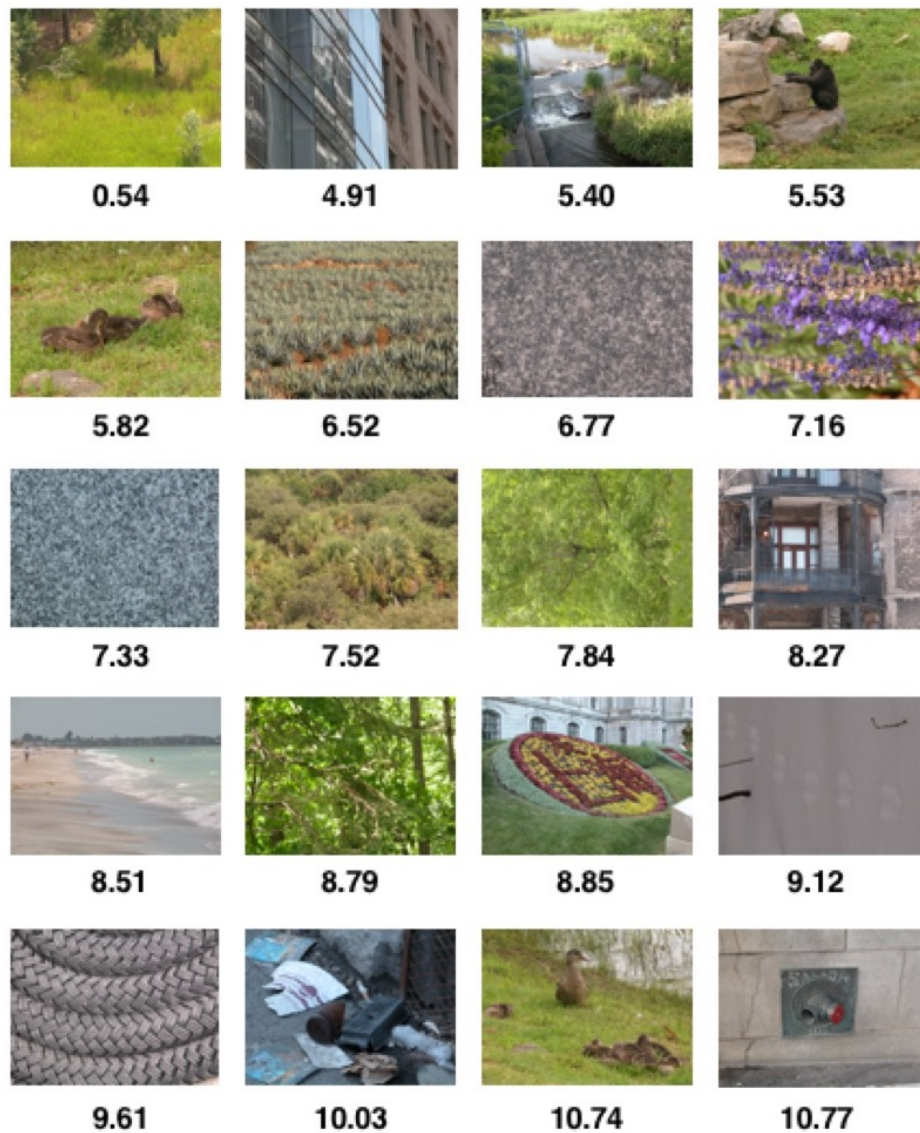


FIGURE 2.9 – Information mutuelle ($\cdot 10^3$) entre contours de luminance (Ed_Lum) et contours de chrominance Rouge-Vert (Ed_RG) des 20 images obtenant les plus faibles valeurs avec le modèle de référence.

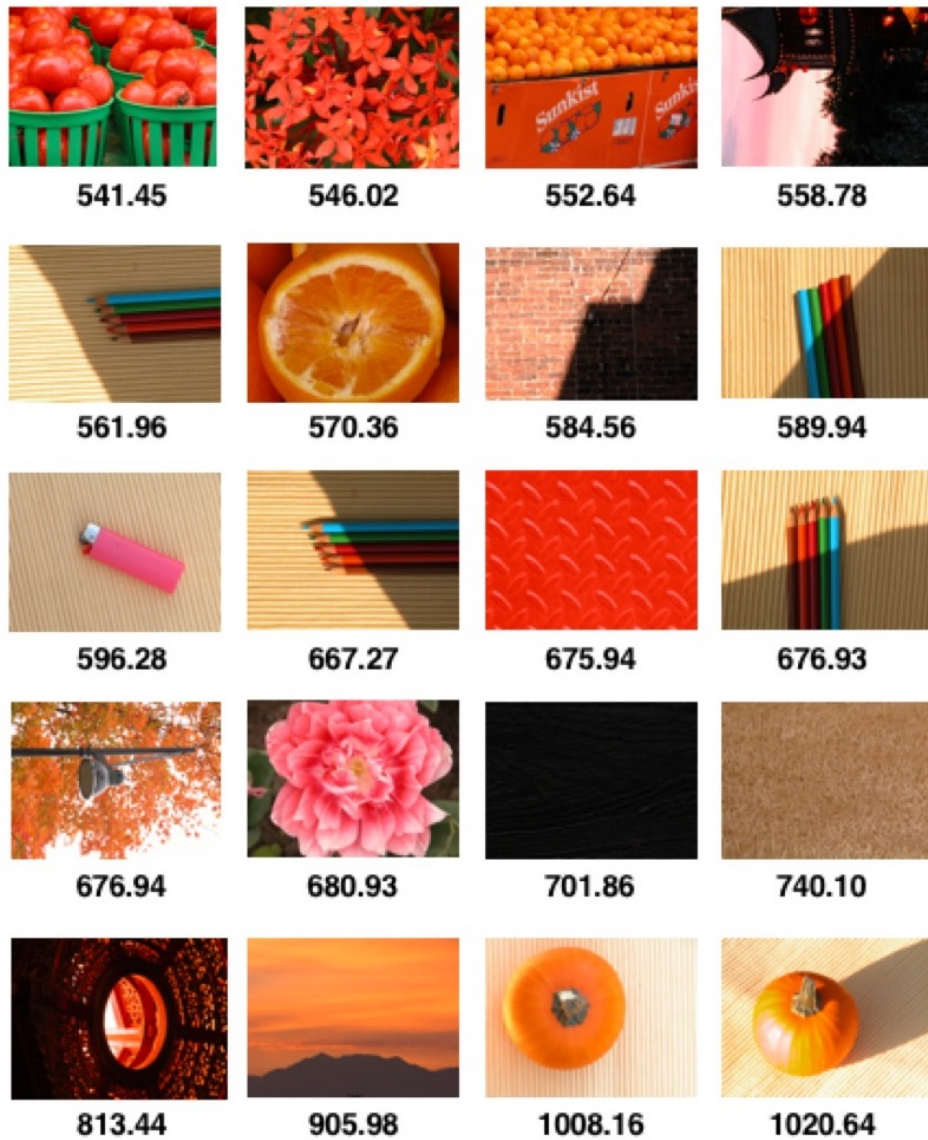
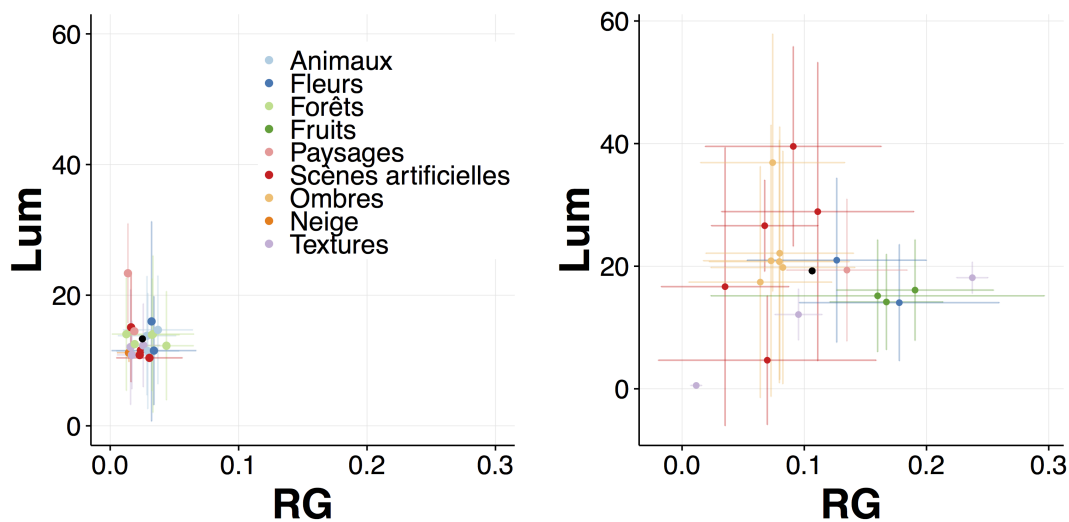


FIGURE 2.10 – Information mutuelle ($\cdot 10^3$) entre contours de luminance (Ed_Lum) et contours de chrominance Rouge-Vert (Ed_RG) des 20 images obtenant les plus fortes valeurs avec le modèle de référence.

moyennes et l'allure des distributions des valeurs de luminance et de chrominance Rouge-Vert sur les images de chacun des groupes. Comme illustré sur la Figure 2.11, les valeurs moyennes de luminance et de chrominance Rouge-Vert absolue mesurées pour chacun des groupes sont significativement différentes ($(|\overline{RG}|, \overline{Lum}) = (0.0251 \pm 0.0089, 13.32 \pm 2.86)$ pour le groupe d'information mutuelle faible et $(|\overline{RG}|, \overline{Lum}) = (0.1064 \pm 0.0564, 19.24 \pm 9.15)$ pour le groupe d'information mutuelle élevée). L'allure de la distribution des pixels de chacune des images des deux groupes dans un repère (RG,Lum) est fournie en annexe (Annexe C) mais la Figure 2.12 présente une image caractéristique pour chacun des groupes.



(a) 20 images de plus faible information mutuelle Ed_Lum/Ed_RG. (b) 20 images de plus forte information mutuelle Ed_Lum/Ed_RG.

FIGURE 2.11 – Distributions des valeurs moyennes de chrominance Rouge-Vert absolue (RG) et de luminance (Lum) des images situées aux extrema du classement d'information mutuelle entre contours de chrominance Rouge-Vert (Ed_RG) et de luminance (Ed_Lum). Les barres d'erreurs correspondent à l'écart-type. Les points noirs sont les valeurs moyennes obtenues pour chacun des groupes de 20 images.

Nous avons à ce stade pu classer nos images en fonction de la dépendance statistique de leurs contours de luminance et de chrominance Rouge-Vert. Afin d'évaluer la pertinence de cet indice, nous allons désormais tester sa sensibilité à différents paramètres de notre algorithme.

2.2.2.2 Effet des paramètres

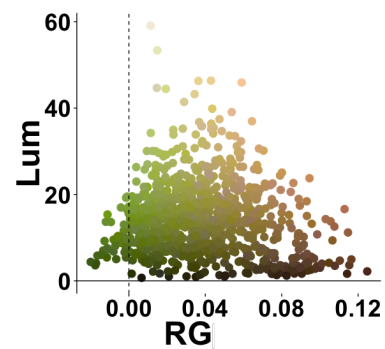
Paramètres du traitement

Effet de l'étape 1 : Conversion dans l'espace d'absorption des cônes (rétine)

Nous comparons les valeurs d'information mutuelle obtenues avec nos deux modèles de



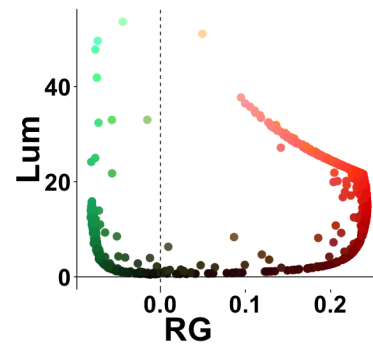
(a) Exemple d'image de faible information mutuelle Ed_Lum/Ed_RG .



(b) Distribution des pixels de l'image dans un repère (RG,Lum).



(c) Exemple d'image de forte information mutuelle Ed_Lum/Ed_RG .



(d) Distribution des pixels de l'image dans un repère (RG,Lum).

FIGURE 2.12 – Exemples d'images de faible (a) et forte (b) information mutuelle entre contours de luminance (Ed_Lum) et de chrominance Rouge-Vert (Ed_RG) et distributions des pixels dans un repère (Chrominance Rouge-Vert (RG), Luminance (Lum)) (b et d).

conversion dans l'espace LMS : le modèle de référence (**Ref**) qui utilise les fonctions fondamentales des cônes de Smith et Pokorny (1975) et le modèle **CIE06**, basé sur les fonctions fondamentales de Stockman et Sharpe (2000) dont l'emploi est préconisé par la CIE. Comme illustré sur la Figure 2.13, les valeurs d'information mutuelle obtenues avec le modèle **CIE06** sont en moyenne légèrement inférieures à celles calculées avec le modèle de référence.

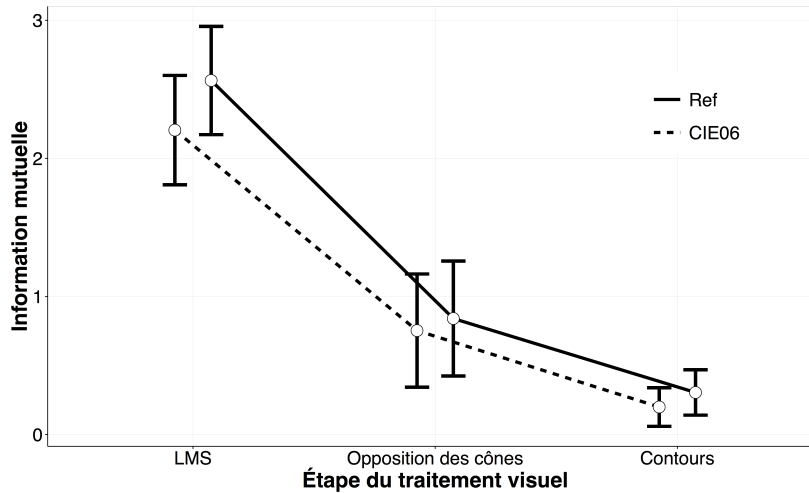


FIGURE 2.13 – Évolution des valeurs d'information mutuelle moyennes le long du traitement en fonction de la conversion dans l'espace LMS (rétine) : comparaison des modèles **Ref** et **CIE06**.

Nous calculons ensuite les corrélations de Spearman entre les valeurs d'information mutuelle moyennes ($\overline{IM}(C_1, C_2, C_3)$) obtenues avec chacun des modèles et les valeurs d'information mutuelle entre les deux premiers canaux ($IM(C_1, C_2)$) afin de mesurer l'impact de ce paramètre sur le classement relatif de nos images. Les valeurs de ces coefficients sont fournies sur la Figure 2.14 qui représente sous forme de nuage de points les valeurs d'information mutuelle obtenues avec le modèle **CIE06** en fonction des valeurs du modèle **Ref** pour l'ensemble de nos images. Nous constatons que les valeurs d'information mutuelle obtenues avec les deux modèles sont très proches et que le classement des images est très peu affecté par ce paramètre.

Effet de l'étape 2 : Conversion dans l'espace d'opposition des cônes (LGN)

Nous comparons ici deux modèles de conversion dans l'espace d'opposition des cônes au modèle de référence (**Ref**) : le modèle linéaire (**Lin**) et le modèle logarithmique (**Log**). Comme précédemment, nous analysons l'évolution des valeurs d'information mutuelle moyennes le long du traitement visuel (Fig. 2.15) et les corrélations de rang entre ces valeurs en fonction du modèle utilisé (Fig. 2.16).

Dans le cas du modèle logarithmique nous obtenons des résultats très proches de ceux du modèle de référence. Le modèle linéaire fournit, en revanche, des valeurs d'information mutuelle supérieures à celles obtenues avec le modèle de référence. Par ailleurs, les corrélations de rang entre ces valeurs et celles du modèle de référence sont relativement faibles ($\rho(\overline{IM}(C_{Ed_Lum}, C_{Ed_RG}, C_{Ed_BY})) = 0.50$).

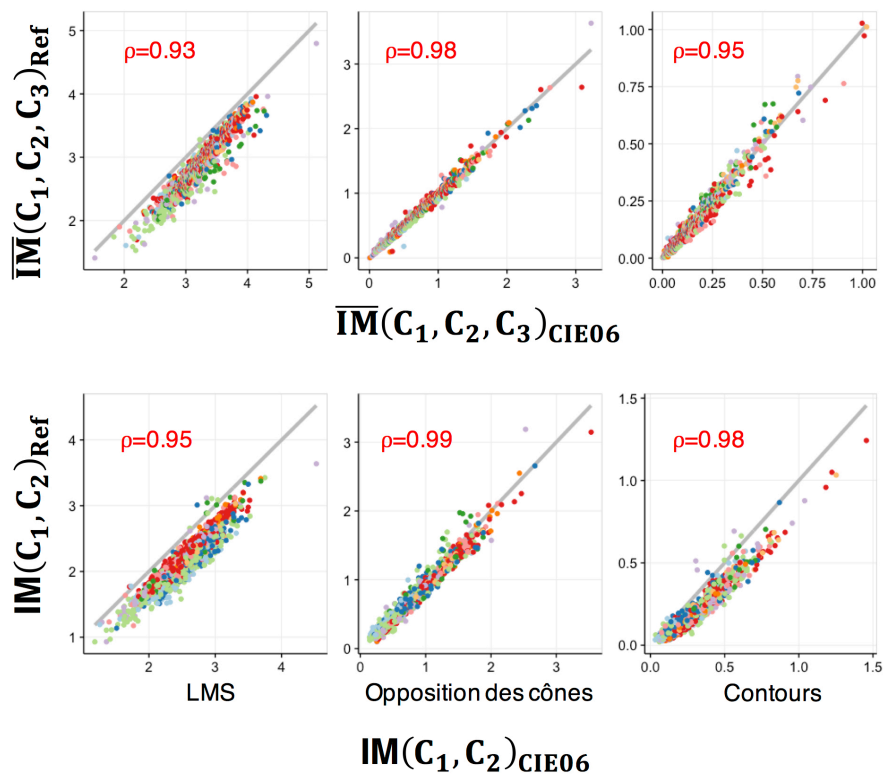


FIGURE 2.14 – Comparaison des valeurs d'information mutuelle moyennes ($\overline{IM}(C_1, C_2, C_3)$) et entre les deux premiers canaux ($IM(C_1, C_2)$) de chacune des images obtenues avec les différentes conversion dans l'espace LMS (rétine) (**CIE06** et **Ref**). Les lignes grises correspondent au cas où les valeurs seraient identiques pour les deux modèles.

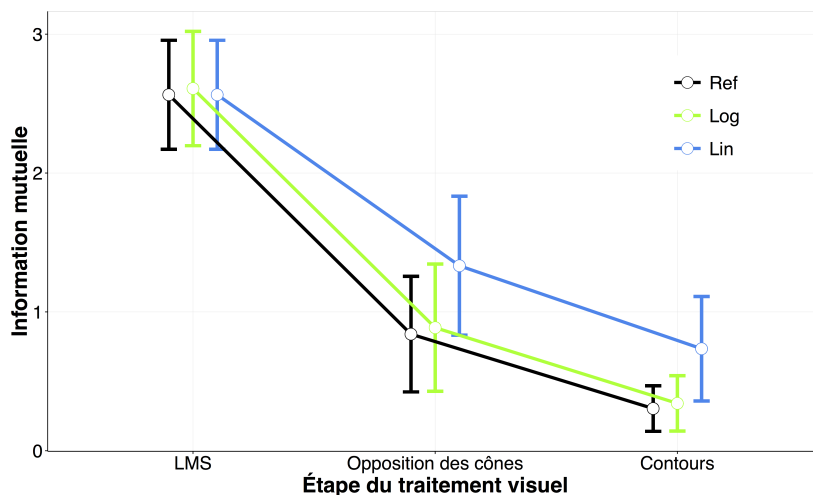
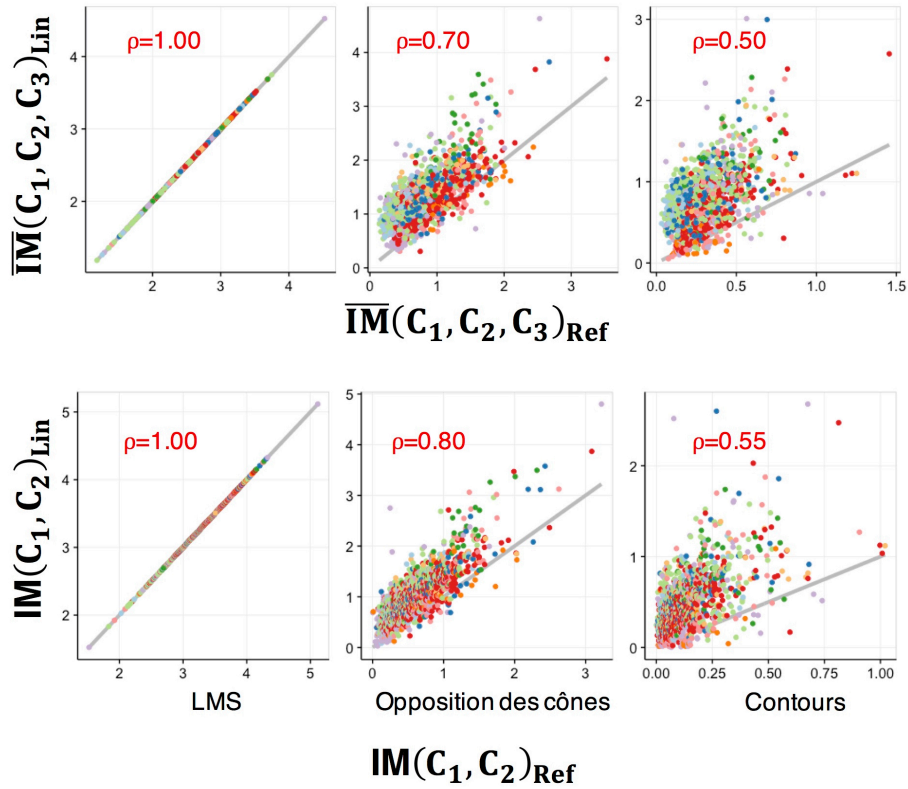
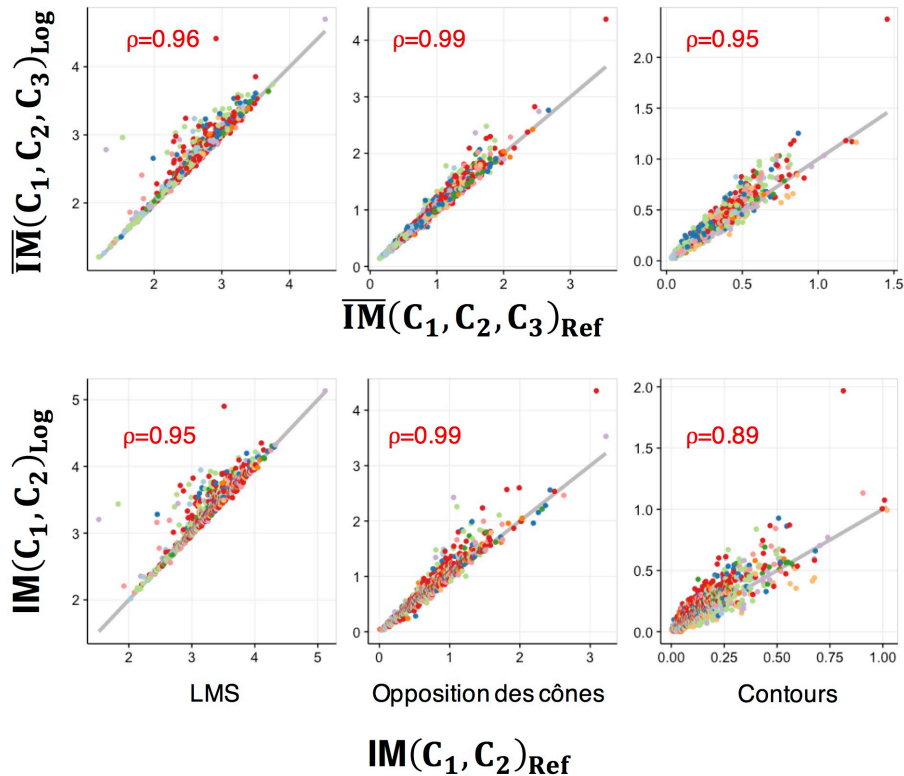


FIGURE 2.15 – Évolution des valeurs d'information mutuelle moyennes le long du traitement visuel en fonction de la conversion dans l'espace d'opposition des cônes (LGN) : comparaison des modèles **Ref**, **Lin** et **Log**.



(a) Comparaison du modèle linéaire (**Lin**) avec le modèle de référence (**Ref**).



(b) Comparaison du modèle logarithmique (**Log**) avec le modèle de référence (**Ref**).

FIGURE 2.16 – Valeurs d'information mutuelle moyennes ($\overline{IM}(C_1, C_2, C_3)$) et entre les deux premiers canaux ($IM(C_1, C_2)$) de chacune des images, obtenues avec les différentes conversions dans l'espace d'opposition des cônes (LGN) (**Ref**, **Lin** et **Log**).

Si l'on considère un autre écran (Philips 201B4), les résultats sont peu modifiés (Fig. 2.17). Les valeurs d'information mutuelle sont donc peu dépendantes de l'écran utilisé.

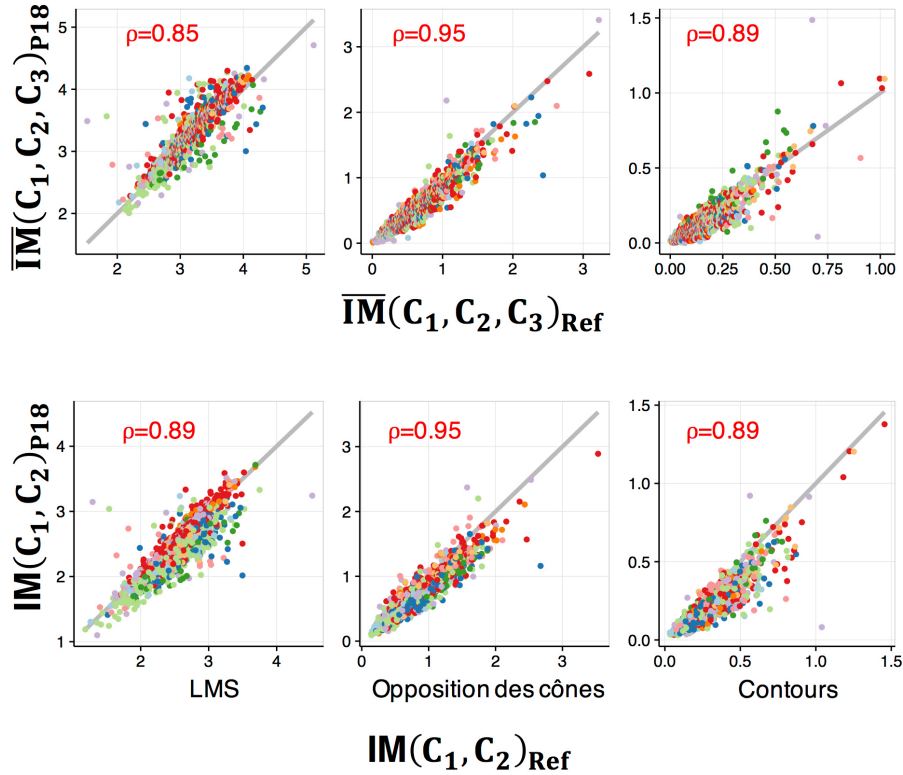


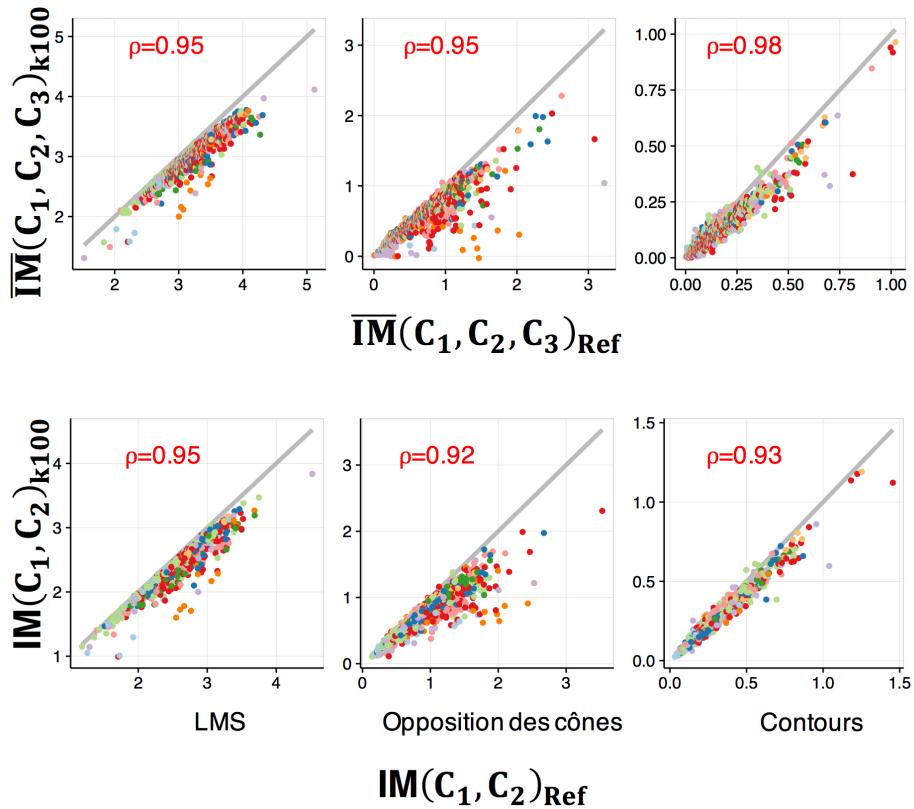
FIGURE 2.17 – Comparaison des valeurs d'information mutuelle moyennes ($\overline{IM}(C_1, C_2, C_3)$) et entre les deux premiers canaux ($IM(C_1, C_2)$) de chacune des images, obtenues avec le modèle de référence dans le cas d'une conversion adaptée à l'écran Philips 201B4 (**P18**) ou à l'écran de référence (**Ref**).

Finalement, nous pouvons observer des différences importantes en fonction de la définition utilisée pour les mécanismes post-récepteurs, majoritairement entre les définitions incluant une étape d'adaptation (**Ref** ou **Log**) et celles ne le faisant pas (**Lin**). La question de l'utilisation d'une définition pertinente des mécanismes post-récepteurs semble donc cruciale pour pouvoir exploiter les mesures d'information mutuelle. Si le but du système visuel est effectivement de réduire la redondance d'information, l'adaptation semble présenter un avantage car les valeurs d'information mutuelle moyennes obtenues avec les définitions tenant compte d'une adaptation (**Ref** ou **Log**) sont plus faibles que celles mesurées avec la définition linéaire (**Lin**).

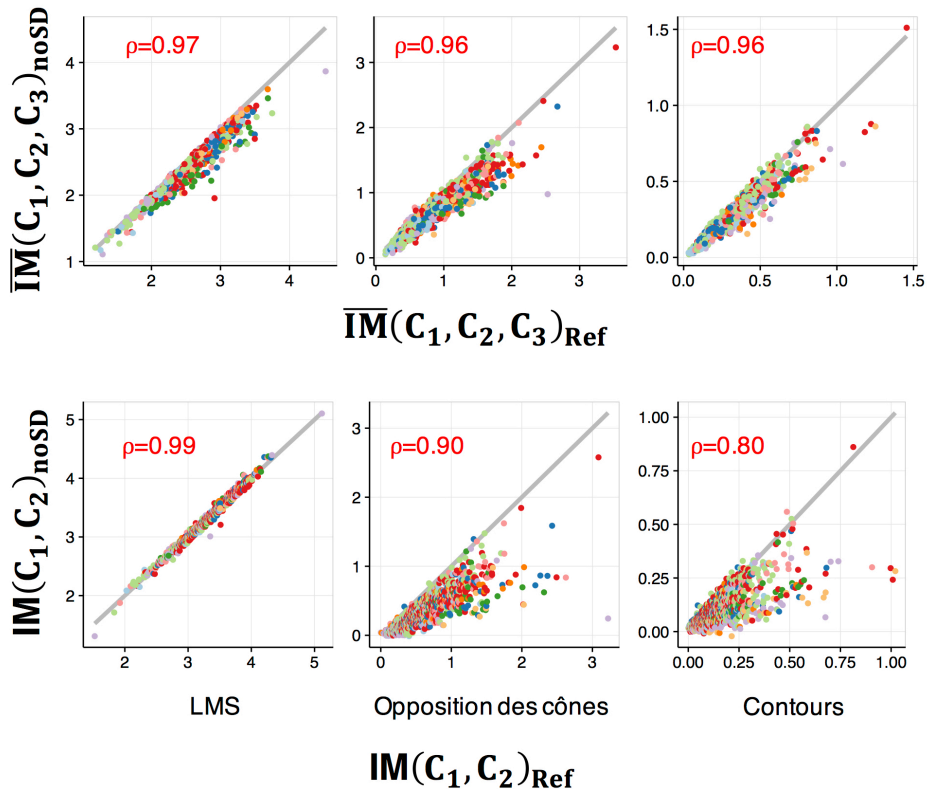
Paramètres de la méthode d'estimation de l'information mutuelle

La valeur du paramètre k semble assez peu impacter nos résultats d'information mutuelle (Fig. 2.18 pour le cas $k=100$). En revanche, l'étape de réduction des données modifie plus sen-

siblement les valeurs d'information mutuelle (Fig. 2.18). Ce résultat n'est pas très surprenant car ce traitement affecte la distribution des données. Il n'est donc pas très pertinent d'inclure une réduction des données dans le calcul d'information mutuelle si nous voulons que l'indice soit représentatif des propriétés de l'image originale.



(a) Effet du paramètre k : Comparaison des valeurs d'information mutuelle obtenues avec le modèle de référence dans le cas $k=25$ (**Ref**) et $k=100$ (**k100**).



(b) Effet de la réduction des données : Comparaison des valeurs d'information mutuelle obtenues avec (**Ref**) ou sans (**noSD**) l'étape de réduction des données.

FIGURE 2.18 – Effet des paramètres de la méthode de calcul d'information mutuelle.

2.2.3 Discussion

2.2.3.1 Synthèse des résultats

Nous avons pu montrer par cette étude que :

- L'indépendance statistique entre les différents canaux, mesurée au travers de l'information mutuelle, variait sensiblement d'une image à l'autre. Nous avons ainsi pu mettre en évidence un artefact dans la méthode décrite par Hansen et Gegenfurtner (2009) et proposer une amélioration de celle-ci en effectuant les calculs d'information mutuelle pour chaque image individuellement.
- La chaîne de traitement proposée est peu sensible au choix de différents paramètres (fonctions fondamentales des cônes, modèle d'adaptation, écran utilisé, nombre de plus proches voisins pour l'estimateur d'information mutuelle) et semble donc fournir une mesure relativement robuste de la quantité de dépendance statistique entre nos canaux.
- Nous avons en revanche pu constater que d'autres paramètres, tels que le choix d'une définition des canaux d'opposition des cônes intégrant ou non un processus d'adaptation modifiait significativement nos résultats. En comparant les valeurs d'information mutuelle obtenues dans chacun des cas, l'étape d'adaptation semble permettre de décorrélérer plus rapidement et plus efficacement l'information entre les différents canaux, qu'elle intervienne au niveau des réponses des cônes ou des mécanismes post-récepteurs.

Nous pouvons donc, en choisissant une définition mécanistique de la chrominance, établir un classement de nos images sur la base de cet indice d'information mutuelle. À notre connaissance, cette analyse est la première à proposer un indice permettant de quantifier l'indépendance statistique des informations portées par les différents canaux post-récepteurs à l'échelle d'une image. Comme nous l'avons vu dans l'introduction de ce chapitre, historiquement, les analyses statistiques sur les scènes naturelles se sont d'abord intéressées au codage efficace de l'information très redondante enregistrée par les cônes, sous forme de canaux post-récepteurs, optimisant sa transmission au système visuel. Quelques études se sont ensuite concentrées sur la comparaison des informations de luminance et de chrominance mais certaines ont utilisé des stimuli artificiels (Buchsbaum et Gottschalk, 1983), un très petit nombre d'images de scènes naturelles (Ruderman et al., 1998 ; Wachtler, Lee et al., 2001 ; Fine et al., 2003) ou des images non calibrées (Heidemann, 2006 ; Zhou et Mel, 2008). De rares travaux ont cherché à quantifier la redondance de l'information portée par les différents canaux en utilisant des mesures telles que l'information mutuelle (Hansen et Gegenfurtner, 2009 ; Peyvandi et al., 2013 ; Foster et al., 2004) ou la corrélation (Johnson, Kingdom et al., 2005) mais dans toutes ces études l'objectif était d'estimer une indépendance globale, à partir d'un grand nombre d'images et/ou dans des conditions d'illumination différentes.

Notre étude est donc la première à mettre en évidence la variabilité inter-images en termes de redondance d'information de luminance et de chrominance. Compte tenu de ce résultat, on peut supposer que l'indépendance de ces informations est également susceptible de varier

au sein même d'une image. Il serait donc intéressant de pouvoir quantifier localement cette indépendance. L'information mutuelle n'est peut-être pas une mesure adaptée à cette analyse car elle nécessite un certain nombre de données pour être estimée de façon suffisamment fiable. De façon plus générale, on peut s'interroger sur la pertinence du choix de l'information mutuelle comme mesure de l'indépendance statistique entre nos variables. Cet indice semble à première vue bien adapté à notre problématique et à l'approche computationnelle de celle-ci mais quelques questions demeurent sur l'information réellement fournie par cet indice :

- Que mesure-t-on exactement ?
- À quoi correspondrait une scène pour laquelle la redondance des informations de luminance et de chrominance serait parfaite ? Nulle ?
- Les valeurs obtenues ne sont-elles pas en partie liées à la présence d'artefacts dans la méthodologie ?

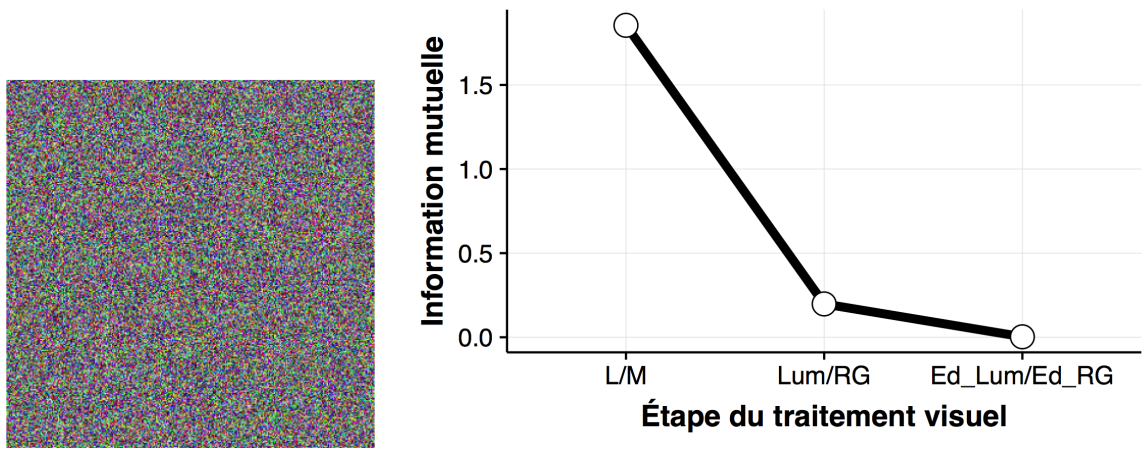
Afin d'essayer de répondre en partie à ces questions, nous nous sommes intéressés au cas d'une image de bruit aléatoire.

2.2.3.2 Image aléatoire

Afin de mieux comprendre ce que l'information mutuelle nous apprend de nos images de scènes naturelles, nous allons considérer le cas d'une image aléatoire. Nous définissons cette image de dimensions 300×300 pixels en attribuant des valeurs aléatoires à chacune des trois dimensions R, G et B de chacun des pixels de l'image. L'image aléatoire obtenue est présentée sur la Figure 2.19, ainsi que les valeurs d'information mutuelle estimées et les distributions des valeurs des pixels sur les canaux C_1 et C_2 à chacune des étapes du traitement visuel.

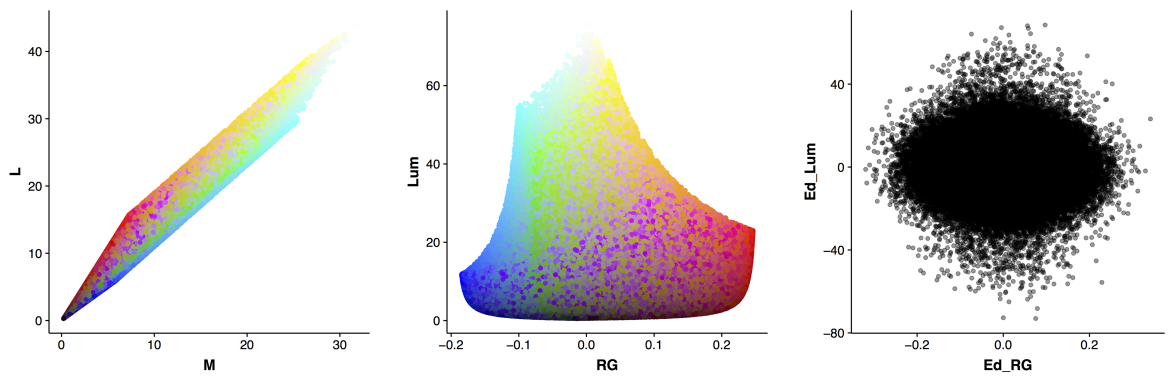
Les distributions de L et M sont fortement corrélées, en raison de la proximité de leurs spectres de sensibilité spectrale (Buchsbaum et Gottschalk, 1983). Même si les mécanismes d'opposition permettent de réduire considérablement la redondance, s'agissant de combinaisons linéaires des valeurs L et M, l'allure de la distribution est également impactée par les propriétés des spectres. Les distributions sont bien moins contraintes après transformation dans l'espace des contours, où l'on réalise des opérations inter-pixels, ceux-ci étant totalement indépendants les uns des autres pour une image aléatoire. L'effet de décroissance de l'information mutuelle le long du traitement visuel semble donc décorréler des propriétés intrinsèques des scènes naturelles et peut également être observé pour une image aléatoire.

De plus, nous travaillons ici avec les valeurs LMS correspondant à l'affichage des images sur notre écran. Comme illustré sur la Figure 2.20, le gamut de l'écran (c'est-à-dire, la partie de l'ensemble des couleurs que l'écran permet de reproduire) pourrait également expliquer en partie cette décroissance, l'espace des couleurs pouvant être représentées étant limité et biaisant donc les distributions de nos variables.



(a) Image aléatoire.

(b) Valeurs d'information mutuelle moyennes à chaque étape du traitement.



(c) Distribution des pixels de l'image dans un repère (L,M).

(d) Distribution des pixels de l'image dans un repère (Lum, RG).

(e) Distribution des pixels de l'image dans un repère (Ed_Lum, Ed_RG).

FIGURE 2.19 – Analyse d'une image aléatoire.

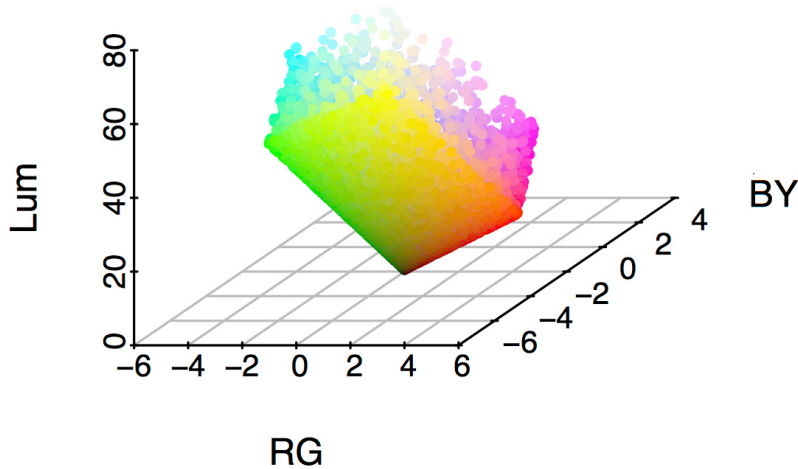


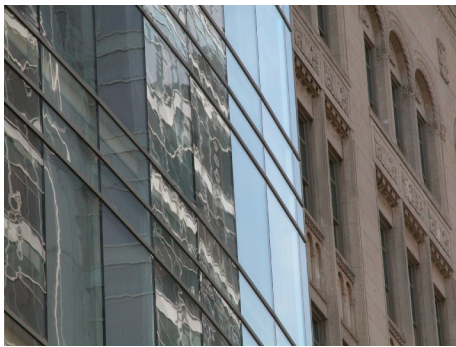
FIGURE 2.20 – Gamut de notre écran de référence (Iiyama Vision Master Pro 513) présenté dans l'espace DKL.

2.2.4 L'information mutuelle : un indice pertinent ?

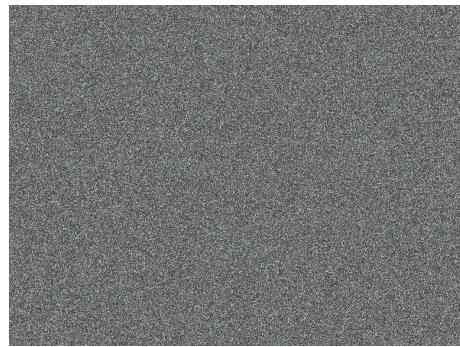
D'autre part, la différence observée entre les distributions des valeurs de luminance et de chrominance Rouge-Vert des pixels des images présentant une faible information mutuelle entre contours et celles présentant une forte information mutuelle semble indiquer que l'indépendance des contours estimée est fortement liée à la distribution des valeurs des pixels. Pour le vérifier, nous avons sélectionné 4 images de faible IM entre contours et 4 images de forte IM, dont nous avons aléatoirement permuté les positions des pixels et mesuré les nouvelles valeurs d'information mutuelle.

Les valeurs d'information mutuelle entre Ed_Lum et Ed_RG pour les images originales et les images aléatoires des deux groupes sont présentées sur la Figure 2.22. On constate que, malgré la permutation des pixels, l'écart entre les valeurs d'information mutuelle des contours estimées pour chacun des deux groupes, reste à peu près constant (valeurs environ 10 fois plus élevées pour le groupe de forte IM que pour le groupe de faible IM). Ce résultat suggère que l'indice d'information mutuelle entre contours ne renseigne pas vraiment sur la structure spatiale des images, mais dépend beaucoup de la distribution des valeurs de luminance et de chrominance des pixels de l'image.

Finalement, nos mesures d'information mutuelle sont très sensibles aux distributions des pixels, et celles-ci semblent être affectées par un grand nombre de paramètres. C'est notamment le cas de l'étape de réduction que nous introduisons avant l'estimation d'information mutuelle, ou encore du choix de la conversion dans l'espace d'opposition des cônes : une transformation log ou une normalisation divise de la chrominance par la luminance modifie significativement la structure des données, et donc les valeurs d'information mutuelle.



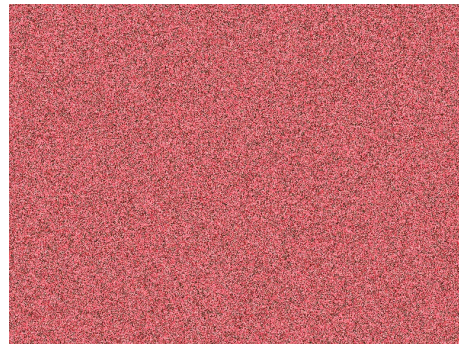
(a) Image 1: faible information mutuelle entre Ed_Lum et Ed_RG.



(b) Image aléatoire construite par permutation des pixels de l'image 1.



(c) Image 2: forte information mutuelle entre Ed_Lum et Ed_RG.



(d) Image aléatoire construite par permutation des pixels de l'image 2.

FIGURE 2.21 – Exemples d'images aléatoires obtenues à partir des images originales.

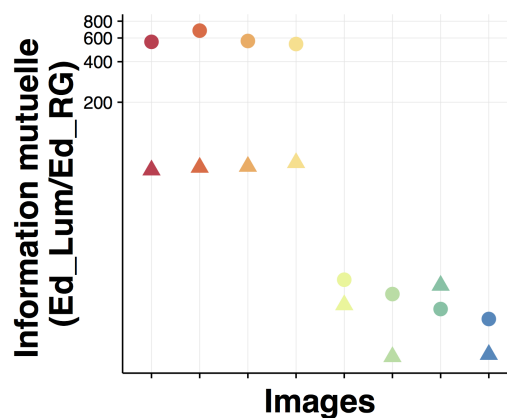


FIGURE 2.22 – Information mutuelle ($\cdot 10^3$, en échelle logarithmique) entre contours de luminosité et de chrominance Rouge-Vert pour 8 images originales (●) et des images aléatoires construites à partir de ces 8 images par permutation aléatoire des positions des pixels (▲).

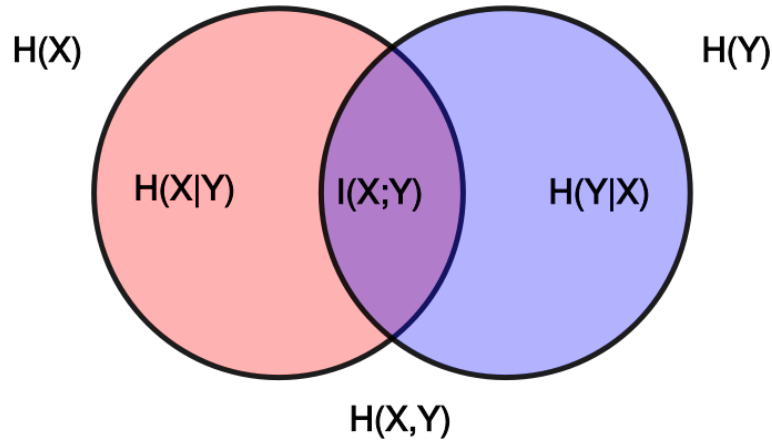


FIGURE 2.23 – Illustration des relations entre les entropies de deux variables X et Y , $H(X)$ et $H(Y)$, leur information mutuelle, $I(X, Y)$, leurs entropies conditionnelles, $H(X | Y)$ et $H(Y | X)$, et leur entropie jointe, $H(X, Y)$.

Par définition, l'information mutuelle dépend des entropies des deux variables considérées et de leur entropie conjointe : $I(X, Y) = H(X) + H(Y) - H(X, Y)$, où $H(X)$ et $H(Y)$ sont les entropies des distributions de X et Y et $H(X, Y)$ est l'entropie conjointe de X et Y (Fig. 2.23). Une image pour laquelle l'entropie de chacune des variables X et Y est plus élevée aura donc une information mutuelle plus élevée. On peut se demander si nos résultats ne dépendent donc pas en grande partie de l'entropie de chacune des variables, et finalement assez peu de leur entropie conjointe. Il pourrait être intéressant de considérer d'autres indices, comme par exemple l'information mutuelle normalisée (IMN) : $IMN(X, Y) = \frac{2 \times I(X, Y)}{[H(X) + H(Y)]}$ qui renseigne sur la part de l'information totale étant commune aux deux variables ou encore l'entropie conditionnelle $H(Y | X) = H(Y) - I(X, Y)$ qui mesure l'incertitude restant sur la valeur de Y lorsque la valeur de X est connue. Ne disposant pas d'algorithme permettant d'estimer de façon fiable les entropies seules, nous n'avons pas poussé ces investigations mais elles mériteraient de l'être.

2.3 Conclusion

Finalement, nous avons pu montrer par cette étude que l'analyse des statistiques des scènes naturelles pouvaient être très informative pour mesurer l'apport de l'information de chrominance par rapport à celle de luminance. Il faut, cependant, fixer avec attention de nombreux paramètres et interroger les choix fait aux différentes étapes du traitement, et leurs conséquences sur les distributions des variables, pour obtenir un indice pertinent. Notre étude a notamment mis en lumière l'influence des processus d'adaptation sur les distributions des

valeurs de luminance et de chrominance des pixels, et leurs conséquences sur les mesures d'information mutuelle estimées. Afin de proposer une définition mécanistique de la chrominance adaptée à la perception des scènes naturelles, nous allons, dans le chapitre suivant, questionner l'efficacité des différentes conversions dans l'espace d'opposition des cônes proposées ici à expliquer des données psychophysiques concernant la perception de la couleur dans des scènes naturelles.

3

Un modèle mécanistique de la chrominance pour la perception visuelle des scènes naturelles

Le seul véritable voyage, le seul bain de Jouvence, ce ne serait pas d'aller vers de nouveaux paysages, mais d'avoir d'autres yeux, de voir l'univers avec les yeux d'un autre, de cent autres, de voir les cent univers que chacun d'eux voit, que chacun d'eux est.

Marcel Proust, *La Prisonnière*, 1923

Sommaire

3.1	Introduction	80
3.2	Expérience 1 - Détection de couleur	84
3.2.1	Hypothèse	84
3.2.2	Design expérimental	84
3.2.3	Analyse des résultats	89
3.2.4	Résultats	96
3.3	Observateur décalé	99
3.4	Expérience 2 - Détection de luminance	103
3.4.1	Hypothèse	103
3.4.2	Design expérimental	103
3.4.3	Résultats	104
3.5	Expérience 3 - Détection de couleur (contrôle)	106

3.5.1	Hypothèse	106
3.5.2	Design expérimental	106
3.5.3	Résultats	107
3.6	Discussion	108
3.6.1	Synthèse des résultats	108
3.6.2	Cas du mécanisme logarithmique - Adaptation des photorécepteurs	109
3.6.3	Perspectives	109
3.7	Conclusion	111

Les travaux présentés dans ce chapitre ont fait l'objet d'un article soumis au journal *Plos Computational Biology* : "*Detecting color in natural scenes : evidence for early interactions between luminance and chrominance*", Camille Breuil, Nathalie Guyader et Simon Barthelmé, 2018.

Dans le chapitre précédent nous avons proposé une méthode pour quantifier l'indépendance de l'information de luminance et de l'information de chrominance dans des images de scènes naturelles (modèle de référence pour l'estimation de l'information mutuelle). Nos résultats ont mis en évidence l'influence d'un certain nombre de paramètres sur les valeurs obtenues et ont, notamment, souligné l'importance de la définition de la chrominance utilisée. Nous testions dans le chapitre précédent trois modèles de conversion dans l'espace d'opposition des cônes, correspondant à trois hypothèses mécanistiques sur l'encodage de l'information chromatique par le système visuel. Nos résultats montrent, du point de vue de la théorie de l'information, un avantage pour les mécanismes intégrant un processus d'adaptation (que ce soit au niveau des photorécepteurs, avec le modèle logarithmique, ou des canaux post-récepteurs, avec le modèle de normalisation divisive), qui permettent de décorréler plus efficacement signaux de luminance et de chrominance. Dans ce chapitre, nous allons chercher à vérifier expérimentalement si des données psychophysiques de perception de la couleur dans des images de scènes naturelles soutiennent effectivement l'existence de l'un ou l'autre de ces mécanismes. Nous réalisons pour cela une série d'expériences dans lesquelles nous avons manipulé les informations de luminance et de chrominance de scènes naturelles, et mesuré les effets de ces manipulations dans des tâches de détection. Nous proposons également un modèle d'encodage de l'information et de prise de décision nous permettant de comparer nos hypothèses mécanistiques.

3.1 Introduction

Afin d'étudier le rôle de l'information chromatique dans le traitement visuel des scènes naturelles, il est nécessaire de disposer d'un modèle de perception de la couleur pertinent dans ce contexte. Pour cela, il faut avant tout définir correctement à quoi correspond l'information de chrominance extraite par le système visuel dans le cas des scènes naturelles. Comme nous l'avons vu dans le chapitre précédent, une première question est de savoir si une définition linéaire est suffisante, ou s'il est nécessaire d'intégrer des étapes non-linéaires, sous forme notamment de normalisation divisive de la chrominance par la luminance.

La question de savoir si les mécanismes post-récepteurs pouvaient être décrits comme de simples combinaisons linéaires des réponses des cônes a été formulée quand les processus d'adaptation des photorécepteurs et la chaîne de traitement post-récepteurs étaient encore méconnus. Compte tenu des connaissances actuelles sur ces sujets, un mécanisme purement linéaire peut, à première vue, sembler une hypothèse optimiste. Elle suppose, en effet, un niveau d'adaptation constant pour les trois types de cônes. Plusieurs études, en psychophysique, ont montré que le modèle linéaire basé sur les excitations des cônes échouait à rendre compte de données expérimentales concernant la perception de la couleur (Ingling, Barley et al., 1996 pour le mécanisme Rouge-Vert, Werner et Wooten, 1979 ; Larimer et al., 1975 ; Knoblauch et Shevell, 2001 pour le mécanisme Bleu-Jaune)

Dans la majorité de ces expériences, ayant utilisé des stimuli simples pour étudier les mécanismes post-récepteurs, ceux-ci étaient décrits en termes de combinaisons linéaires de contrastes de cônes relativement au fond neutre, et non directement en termes d'excitation des cônes (Switkes et al., 1988 ; Cole, Stromeyer III et al., 1990 ; Cole, Hine et al., 1993 ; Mullen et Losada, 1994). Le canal de chrominance Rouge-Vert, par exemple, est défini dans cet espace de représentation par $RG = \frac{\Delta L}{L} - \frac{\Delta M}{M}$, où les dénominateurs correspondent aux niveaux d'excitation respectifs des cônes L et M résultant de la présentation du fond neutre d'adaptation et les numérateurs à la variation de ces niveaux d'excitation lors de la présentation d'un stimulus simple sur le fond neutre de l'écran.

Bien que la question de la définition des mécanismes post-récepteurs à partir des signaux issus des cônes ait fait l'objet de nombreuses études sur des stimuli simples, tels que des *patches* colorés présentés sur des fonds uniformes, le cas des scènes naturelles a été très peu étudié et nécessite des investigations spécifiques. Il n'est, en effet, pas évident que des théories basées sur des données obtenues à partir de stimuli simples puissent se généraliser à la prédiction des processus entrant en jeu dans le traitement de stimuli plus complexes. La perception des scènes présente de nombreux challenges que ne permettent pas d'appréhender les études se basant uniquement sur des stimuli artificiels, ce qui a conduit ces dernières années à une remise en question de la pertinence de ces approches pour comprendre le fonctionnement du système visuel dans un environnement naturel (Olshausen et Field, 2005) et suscité des controverses (Felsen et Dan, 2005 ; Rust et Movshon, 2005).

Comme souligné par Brainard (1996) ou Olmos et Kingdom (2004a), les définitions basées sur les contrastes de cônes sont pertinentes lorsque l'on considère des stimuli simples mais elles se révèlent inadaptées pour l'étude de stimuli plus complexes, tels que les scènes naturelles. Le niveau d'adaptation est, en effet, susceptible de varier au sein de l'image (notamment en raison de la présence d'ombres), rendant difficile le choix du facteur de normalisation. Il a ainsi été proposé de normaliser l'excitation de chacun des cônes par sa valeur moyenne sur l'ensemble du stimulus (Buchsbaum, 1980 ; D'Zmura et Lennie, 1986 ; Brainard et Wandell, 1986 ; Ruderman et al., 1998) ou encore par sa valeur maximale dans le stimulus (Land et McCann, 1971), sans que nous n'ayons pour autant aucune garantie que l'un ou l'autre de ces choix modélise correctement le traitement opéré par le système visuel.

De plus, le signal enregistré par les cônes étant, dans un contexte naturel, inévitablement très bruité (caractère aléatoire de l'absorption, chimie des photorécepteurs,...) on peut s'in-

terroger sur la pertinence de réaliser une adaptation précoce sur les réponses d'excitation des photorécepteurs. Contrairement à la façon dont elle souvent présentée, l'adaptation est un processus en partie stochastique. Réalisée sur un signal bruité, elle risque donc de causer des fluctuations de gain indépendantes du signal initial, amplifiant ainsi le bruit dans le signal transmis. Dans ces conditions, l'adaptation pourrait donc s'avérer finalement délétère, et il serait alors préférable pour le système visuel d'accepter qu'il y a une mauvaise adéquation entre le signal transmis par les cônes et la distribution du stimulus d'origine, plutôt que de s'adapter. C'est en particulier pertinent si l'adaptation peut se faire plus tardivement dans le traitement, bénéficiant ainsi du regroupement d'un plus grand nombre de réponses neuronales (Rieke et Rudd, 2009).

C'est notamment pour pallier à ces difficultés que Párraga, Brelstaff et al. (1998) ont proposé, pour l'étude des scènes naturelles, une définition des canaux de chrominance basée sur les excitations des cônes et intégrant une normalisation divisive. Comme expliqué ci-dessous, on peut dire que ces définitions permettent de s'affranchir des ombres, au sens où elles décorrèlent les variations de chrominance des variations de luminance. C'est principalement cette raison qui a motivé leur définition, mais des arguments en faveur de l'existence de mécanismes de normalisation divisive de ce type dans les systèmes sensoriels ont depuis été exposés dans d'autres études (Schwartz et Simoncelli, 2001 ; Solomon, 2005, voir Carandini et Heeger, 2012 pour une revue).

Afin de mieux comprendre en quoi les définitions de chrominance intégrant une normalisation divisive peuvent être qualifiées d'*éliminatrices des ombres*, revenons aux relations entre excitations des cônes et propriétés de réflectance et d'absorbance des objets. Les excitations respectives de chacun des trois types de cônes sont données par :

$$\begin{aligned}
 L &= \int_{\lambda} I(\lambda)R(\lambda)l(\lambda)d\lambda \\
 M &= \int_{\lambda} I(\lambda)R(\lambda)m(\lambda)d\lambda \\
 S &= \int_{\lambda} I(\lambda)R(\lambda)s(\lambda)d\lambda
 \end{aligned}
 \tag{3.1}$$

où l , m et s sont les fonctions de sensibilité des cônes de Smith et Pokorny (1975), λ est la longueur d'onde, $I(\lambda)$ est le spectre de radiance de l'illuminant et $R(\lambda)$ le spectre de réflectance de l'objet.

Dans l'exemple présenté sur la Figure 3.1, les propriétés de réflectance des pixels A et B sont identiques. En revanche, les illuminations sont différentes car B, contrairement à A, se situe dans l'ombre portée du cube. Pour simplifier, supposons que le spectre d'illumination est $I(\lambda)$ pour le pixel A et $\frac{I(\lambda)}{2}$ pour le pixel B. Avec les définitions linéaires des mécanismes post-récepteurs, l'information de luminance et de chrominance Rouge-Vert du pixel A sont donc :

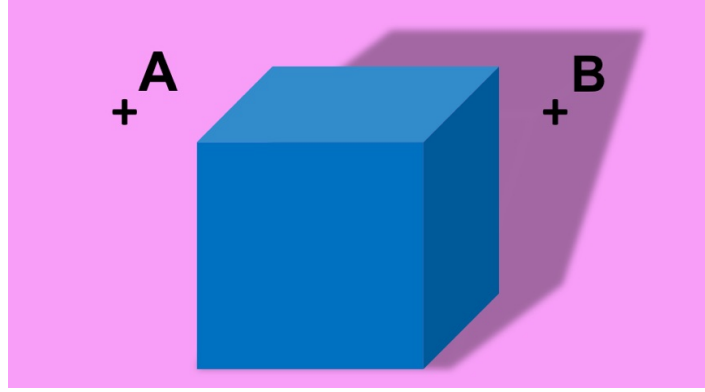


FIGURE 3.1 – Effet d'élimination des ombres de la normalisation divisive

$$\begin{aligned}
 Lum(A) &= L(A) + M(A) = \int_{\lambda} I(\lambda)R(\lambda)l(\lambda)d\lambda + \int_{\lambda} I(\lambda)R(\lambda)m(\lambda)d\lambda \\
 RG(A) &= L(A) - M(A) = \int_{\lambda} I(\lambda)R(\lambda)l(\lambda)d\lambda - \int_{\lambda} I(\lambda)R(\lambda)m(\lambda)d\lambda
 \end{aligned} \tag{3.2}$$

et celles du pixels B :

$$\begin{aligned}
 Lum(B) &= L(B) + M(B) = \int_{\lambda} \frac{I(\lambda)}{2}R(\lambda)l(\lambda)d\lambda + \int_{\lambda} \frac{I(\lambda)}{2}R(\lambda)m(\lambda)d\lambda \\
 RG(B) &= L(B) - M(B) = \int_{\lambda} \frac{I(\lambda)}{2}R(\lambda)l(\lambda)d\lambda - \int_{\lambda} \frac{I(\lambda)}{2}R(\lambda)m(\lambda)d\lambda
 \end{aligned} \tag{3.3}$$

Nous avons donc $Lum(B) = \frac{Lum(A)}{2}$ et $RG(B) = \frac{RG(A)}{2}$, mais le ratio reste constant : $\frac{RG(B)}{Lum(B)} = \frac{RG(A)}{Lum(A)}$. En utilisant la définition de la chrominance avec normalisation divisive, les pixels A et B ont donc la même chrominance, et ce malgré la présence de l'ombre. On peut donc dire que cette définition enlève les ombres, et présume qu'un objet dont les propriétés de réflectance sont constantes, sera perçu comme uniformément coloré, en dépit de la présence éventuelle d'ombres.

Plusieurs études ont utilisé les définitions proposées par Párraga, Brelstaff et al. (1998) pour étudier les statistiques des images naturelles (Párraga, Troscianko et al., 2002 ; Johnson, Kingdom et al., 2005 ; Hansen et Gegenfurtner, 2009) mais, à notre connaissance, aucune étude ne s'est à ce jour attachée à évaluer la pertinence de ces définitions pour expliquer des données expérimentales concernant la perception de scènes naturelles. Nous proposons donc ici une étude visant à comparer la capacité des définitions linéaires à celle des définitions divisives à expliquer des données psychophysiques de détection de luminance et de chrominance dans des scènes naturelles.

Nous avons pour cela mené 3 expériences dans lesquelles nous avons manipulé l'information de luminance et de chrominance dans des images de scènes naturelles. Nous proposons ensuite

un modèle général d'encodage de l'information et de prise de décision pour notre tâche nous permettant de tester la validité de nos deux hypothèses mécanistiques.

3.2 Expérience 1 - Détection de couleur

Dans cette première expérience, nous avons utilisé un paradigme de choix forcé entre deux propositions (2AFC) avec une méthode des stimuli constants pour mesurer les seuils de détection de couleur des participants en faisant varier le niveau de luminance moyen dans des images de scènes naturelles.

3.2.1 Hypothèse

Nous formulons l'hypothèse que, s'il y a effectivement une étape de normalisation divisive de l'information de chrominance par la luminance au cours du traitement visuel, alors la chrominance perçue devrait diminuer lorsque le niveau moyen de luminance augmente, et les seuils de détection de couleur mesurés devraient en conséquence augmenter en fonction du niveau de luminance moyen.

3.2.2 Design expérimental

3.2.2.1 Participants

5 personnes ont participé à cette expérience : 2 femmes et 3 hommes, âgés de 20 à 40 ans. Ils avaient tous une acuité visuelle normale ou corrigée à la normale. Leur vision des couleurs a été évaluée avec le test Farnsworth–Munsell 100-Hue et était normale. Avant l'expérience, deux des sujets ont participé à une étude pilote visant à déterminer les paramètres adaptés pour que les performances couvrent une gamme approximativement comprise en 55 et 95 % de réussite. Chacun des participants a donné son consentement éclairé à prendre part à l'expérience.

3.2.2.2 Dispositif

Les participants étaient assis dans une pièce sombre à 30,5 cm d'un écran à tube cathodique (CRT) de 21 pouces (Iiyama Vision Master Pro 513), avec une résolution de 1024×768 pixels et un taux de rafraîchissement de 100 Hz. Leurs têtes étaient confortablement stabilisées au niveau du menton par un petit coussinet. L'écran avait été préalablement calibré à l'aide d'un spectro-radiomètre CS Konica Minolta. La description détaillée de la procédure de calibration ainsi que les valeurs des paramètres de l'écran mesurés sont fournis en annexe (Annexes A et B). Tous les stimuli étaient présentés sur un fond gris neutre de coordonnées $RGB=[127,127,127]$ et chacun couvrait approximativement 18×18 degrés d'angle visuel pour l'observateur. L'écran était systématiquement allumé au moins une demi-heure avant le début

de chaque session. L'expérience a été programmée avec la Psychtoolbox de Matlab (Brainard, 1997).

3.2.2.3 Protocole

Chacun des essais de l'expérience commençait par l'affichage d'une croix de fixation au centre de l'écran durant 500 ms. Deux stimuli étaient ensuite présentés simultanément de chaque côté de l'écran, à une excentricité constante de 3 degrés d'angle visuel (mesurée entre le centre de l'écran et le bord le plus proche du stimulus) pour une durée de 500 ms. À chaque essai, les deux stimuli présentés étaient générés à partir de la même image de scène naturelle. L'un d'eux, contenait uniquement une partie de l'information de luminance de l'image originale (image de référence), le second, contenait la même information de luminance, plus une petite proportion de l'information de chrominance Rouge-Vert de l'image originale (image test). Il était demandé aux participants d'indiquer laquelle des deux images était colorée en pressant la touche gauche ou droite du clavier. Après 500 ms de présentation, ou dès que le sujet avait donné sa réponse, les stimuli étaient masqués. Un feedback sonore indiquait aux participants une réponse correcte. Une brève pause de 500 ms séparait chaque essai. Une illustration schématique du déroulement d'un essai est présenté sur la figure 3.2. Chaque participant a effectué 3 sessions de l'expérience à des jours différents, chacune des sessions étant constituée de 750 essais.

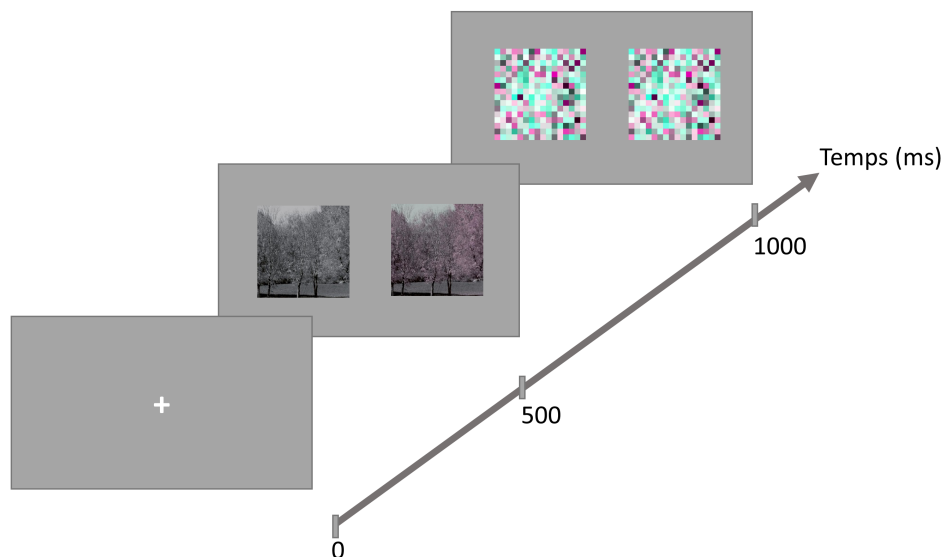


FIGURE 3.2 – Déroulement d'un essai. Après l'affichage d'une croix de fixation au centre de l'écran, image de référence et image test sont présentées de part et d'autre de l'écran durant 500 ms avant d'être masquées.

3.2.2.4 Stimuli

Les stimuli utilisés pour cette expérience ont été générés à partir de 10 images de la base d'images couleur calibrées McGill (<http://tabby.vision.mcgill.ca>, Olmos et Kingdom, 2004a). Ces 10 images ont été choisies dans différentes catégories (Forêts, Fleurs, Scènes artificielles et Animaux), et de sorte à couvrir un large spectre de valeurs de luminance et chrominance moyennes (Fig. 3.3).

Les images RGB originales, enregistrées dans le format TIFF et de dimensions 768×576 pixels, ont dans un premier temps été redimensionnées (384×384 pixels) à l'aide d'une méthode d'interpolation bicubique. Elles ont ensuite été converties dans l'espace LMS d'absorption des cônes en utilisant la procédure adaptée à notre écran d'expérimentation (Annexe B). Les réponses des cônes ont été calculées en utilisant les fonctions fondamentales de Smith et Pokorny (1975)). Les images ont finalement été converties dans l'espace d'opposition des cônes DKL (Derrington et al., 1984) en utilisant les équations suivantes :

$$\begin{aligned} Lum &= L + M \\ RG &= L - \alpha M \\ BY &= 2S - \beta(L + M) \end{aligned} \quad (3.4)$$

où (α, β) est un couple de paramètres déterminé par les propriétés de l'écran sur lequel les stimuli sont présentés.

Tous les stimuli utilisés pour cette expérience ont été générés dans cet espace tridimensionnel à partir des 10 images originales sélectionnées :

- Nous avons testé 3 conditions de luminance, obtenues en faisant varier paramétriquement le ratio de luminance initiale de chacun des pixels ($\eta = 50\%, 75\%, 100\%$).
- La quantité d'information de chrominance Rouge-Vert contenue dans les images test a été modulée en faisant varier sur 5 niveaux le paramètre ζ , correspondant au ratio de chrominance initiale de chacun des pixels ($\zeta = 2\%, 4\%, 6\%, 8\%, 10\%$).

La Figure 3.4 montre l'exemple de tous les stimuli générés à partir de l'une des 10 images originales. Pour l'ensemble de nos stimuli, nous avons fixé à zéro la chrominance Bleu-Jaune (BY), afin de simplifier le design expérimental et l'analyse des résultats. Pour la suite, nous ne ferons donc plus mention de la chrominance Bleu-Jaune.

Finalement, pour chaque essai, si l'on considère un pixel t , il peut être représenté par ses coordonnées $\begin{pmatrix} Lum(t) \\ RG(t) \end{pmatrix}$ dans l'espace DKL, à la fois dans l'image de référence ($DKL_r(t)$) et dans l'image test ($DKL_t(t)$) :

$$DKL_r(t) = \begin{pmatrix} \eta Lum_i(t) \\ 0 \end{pmatrix} \text{ et } DKL_t(t) = \begin{pmatrix} \eta Lum_i(t) \\ \zeta RG_i(t) \end{pmatrix} \quad (3.5)$$

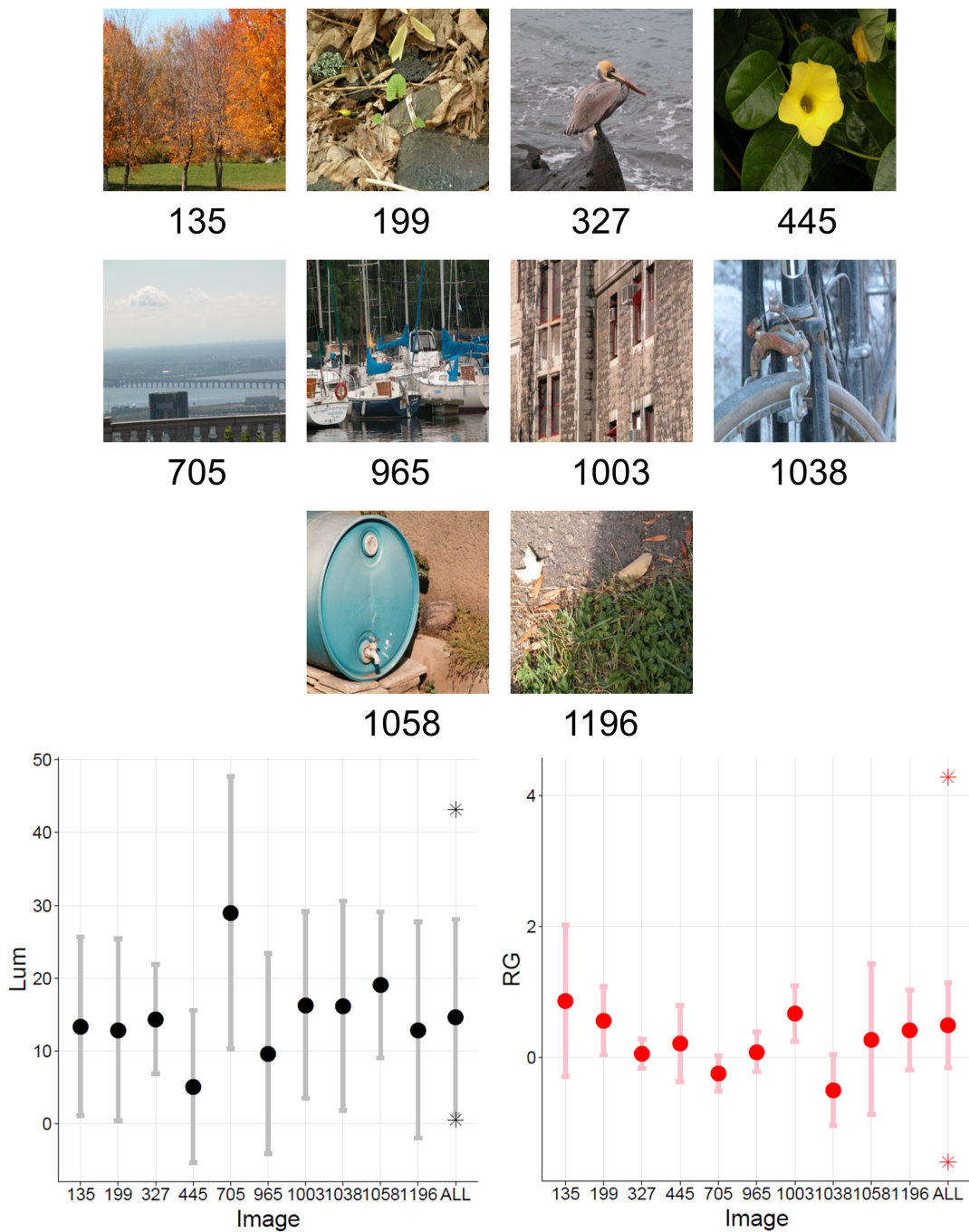


FIGURE 3.3 – Les 10 images originales redimensionnées, présentées avec leurs valeurs moyennes de luminance (Lum) et de chrominance Rouge-Vert (RG) (définies dans l'espace DKL). Les barres d'erreur correspondent à l'écart-type sur l'ensemble des pixels d'une image. ALL : Valeurs moyennes sur l'ensemble des images de la base McGill. * : Valeurs pour les images situées aux extrema des distributions.

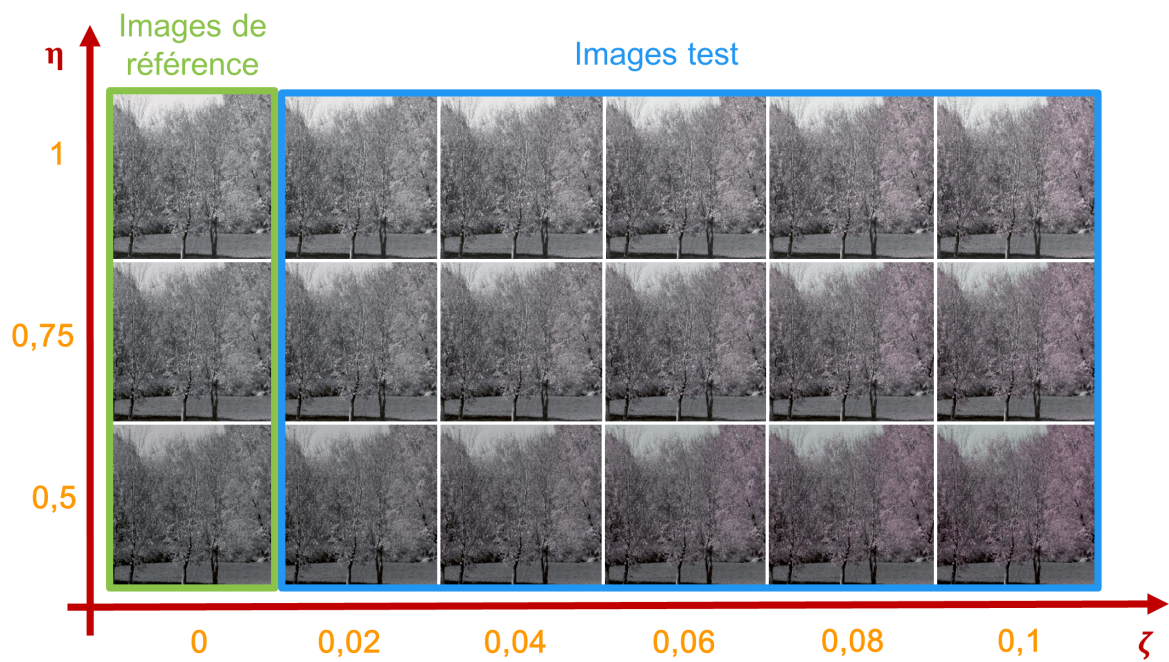


FIGURE 3.4 – Exemple des stimuli obtenus à partir de l'une des 10 images originales, en faisant varier les paramètres η et ζ . Les images de référence ne contiennent qu'un ratio η de l'information de luminance initiale de l'image. Les images test correspondantes contiennent le même ratio η de l'information de luminance initiale, plus un ratio ζ de l'information de chrominance initiale de l'image.

Avec $Lum_i(t)$ et $RG_i(t)$ la luminance et la chrominance initiales du pixel t dans l'image originale, exprimées dans l'espace DKL.

3.2.3 Analyse des résultats

Nous avons mesuré pour chacun des participants sa performance, c'est-à-dire la proportion d'images colorées correctement identifiées dans les différentes conditions de luminance de l'expérience. D'après notre hypothèse, si la chrominance perçue par l'observateur est effectivement divisée par la luminance, les performances devraient donc faiblir lorsque le niveau moyen de luminance augmente. Nous nous intéresserons donc, dans un premier temps, aux performances de détection de chacun des participants, pour chacun des trois niveaux de luminance séparément.

Il nous faudra ensuite nous interroger sur la façon dont l'information contenue dans les stimuli a pu être traitée par le système visuel pour expliquer nos données expérimentales. Nous proposons pour cela un modèle d'encodage de l'information et de prise de décision décrit ci-dessous.

3.2.3.1 Modèle

Nous présentons ici un modèle pour expliquer comment un observateur réalise notre expérience. Notre tâche de détection implique deux étapes principales de traitement :

- **Encodage de l'information.** Dans un premier temps, l'information visuelle associée à chacun des deux stimuli présentés doit être encodée sous forme d'information sensorielle accessible à l'observateur.
- **Prise de décision.** Dans un second temps, l'observateur doit prendre une décision concernant lequel des deux stimuli il perçoit comme coloré, en se basant sur l'information sensorielle précédemment acquise.

Afin de développer notre modèle, nous considérons le cas d'un observateur bayésien idéal (Green et Swets, 1988 ; Geisler, 2003), faisant le meilleur usage possible de l'information disponible, compte tenu de connaissances préalables (*a priori*) et sous certaines contraintes.

Encodage de l'information

Nous supposons que chaque pixel d'un stimulus est, dans un premier temps, encodé individuellement dans l'espace sensoriel de l'observateur. Nous modélisons cette étape par une fonction de transformation interne, φ , et nous supposons de plus que chacune des représentations sensorielles associées aux pixels est perturbée par un bruit interne. Concrètement, cette fonction φ correspond à la conversion d'un pixel de l'espace DKL à l'espace sensoriel de l'observateur.

Pour chaque essai, les représentations sensorielles d'un pixel t dans l'image de référence (s_r) et dans l'image test (s_t) sont les suivantes :

$$s_r(t) = \varphi(b(t)) + \varepsilon_0(t) \quad (3.6)$$

et

$$s_t(t) = \varphi(b(t) + g(t)) + \varepsilon_1(t) \quad (3.7)$$

où φ est la fonction de transformation interne de l'observateur, b le signal de base commun aux deux stimuli et g le signal ajouté au stimulus test par rapport au stimulus de référence. ε_0 et ε_1 sont les bruits internes affectant chacune des deux représentations sensorielles.

Dans le cadre de notre expérience, le signal b correspond à l'information de luminance présente dans les deux stimuli et le signal g à l'information de chrominance ajoutée au stimulus test. Un exemple de signaux b et g pour un des essais de notre expérience et une représentation schématique des étapes d'encodage de ces signaux dans l'espace sensoriel d'un observateur sont présentés sur la Figure 3.5.

Prise de décision

La tâche de l'observateur consiste à prendre une décision concernant lequel des deux stimuli qui lui sont présentés contient le signal ajouté g ou, en d'autres termes, lequel correspond à la représentation s_t . Intéressons nous à l'information dont dispose *a priori* l'observateur pour prendre sa décision. Contrairement à une tâche de détection classique, ici, le signal ajouté g est *inconnu*, car un observateur ne peut pas *a priori* prédire à partir du signal de luminance l'allure du signal de chrominance. Pour compenser cette méconnaissance de la nature exacte du signal, l'observateur peut, en revanche, avoir un modèle *statistique* du signal, représentant un *a priori* sur la distribution du signal de chrominance. Mathématiquement parlant, ceci revient à estimer la probabilité d'observer le signal s_t , compte tenu de l'ensemble des valeurs possibles pour le signal g . Ce problème est insoluble si l'on considère l'intégralité des signaux possibles mais, nous allons nous placer ici dans le cas d'un signal gaussien. Nous reviendrons un peu plus loin sur les implications de cette hypothèse. Par simplicité, nous supposons donc ici que le signal g est du bruit blanc, c'est-à-dire que pour chaque pixel t , $g(t)$ a une distribution normale ($g \sim \mathcal{N}(0, \tau^2)$), et que les pixels sont indépendants les uns des autres et de $b(t)$.

Par ailleurs, nous supposons également qu'en approximant la quantité $\varphi(b(t) + g(t))$ par un développement limité d'ordre 1 ($\varphi(b(t) + g(t)) = \varphi(b(t)) + g(t)\varphi'(b(t)) + \varepsilon$), l'erreur (ε) que nous faisons est toujours faible comparée au bruit interne ($\varepsilon_1(t)$), ce qui nous permet d'utiliser légitimement cette estimation. Nous nous en servons pour calculer la distribution de la variable de décision et nous noterons par la suite $f(t) = g(t)\varphi'(b(t))$ pour simplifier les équations.

Nous considérons ici le cas d'un observateur bayésien idéal qui réalise la tâche de façon optimale, compte tenu de l'information disponible et des contraintes existantes. La meilleure façon de décider laquelle des deux images présentées est colorée est de calculer une variable

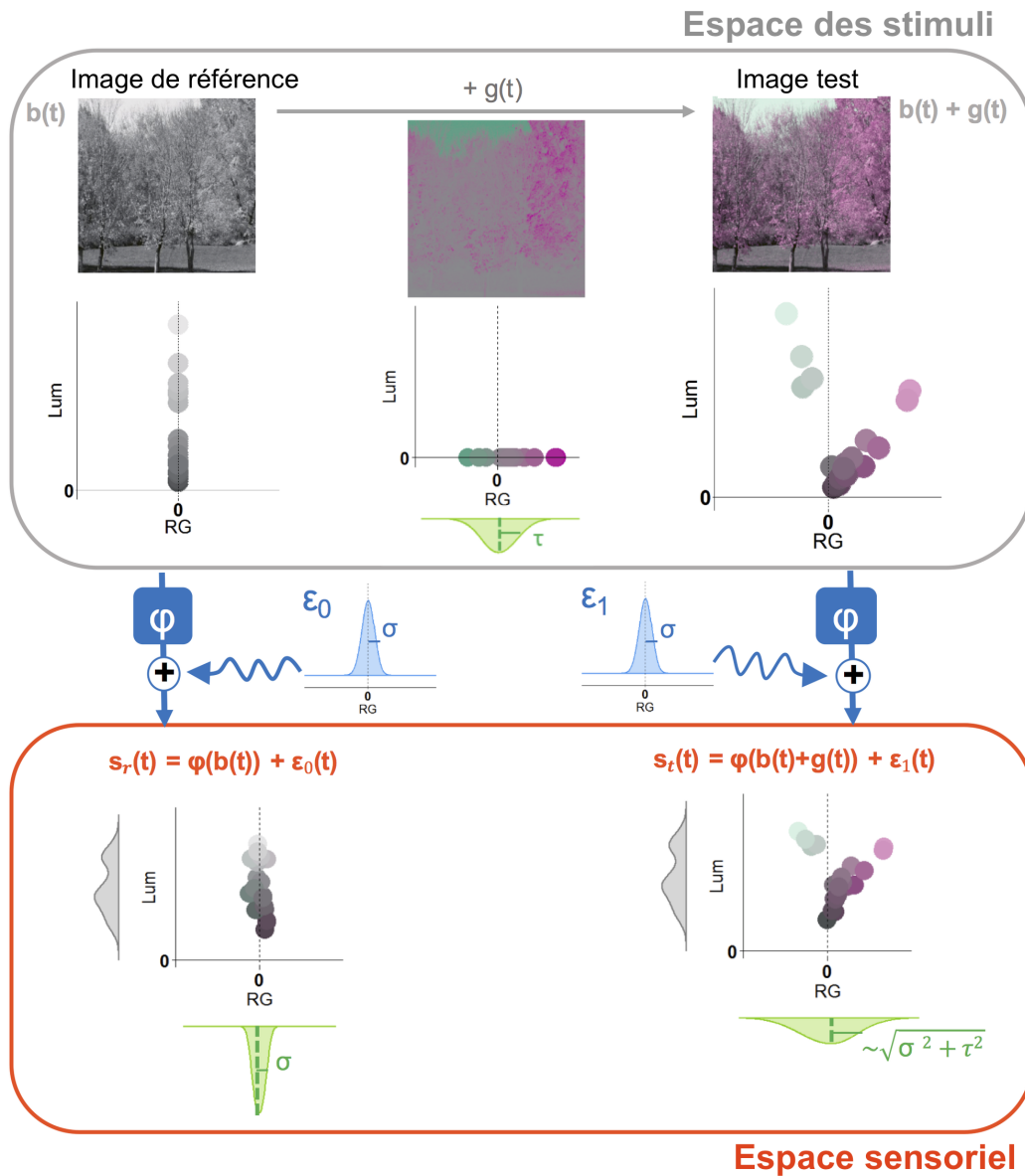


FIGURE 3.5 – Représentation schématique des étapes d’encodage des stimuli dans l’espace sensoriel d’un observateur. Les signaux associés à l’image de référence et à l’image test sont représentés par b et $b + g$ dans l’espace des stimuli. Nous supposons que le signal g peut être considéré comme normalement distribué ($g \sim \mathcal{N}(0, \tau^2)$). Chaque pixel t , représenté par deux coordonnées (RG et Lum) dans l’espace des stimuli, va être converti par la fonction de transformation interne, φ , dans l’espace sensoriel de l’observateur. On suppose que chacune de ces représentations est affectée d’un bruit interne, ε_0 ou ε_1 ($\varepsilon_0, \varepsilon_1 \sim \mathcal{N}(0, \sigma^2)$). En conséquence, les représentations sensorielles associées à chacun des stimuli sont statistiquement identiques en termes d’information de luminance, mais l’information de chrominance sensorielle est différente. Nous supposons donc, que c’est cette information que l’observateur utilise pour prendre sa décision, et qu’il le fait en calculant une variable de décision, μ , qui somme ses observations sur l’ensemble des pixels de l’image.

de décision, μ , correspondant au ratio des probabilités des deux hypothèses :

$$\mu = \frac{p(L | s_r, s_t)}{p(R | s_r, s_t)} \quad (3.8)$$

avec L le cas où l'image test à gauche et R le cas où l'image test est à droite, étant données les observations sensorielles (s_r, s_t) . Si $\mu > 1$ (resp. $\mu < 1$), l'observateur répond que l'image test est à gauche (resp. à droite).

Dans le cas d'un essai où l'image test est présentée à gauche, l'observateur répondra donc juste lorsque $\mu > 1$ (i.e. $\log(\mu) > 0$). La probabilité qu'il détecte l'image test correspond au nombre de fois que ceci se produit ($p(\log(\mu)) > 0$).

En raison du bruit interne, à chaque essai, les signaux $s_r(t)$ et $s_t(t)$ sont normalement distribués autour des moyennes respectives $\varphi(b(t))$ et $\varphi(b(t)) + f(t)$, avec une variance σ^2 .

Pour un essai donné, la probabilité que l'image de référence soit à droite ($p_R(s_r)$) et l'image test à gauche ($p_L(s_t)$) est donc finalement :

$$p(L | s_r, s_t) = p_R(s_r)p_L(s_t) \quad (3.9)$$

avec

$$p_R(s_r) = \prod_t \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(s_r(t) - \varphi(b(t)))^2}{2\sigma^2}\right) \quad (3.10)$$

et

$$p_L(s_t) = \prod_t \frac{1}{\sqrt{2\pi(\sigma^2 + \varphi'(b(t))^2\tau^2)}} \exp\left(-\frac{(s_t(t) - \varphi(b(t)))^2}{2(\sigma^2 + \varphi'(b(t))^2\tau^2)}\right) \quad (3.11)$$

La probabilité que l'observateur détecte l'image test est ensuite donnée par $p(\log(\mu) > 0)$ où

$$\begin{aligned} \log(\mu) &= \log(p_R(s_r)p_L(s_t)) - \log(p_R(s_t)p_L(s_r)) \\ &= \frac{1}{2\sigma^2} \left[\sum_t a(t) (s_t(t) - \varphi(b(t)))^2 - \sum_t a(t) (s_r(t) - \varphi(b(t)))^2 \right] \end{aligned} \quad (3.12)$$

$$\text{avec } a(t) = \frac{\varphi'(b(t))^2}{\varphi'(b(t))^2 + \left(\frac{\sigma}{\tau}\right)^2}.$$

Nous faisons l'hypothèse que l'observateur connaît sa représentation sensorielle de chacun des pixels du signal de base, $\varphi(b(t))$. Les quantités $s_t(t) - \varphi(b(t))$ et $s_r(t) - \varphi(b(t))$ associées à chaque pixel peuvent alors être respectivement estimées par $f(t) + \varepsilon_1(t)$ et $\varepsilon_0(t)$.

Nous supposons également que $a(t)(f(t) + \varepsilon_1(t))^2$ et $a(t)\varepsilon_0(t)^2$ sont deux variables indépendantes et identiquement distribuées. En considérant un grand nombre de pixels, T , nous

pouvons alors appliquer le théorème Central Limit, qui nous donne les distributions asymptotiques suivantes :

$$\sum_{t \in T} a(t) (f(t) + \varepsilon_1(t))^2 \sim \mathcal{N} \left(\sum_{t \in T} [a(t)(\sigma^2 + f(t)^2)], 2\sigma^2 \sum_{t \in T} [a(t)^2(\sigma^2 + 2f(t)^2)] \right) \quad (3.13)$$

et

$$\sum_{t \in T} a(t) \varepsilon_0(t)^2 \sim \mathcal{N} \left(\sigma^2 \sum_{t \in T} a(t), 2\sigma^4 \sum_{t \in T} a(t)^2 \right)$$

Finalement, nous obtenons donc approximativement

$$\log(\mu) \sim \mathcal{N} \left(\frac{\sum_{t \in T} (a(t)f(t)^2)}{2\sigma^2}, \frac{\sum_{t \in T} [a(t)^2(\sigma^2 + f(t)^2)]}{\sigma^2} \right) \quad (3.14)$$

et la probabilité que l'observateur réponde correctement est donnée par

$$\Phi \left(\frac{\sum_{t \in T} [a(t)f(t)^2]}{2\sigma \sqrt{\sum_{t \in T} [a(t)^2(\sigma^2 + f(t)^2)]}} \right) \quad (3.15)$$

où Φ est la fonction distribution normale cumulée (Kingdom et Prins, 2016).

Ce résultat est vrai dans le cas où $b(t)$ est un bruit blanc. L'équation 3.15 se simplifie grandement si l'on fait l'hypothèse que $f(t) \sim \mathcal{N}(0, \sigma^2)$, ce qui suppose implicitement que les signaux b et g sont corrélés. Comme nous le verrons plus loin, notre modèle n'est pas vraiment pénalisé par cette hypothèse, et la probabilité que l'observateur réponde correctement est alors donnée par l'expression plus simple

$$\Phi \left(\frac{F}{2\sigma \sqrt{T\sigma^2 + F}} \right) \quad (3.16)$$

où $F = \sum_{t \in T} f(t)^2$ correspond à la quantité d'information sensorielle (autrement dit, un contraste utile).

Notons que F dépend à la fois des caractéristiques des stimuli (signaux b et g) et de celles de l'observateur (fonction de transformation interne φ). Ainsi, si l'on considère l'une de nos images originales, pour des paramètres η et ζ donnés, si l'on connaît la fonction de transformation φ de l'observateur, on peut facilement calculer la probabilité que l'observateur réponde correctement.

Nous avons jusqu'ici fourni un modèle assez standard pour une tâche de détection. Dans la section suivante nous allons préciser quelques détails spécifiques à notre expérience. Nous devons, en particulier, définir la fonction de transformation interne, φ , qui correspond, dans notre modèle, à l'étape d'encodage des informations de luminance et de chrominance dans l'espace sensoriel de l'observateur. Afin d'évaluer l'efficacité des différentes définitions de chrominance présentées dans le chapitre précédent à justifier nos résultats expérimentaux, nous

allons comparer deux modèles pour l'encodage de l'information de chrominance de chaque pixel dans l'espace sensoriel de l'observateur (c'est-à-dire, deux fonctions φ différentes). Notre premier modèle est basé sur une définition linéaire des canaux d'opposition des cônes, alors que le second inclut une normalisation divisive de la chrominance par la luminance.

3.2.3.2 Les deux hypothèses mécanistiques

Nous allons, dans cette partie, présenter les deux hypothèses mécanistiques concernant l'encodage de l'information de chrominance que nous allons confronter, ainsi que les implications de chacune sur le modèle général décrit ci-dessus.

Mécanisme linéaire

Nous appelons, un peu abusivement, mécanisme *linéaire*, la transformation interne du stimulus dans l'espace de représentation sensorielle de l'observateur consistant en une simple compression logarithmique des coordonnées DKL du stimulus. Cette transformation correspond à un processus d'adaptation intervenant au niveau des canaux post-récepteurs et suivant simplement une loi de Weber-Fechner. Dans le cadre théorique du modèle général de détection des couleurs proposé précédemment, cette hypothèse équivaut à considérer la fonction de transformation interne suivante :

$$\varphi_{lin} \begin{pmatrix} Lum(t) \\ RG(t) \end{pmatrix} = \begin{pmatrix} \log(|Lum(t)| + E_1) \\ \log(|RG(t)| + E_2) \end{pmatrix} \quad (3.17)$$

avec E_1 et E_2 deux paramètres liés à la fraction de Weber. Ici, pour simplifier, nous fixons les valeurs de E_1 et E_2 à 1 (différentes valeurs ont été testées pour contrôle, sans que nos conclusions n'en soient modifiées).

Mécanisme divisif

Nous appelons mécanisme *divisif* la transformation interne du stimulus dans l'espace de représentation sensorielle basée sur la définition de chrominance proposée par Párraga, Brelstaff et al. (1998). Cette hypothèse implique que l'encodage de l'information de chrominance inclut une étape de normalisation par l'information de luminance. Nous supposons, par ailleurs, que la transformation interne de l'information de luminance est identique à celle considérée dans le cas d'un mécanisme linéaire. La fonction de transformation interne associée à l'hypothèse du mécanisme divisif est donc la suivante :

$$\varphi_{div} \begin{pmatrix} Lum(t) \\ RG(t) \end{pmatrix} = \begin{pmatrix} \log(|Lum(t)| + E_1) \\ \frac{RG(t)}{Lum(t) + \omega} \end{pmatrix} \quad (3.18)$$

Nous ajoutons ici le terme ω afin d'éviter une division par zero dans le cas où l'information de luminance serait nulle. Pour simplifier, nous prenons $\omega = 1$ (différentes valeurs ont été testées pour contrôle, sans que nos conclusions n'en soient modifiées).

Afin de tester nos deux hypothèses, nous avons comparé les prédictions de chacune d'entre elles avec nos données expérimentales.

3.2.3.3 Comparaison des hypothèses mécanistiques

Ajustement des modèles aux données expérimentales

Dans un premier temps, pour chacune des hypothèses mécanistiques, nous avons cherché à optimiser la fonction psychométrique de probabilité de détection de l'image test, $\Phi\left(\frac{F}{2\sigma\sqrt{T\sigma^2 + F}}\right)$, en fonction de la quantité d'information de chrominance sensorielle, F . Nous estimons pour cela la valeur du paramètre σ , correspondant à l'écart-type du bruit interne de chacun des participants, en utilisant la méthode du maximum de vraisemblance pour que les courbes obtenues se conforment le plus possible aux données expérimentales.

Pour cette première expérience, à chacun des essais les deux images présentées contiennent la même information de luminance (ηLum_i) et une petite proportion de l'information de chrominance initiale (ζRG_i) est ajoutée à l'image test. Avec les notations introduites dans la section Modèle nous avons donc :

$$b(t) = \begin{pmatrix} \eta Lum_i(t) \\ 0 \end{pmatrix} \text{ et } g(t) = \begin{pmatrix} 0 \\ \zeta RG_i(t) \end{pmatrix} \quad (3.19)$$

En conséquence, sous l'hypothèse d'un mécanisme linéaire, la quantité d'information de chrominance sensorielle résultant de la présentation des stimuli b et $b + g$ est $F = \sum_t \left(\frac{\zeta RG_i(t)}{E_2}\right)^2$. En considérant l'hypothèse d'un mécanisme divisif, la quantité F prédite par le modèle est différente : $F = \sum_t \left(\frac{\zeta RG_i(t)}{\eta Lum_i(t) + \omega}\right)^2$.

Évaluation de la qualité de l'ajustement des modèles aux données expérimentales

Afin de comparer la qualité de l'ajustement des deux modèles aux données expérimentales, nous calculons les AIC (*Akaike Information Criteria* ou critère d'information d'Akaike) (Akaike, 1974). L'AIC est une mesure de la qualité d'un modèle statistique qui pénalise les modèles en fonction du nombre de paramètres, afin de satisfaire au critère de parcimonie. Plus un modèle présente un AIC faible, meilleur il sera. La formule que nous utilisons pour calculer l'AIC est la suivante :

$$AIC = -2 \log L + 2k \quad (3.20)$$

où L est le maximum de la fonction de vraisemblance du modèle et k le nombre de paramètres. Notons que dans le cas de notre expérience, pour chacun des deux modèles associés à nos hypothèses mécanistiques, nous avons $k=1$ avec σ le seul paramètre à optimiser.

3.2.4 Résultats

3.2.4.1 Performances de détection

Nous mesurons dans un premier temps les performances de détection, en termes de proportion de réponses correctes, en fonction du ratio d'information de chrominance initiale ajouté (ζ), pour chacun des participants (S1,S2,...S5) et dans les trois conditions de luminance (η).

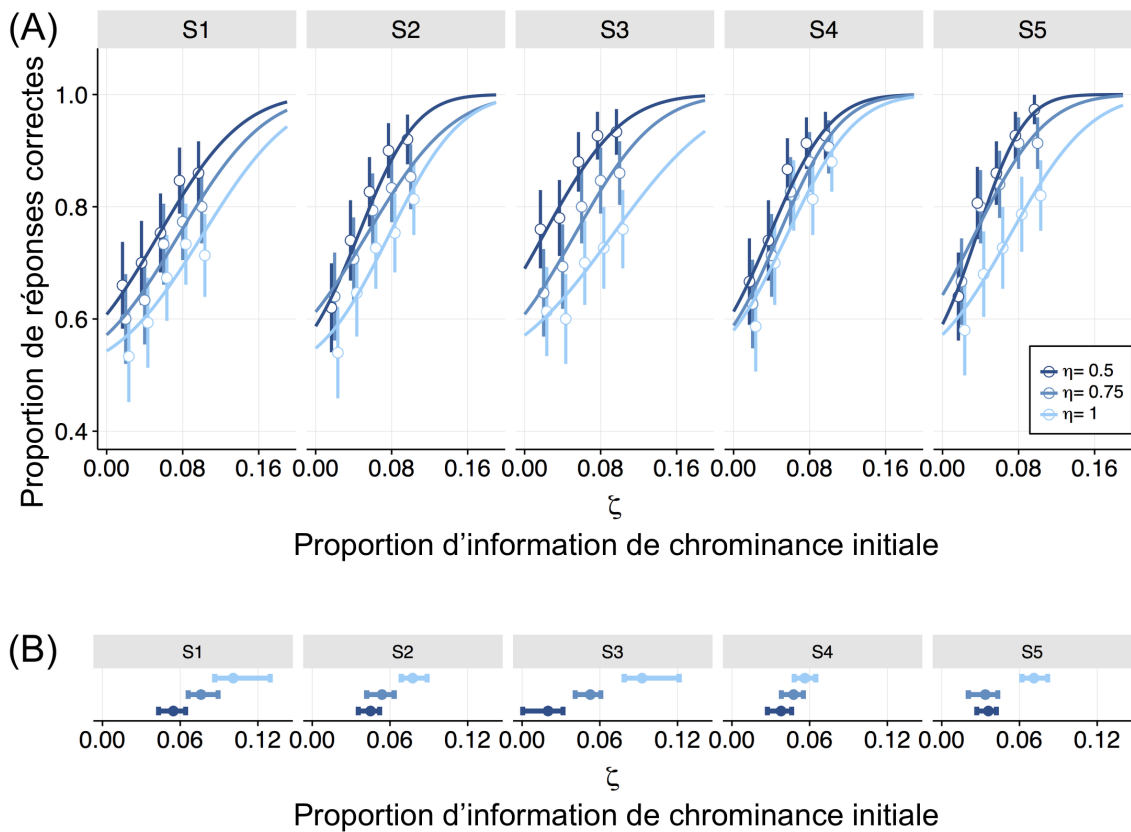


FIGURE 3.6 – Performances de détection des participants.

A : Proportion de réponses correctes (\circ) en fonction du pourcentage d'information de chrominance initiale, pour chacune des trois conditions de luminance (η), pour chaque participant (S1,S2,...,S5). Les barres d'erreur correspondent à l'écart-type. B : Estimation des seuils de détection individuels (à 75% de réponses correctes) pour chacune des conditions de luminance (η). Les intervalles ont été obtenus par bootstrap paramétrique ($n=1000$) et correspondent aux 95ème percentiles inférieur et supérieur de la distribution.

Nous constatons sur la Figure 3.6 que, pour nos 5 participants, les performances de détection décroissent avec le niveau moyen de luminance. Ce résultat suggère l'existence d'une interaction de type divisif de la chrominance perçue avec la luminance.

3.2.4.2 Ajustement des modèles aux données expérimentales

Pour chacune des deux hypothèses mécanistiques, nous optimisons pour chaque participant la fonction psychométrique de probabilité de détection de l'image colorée en fonction de F , la quantité d'information de chrominance sensorielle. Nous estimons pour cela le paramètre σ de la fonction $\phi\left(\frac{F}{2\sigma\sqrt{F+T\sigma^2}}\right)$ en utilisant la méthode du maximum de vraisemblance.

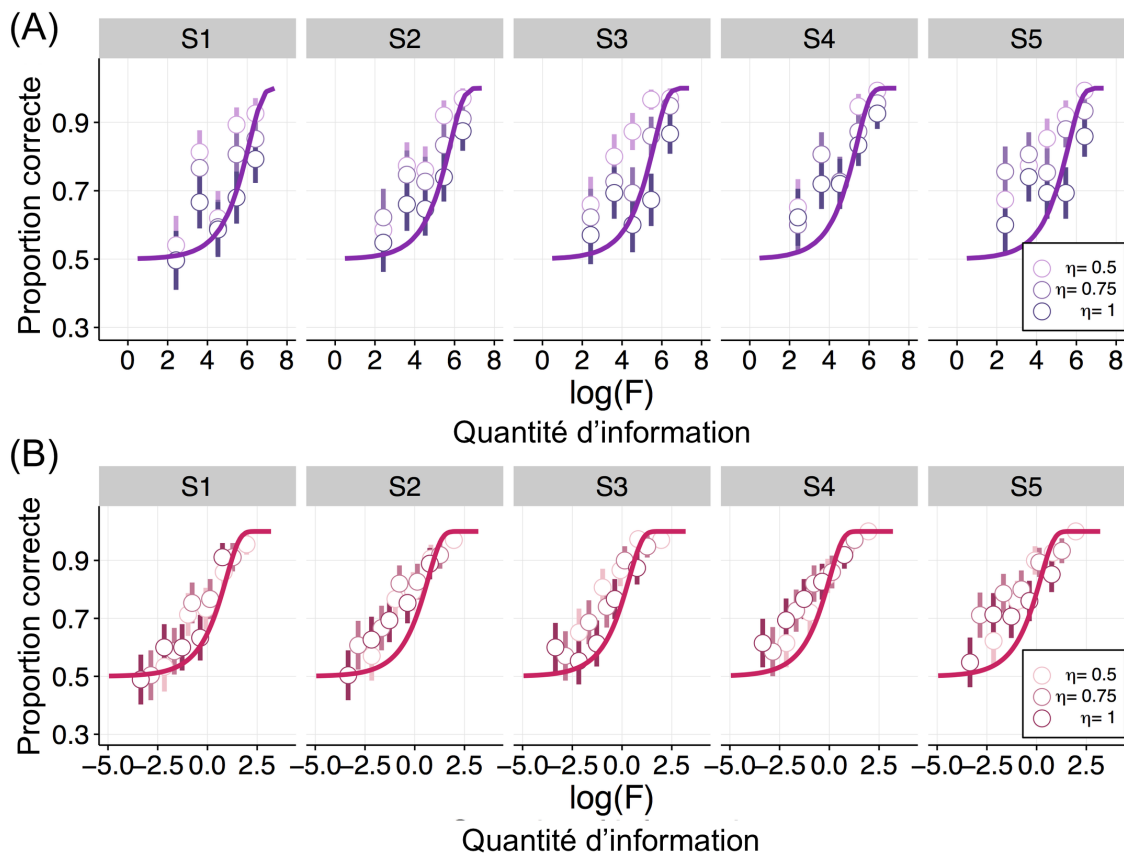


FIGURE 3.7 – Comparaison de l’ajustement des deux modèles associés aux hypothèses mécanistiques aux données expérimentales (A : Hypothèse mécanistique linéaire. B : Hypothèse mécanistique divisive.).

Proportion de réponses correctes (\circ) en fonction du logarithme de la quantité d’information de chrominance sensorielle ($\log(F)$) avec les courbes psychométriques $\phi\left(\frac{F}{2\sigma\sqrt{F+T\sigma^2}}\right)$ ajustées par maximum de vraisemblance pour chacun des observateurs (S1,S2,...,S5).

La Figure 3.7 montre que nos deux hypothèses mécanistiques nous permettent d’obtenir un ajustement convenable des courbes psychométriques aux données expérimentales, notamment compte tenu de la simplicité des modèles qui ne possèdent qu’un seul paramètre (σ).

3.2.4.3 Comparaison des AIC

Nous calculons les AIC associés à chacune des deux hypothèses mécanistiques, pour chaque participant individuellement. Nous constatons que pour tous nos participants l'AIC du modèle divisif (AICdiv) est inférieur à celui du modèle linéaire (AIClin) (Fig. 3.8). Ce résultat suggère donc un avantage du modèle divisif par rapport au modèle linéaire pour expliquer nos données expérimentales.

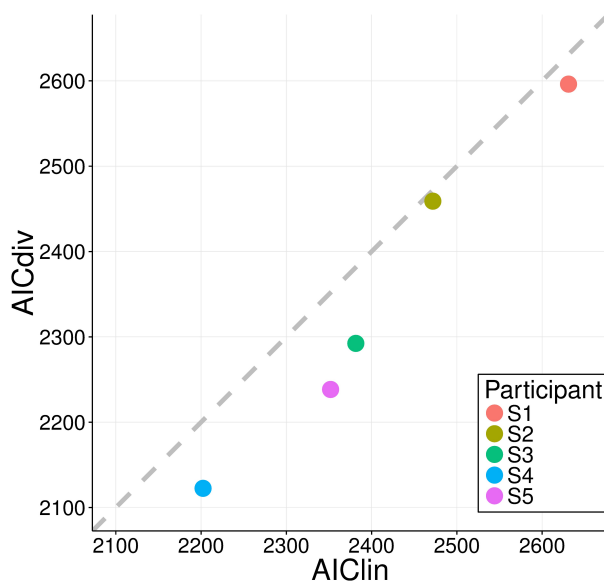


FIGURE 3.8 – Comparaison des AIC calculés pour chacun des participants sous l’hypothèse mécanistique linéaire (AIClin) et sous l’hypothèse mécanistique divisive (AICdiv). La ligne en pointillés correspond au cas où les valeurs seraient identiques avec les deux modèles.

3.2.4.4 Discussion

Nous avons, à ce stade, proposé un modèle d’encodage de l’information et de prise de décision que nous avons pu exploiter afin de tester l’efficacité de nos deux hypothèses mécanistiques à rendre compte de données expérimentales acquises dans une tâche de détection de couleur dans des images de scènes naturelles. Nos résultats mettent en évidence un avantage du mécanisme divisif par rapport au mécanisme linéaire. Nous devons cependant souligner, qu’une hypothèse implicite du modèle proposé jusqu’ici, est que les mécanismes d’encodage de l’information des participants sont parfaitement alignés avec les directions cardinales de notre espace DKL. Ce postulat nous a permis de fournir une description simple du cadre général de notre modèle, mais nous devons le mettre en doute et considérer l’impact qu’il a pu avoir sur nos conclusions.

3.3 Observateur décalé

Bien qu'il y ait consensus sur l'existence des mécanismes post-récepteurs, la façon dont les informations issues des cônes sont combinées et, notamment, le poids de chacun des types de cônes dans les différents canaux, sont très incertains. Deux hypothèses tacites sont faites avec l'utilisation de l'espace DKL :

- Le canal de luminance correspond à la fonction d'efficacité lumineuse $V(\lambda)$ (ou une de ses nombreuses variantes), qui somme les réponses des cônes de types L et M.
- Il existe une forme d'adaptation du système visuel au fond neutre qui impose le poids relatif de chacun des cônes dans les canaux post-récepteurs de chrominance : les coefficients α et β sont, en effet, déterminés par les propriétés de l'écran et du point neutre (Annexe A).

Plusieurs études ont pourtant cherché à mesurer le poids de chacun des types de cônes dans les mécanismes post-récepteurs de luminance et de chrominance, et ont obtenu des résultats très variés, mettant en évidence l'effet de nombreux paramètres (propriétés spatiales et temporelles des stimuli, méthode utilisée,...) et une grande variabilité inter-individuelle.

Une étude ayant cherché à mesurer le poids relatif des cônes L par rapport aux cônes M dans la fonction d'efficacité lumineuse de 40 observateurs placés dans des conditions identiques a, par exemple, obtenu des valeurs comprises entre 0.56 et 17.75 (Sharpe et al., 2005).

Pour le mécanisme de chrominance Rouge-Vert, comme mentionné dans le Chapitre 1, plusieurs définitions avec des pondérations différentes des cônes L et M ont été proposées. Des études utilisant des définitions basées sur les contrastes de cônes ont rapporté une contribution à peu près similaire des cônes L et M (Stromeyer, Cole et al., 1985 ; Chaparro et al., 1994 ; Cole, Hine et al., 1993 ; Sankeralli et Mullen, 1996) mais, dans le cas de l'utilisation de l'espace DKL, il n'est pas évident que le poids relatif de chaque type de cône dans le mécanisme d'un observateur donné soit entièrement déterminé par l'adaptation au fond neutre. Webster, Miyahara et al., 2000 souligne notamment que, comme pour la définition de la luminance, les axes de chrominance sont définis dans le cadre théorique d'un observateur standard et donc, sujets eux aussi à une grande variabilité inter-individuelle qui doit être prise en compte. Ils montrent notamment que les variations mesurées sont suffisamment importantes pour influencer l'interprétation de résultats concernant de possibles interactions entre canaux post-récepteurs.

Il est donc possible, qu'à la fois le canal de luminance et le canal de chrominance Rouge-Vert d'un observateur, ne soient pas parfaitement alignés avec les directions cardinales de notre espace DKL. Une conséquence de ce décalage est, qu'en utilisant l'espace DKL pour générer nos stimuli, lorsque nous souhaitons modifier exclusivement l'information de luminance ou de chrominance d'un pixel, nous n'avons pas accès à l'information réellement modifiée sur les canaux de l'observateur. Comme illustré sur la Figure 3.9, il en résulte que l'observateur a accès à une information sensorielle de luminance et de chrominance différentes de celles prédites dans l'espace DKL, et qu'il peut donc utiliser cette information pour réaliser la tâche

de détection.

Afin de tenir compte de cette variabilité, nous allons ajouter à notre modèle une première étape de conversion de l'espace DKL vers l'espace tri-dimensionnel d'opposition des cônes linéaire de l'observateur. Nous introduisons pour cela une matrice \mathbf{K} , qui correspond à la matrice de passage de l'espace DKL à l'espace défini par les directions des canaux post-récepteurs d'un observateur. Nous supposons que, dans les conditions d'observation considérées, pour un observateur donné, il existe un couple de paramètres (a, c) , tels que ses canaux de luminance et de chrominance puissent être modélisés par :

$$\begin{aligned} Lum_s &= L + \frac{1-c}{c}M \\ RG_s &= L - \frac{1-a}{a}M \end{aligned} \tag{3.21}$$

où $(a, c) \in [0, 1]$, le cas $(a, c) = (\frac{1}{\alpha+1}, 0.5)$ correspondant à un alignement parfait des canaux de l'observateur avec les directions cardinales de l'espace DKL.

Avec la matrice de passage $\mathbf{K} = \begin{pmatrix} c(\alpha-1)+1 & 2c-1 \\ a(\alpha+1)-1 & 1 \end{pmatrix}$ nous avons donc $\begin{pmatrix} Lum_s \\ RG_s \end{pmatrix} = \mathbf{K} \begin{pmatrix} Lum \\ RG \end{pmatrix}$, où $\begin{pmatrix} Lum \\ RG \end{pmatrix}$ sont les valeurs de luminance et de chrominance Rouge-Verte exprimées dans l'espace DKL et $\begin{pmatrix} Lum_s \\ RG_s \end{pmatrix}$ la projection de ces valeurs sur les canaux d'opposition des cônes correspondant de l'observateur.

Si nous reprenons les notations introduites dans la section Modèle, les signaux perçus par l'observateur ne sont plus b et $b+g$ mais b_s et b_s+g_s , avec $b_s = \mathbf{K}b$ et $g_s = \mathbf{K}g$. Avec les signaux b et g de la première expérience nous avons donc

$$b_s(t) = \begin{pmatrix} [c(\alpha-1)+1]\eta Lum_i(t) \\ [a(\alpha+1)-1]\eta Lum_i(t) \end{pmatrix} \text{ et } g_s(t) = \begin{pmatrix} (2c-1)\zeta RG_i(t) \\ \zeta RG_i(t) \end{pmatrix}$$

En conséquence, les deux stimuli ne diffèrent plus seulement en termes d'information de chrominance, mais également en termes d'information de luminance. La Figure 3.10 illustre les conséquences de l'intégration de cette nouvelle étape sur notre modèle d'encodage des informations de luminance et de chrominance dans l'espace sensoriel de l'observateur.

Dans le cas du modèle d'encodage proposé précédemment, sous l'hypothèse d'un mécanisme linéaire, l'information de luminance sensorielle dont un observateur décalé dispose est $F = A \sum_t \left(\frac{\zeta RG_i(t)}{\eta Lum_i(t) + \omega} \right)^2$ avec $A = \left(\frac{c(\alpha-1)+1}{2c-1} \right)^2$. La quantité d'information sensorielle, F , est donc identique à celle obtenue en considérant la chrominance d'un observateur non décalé sous l'hypothèse mécanistique divisive, au terme A près. Pour chaque valeur de A (donc chaque valeur de c), nous pouvons donc trouver une valeur de σ permettant d'optimiser l'ajustement de la fonction psychométrique avec les données expérimentales.

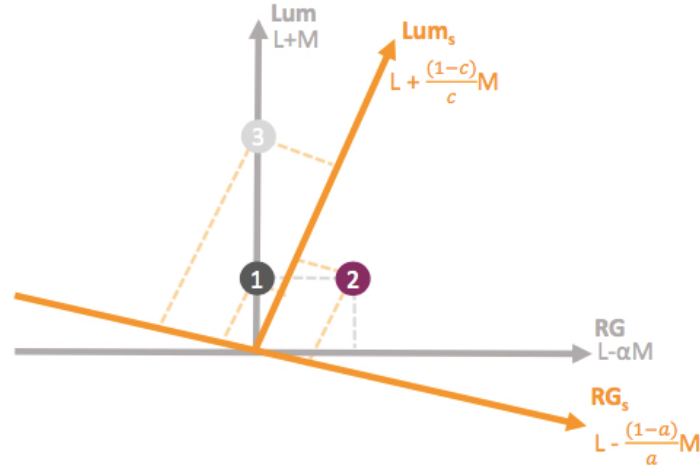


FIGURE 3.9 – Illustration des conséquences du décalage des canaux des mécanismes post-récepteurs d'un observateur avec les directions cardinales de l'espace DKL. Les pixels 1 et 2 ont la même information de luminance dans l'espace DKL (Lum), mais deux informations différentes si on les projette dans la direction du mécanisme de luminance de l'observateur (Lum_s). Les pixels 1 et 3 sont achromatiques dans l'espace DKL, mais contiennent tous deux une information de chrominance Rouge-Vert différente pour l'observateur décalé (RG_s).

Toujours sous l'hypothèse d'un mécanisme linéaire, et en considérant un observateur décalé, l'information de chrominance sensorielle dont il dispose est désormais $F = \sum_t \left(\frac{\zeta RG_i(t)}{|a(\alpha + 1) - 1| \eta Lum_i(t) + E} \right)^2$. Pour chaque valeur de a , nous pouvons à nouveau trouver une valeur du paramètre σ permettant d'ajuster la courbe psychométrique avec les données expérimentales.

Nous ne pouvons donc pas exclure la possibilité qu'il y ait en réalité un biais dans la construction des stimuli de notre expérience, qui ne permettrait en fait pas de décorrélérer complètement les variations de luminance et de chrominance compte tenu de la variabilité inter-individuelle dans la pondération des différents types de cônes dans les canaux post-récepteurs. Sous cette hypothèse, il est possible que l'information sensorielle utilisée par un participant pour réaliser la tâche soit différente de celle escomptée. Dans ce cas, à la fois l'information sensorielle de luminance et celle de chrominance prédites par un mécanisme linéaire, pourraient expliquer nos données. Afin d'écartier définitivement ces deux éventualités, nous avons conduit deux expériences supplémentaires.

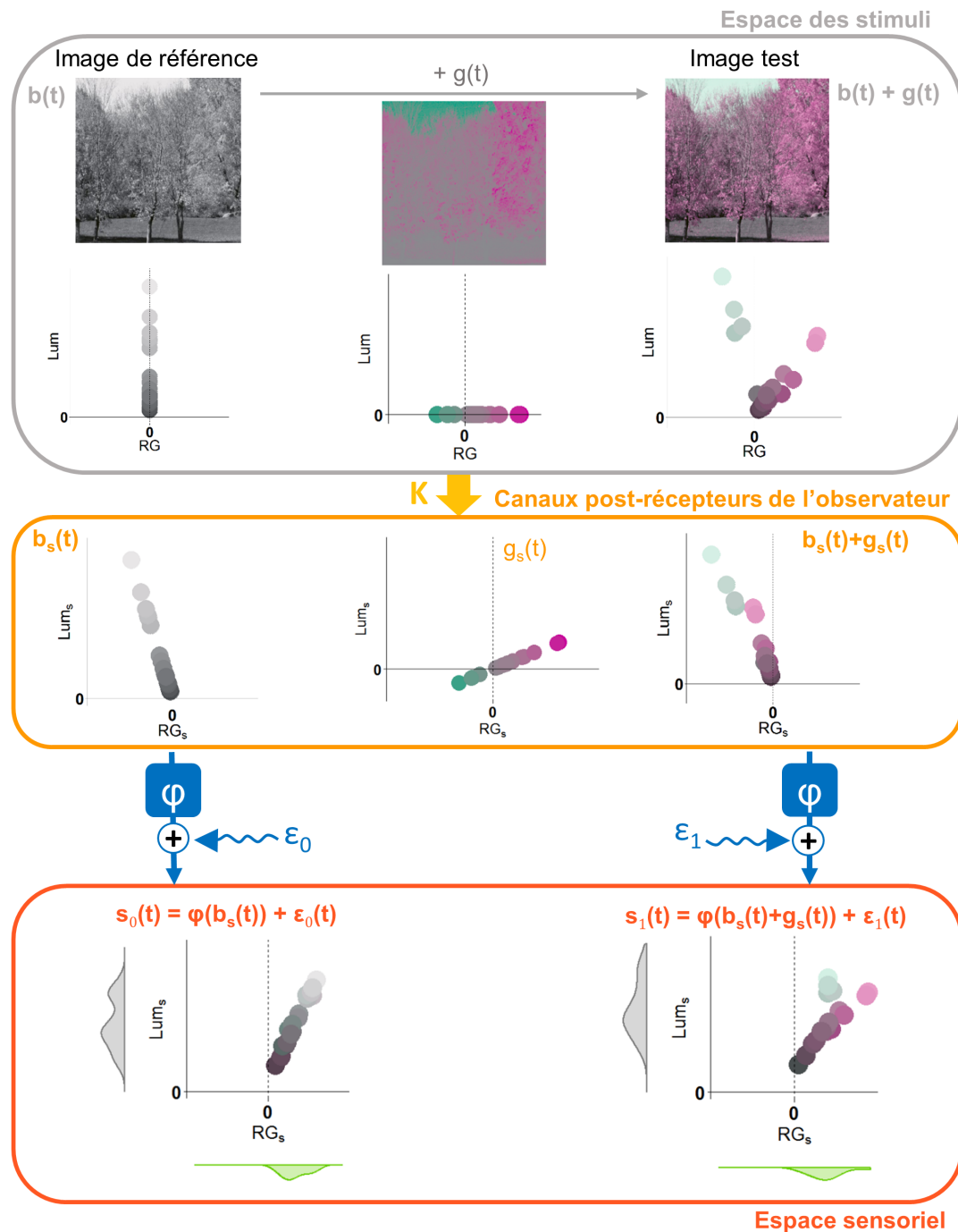


FIGURE 3.10 – Représentation schématique des étapes d’encodage des stimuli dans l’espace sensoriel d’un observateur décalé. Nous rajoutons, par rapport au modèle précédent, une étape d’encodage dans les directions des mécanismes post-récepteurs de l’observateur, présentant un décalage avec les directions cardinales de l’espace DKL utilisé pour la représentation des stimuli. Cette étape est modélisée par la matrice \mathbf{K} . Les autres étapes sont identiques mais, en conséquence du décalage, à la fois les distributions statistiques de chrominance et de luminance des deux stimuli peuvent désormais différer dans l’espace sensoriel de l’observateur.

3.4 Expérience 2 - Détection de luminance

3.4.1 Hypothèse

Pour exclure la possibilité qu'un participant ait pu réaliser la première expérience en utilisant uniquement l'information de luminance sensorielle sous l'hypothèse d'un mécanisme linéaire, nous réalisons une seconde expérience, dans laquelle nous avons explicitement demandé aux participants d'utiliser l'information de luminance, afin de mesurer leurs seuils de détection. Si c'est effectivement cette information que les participants ont utilisée pour réaliser la première expérience, alors les seuils de détection mesurés dans les deux expériences devraient être concordants.

3.4.2 Design expérimental

3.4.2.1 Protocole

Deux des participants ayant pris part à l'expérience précédente (une femme, un homme) participent à cette seconde expérience. Le dispositif utilisé est identique à celui décrit pour la première expérience. Pour cette deuxième expérience, la consigne donnée aux participants est d'identifier, parmi deux images en niveaux de gris de la même scène laquelle est la plus lumineuse. Comme précédemment, nous utilisons une procédure de type 2AFC : les stimuli sont présentés simultanément de part et d'autre de l'écran durant 500 ms.

3.4.2.2 Stimuli

Pour cette expérience, les stimuli sont générés à partir des 10 mêmes images originales que dans l'expérience 1. En notant, comme précédemment, b le signal associé à l'image de référence et $b + g$ le signal correspondant à l'image test, nous avons cette fois :

$$b(t) = \begin{pmatrix} \eta Lum_i(t) \\ 0 \end{pmatrix} \text{ et } g(t) = \begin{pmatrix} \zeta |RG_i(t)| \\ 0 \end{pmatrix} \quad (3.22)$$

où $\eta \in (0.5, 0.75)$ et $\zeta \in (1, 2, 3, 4, 5)$. Les valeurs des paramètres η et ζ ont été choisies pour que les performances des participants soient approximativement comprises entre 55 et 95% de réponses correctes. Un aperçu des différents stimuli obtenus à partir d'une des images originales est présenté sur la Figure 3.11.

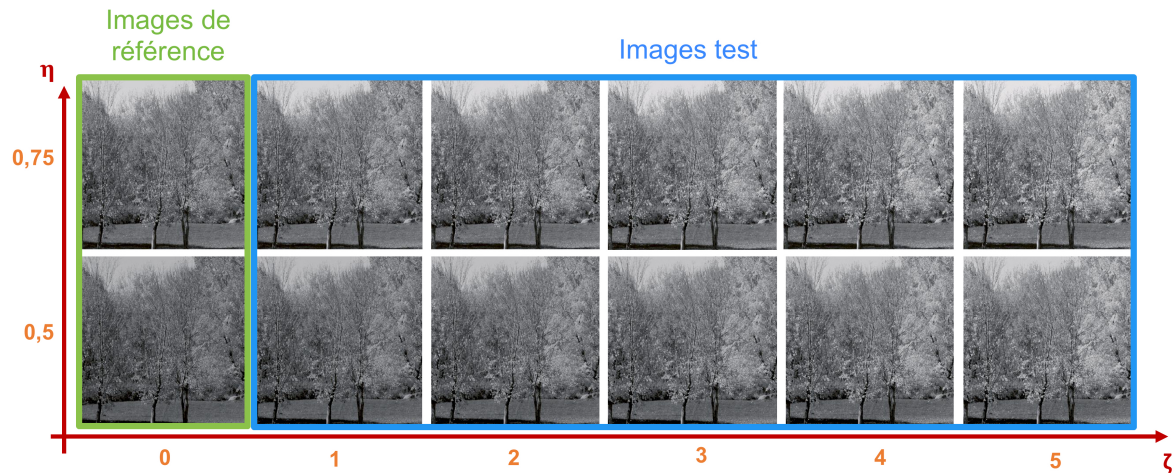


FIGURE 3.11 – Exemple de stimuli générés à partir d’une des 10 images originales pour l’expérience 2. Les images de référence ne contiennent, comme dans l’expérience 1, qu’une proportion η de l’information de luminance initiale, et les images test correspondantes contiennent désormais cette même information de luminance, à laquelle s’ajoute un ratio ζ de l’information de chrominance absolue initiale, projetée cette fois sur le canal de luminance de l’espace DKL.

3.4.3 Résultats

Pour les deux participants de la seconde expérience, nous présentons les données de la première expérience, avec les courbes psychométriques obtenues en ajustant le modèle aux données de la seconde expérience, pour huit valeurs du paramètre c , comprises entre 0 et 1. (Figure 3.12). Nous constatons ainsi, que les performances mesurées dans la deuxième expérience, sont incompatibles avec l’hypothèse d’un mécanisme linéaire basé sur l’information de luminance sensorielle dans la première expérience. En d’autres termes, si les participants avaient effectivement réalisé la première tâche en utilisant uniquement cette information, leurs performances devraient être bien plus élevées pour la seconde expérience. Ce résultat nous permet d’exclure la possibilité que, même si un observateur percevait, en raison d’un décalage de ses canaux post-récepteurs avec les directions cardinales de l’espace DKL et par un mécanisme linéaire, une différence d’information de luminance sensorielle entre les stimuli de la première expérience, il n’ait pu utiliser que cette information pour réaliser la tâche.

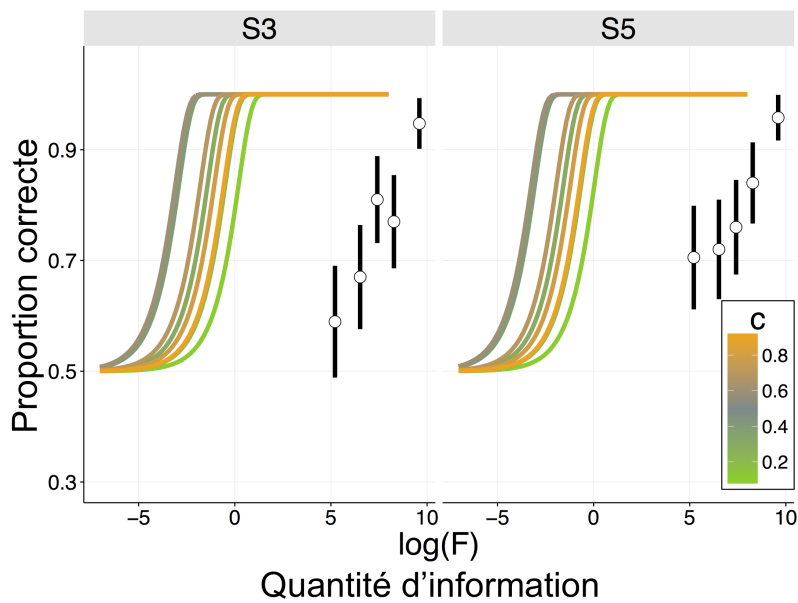


FIGURE 3.12 – Proportion d’images test correctement identifiées par les participants de l’expérience 2 en fonction du logarithme de la quantité d’information sensorielle, F , et sous l’hypothèse d’un mécanisme linéaire. En noir, les données de l’expérience 2. En couleur, les fonctions psychométriques obtenues par maximum de vraisemblance pour chacun des participants à partir de ses données enregistrées lors de l’expérience 1. Chaque couleur correspond à une valeur possible du paramètre c , qui exprime le décalage du mécanisme de luminance de l’observateur avec la direction cardinale de luminance de l’espace DKL. Pour toutes les valeurs de c considérées, les courbes psychométriques obtenues ne permettent pas d’expliquer les données de la seconde expérience.

3.5 Expérience 3 - Détection de couleur (contrôle)

Dans cette troisième expérience, nous cherchons à exclure la possibilité qu'un observateur ait pu effectuer la première tâche de détection avec un mécanisme linéaire en se fiant uniquement à l'information de chrominance sensorielle, qui serait différente de celle espérée compte tenu d'un décalage possible entre les canaux post-récepteurs de l'observateur et les axes de l'espace DKL.

3.5.1 Hypothèse

Supposons que ce soit ce mécanisme qu'un participant ait utilisé pour réaliser la première expérience. Nous pouvons alors trouver, quelle que soit la valeur de a , une valeur du paramètre σ permettant d'ajuster au mieux la courbe psychométrique aux données de la première expérience. Une autre expérience de détection de la couleur basée sur le même mécanisme devrait alors fournir des données concordantes avec les courbes psychométriques obtenues précédemment. Afin de tester cette hypothèse, nous allons réaliser une nouvelle expérience de détection de couleur dans des scènes naturelles, en distribuant cette fois l'information de chrominance initiale d'un pixel sur les axes de luminance et de chrominance de notre espace DKL. Nous introduisons pour cela un nouveau paramètre, u , correspondant au poids relatif de l'information de chrominance initiale projetée sur chacun des axes de l'espace DKL. Notre prédiction est que, si ce mécanisme a effectivement été utilisé lors de la première expérience, alors les seuils de détection mesurés dans cette troisième expérience ne devraient pas dépendre du paramètre u . C'est-à-dire, que le paramètre σ estimé pour un participant à partir des données de la première expérience, devrait permettre d'ajuster une unique fonction psychométrique à toutes les données de cette troisième expérience.

3.5.2 Design expérimental

3.5.2.1 Protocole

Les 5 personnes ayant déjà pris part à la première expérience ont participé à cette troisième expérience. Le dispositif expérimental reste identique aux expériences précédentes. Comme pour la première expérience, nous allons utiliser un paradigme de type 2AFC et demander aux participants laquelle, parmi deux images de la même scène, est colorée.

3.5.2.2 Stimuli

Les stimuli ont, cette fois, été générés en distribuant l'information de chrominance initiale de l'image sur les canaux de luminance et de chrominance de l'espace DKL. En notant à

nouveau b le signal de l'image de référence et $b+g$ celui de l'image test, nous avons :

$$b(t) = \begin{pmatrix} \eta Lum_i(t) \\ 0 \end{pmatrix} \text{ et } g(t) = \begin{pmatrix} u\zeta |RG_i(t)| \\ (1-u)\zeta |RG_i(t)| \end{pmatrix} \quad (3.23)$$

avec $\eta = 0.75$, $u \in (0, 0.5, 0.9)$ et $\zeta \in (0.02, 0.04, 0.06, 0.08, 0.1)$. Les valeurs des paramètres η , u et ζ ont été choisies pour que les performances des participants soient approximativement comprises entre 55 et 95% de réponses correctes. La Figure 3.13 montre l'exemple des stimuli obtenus à partir d'une image originale pour les différentes valeurs des paramètres u et ζ .

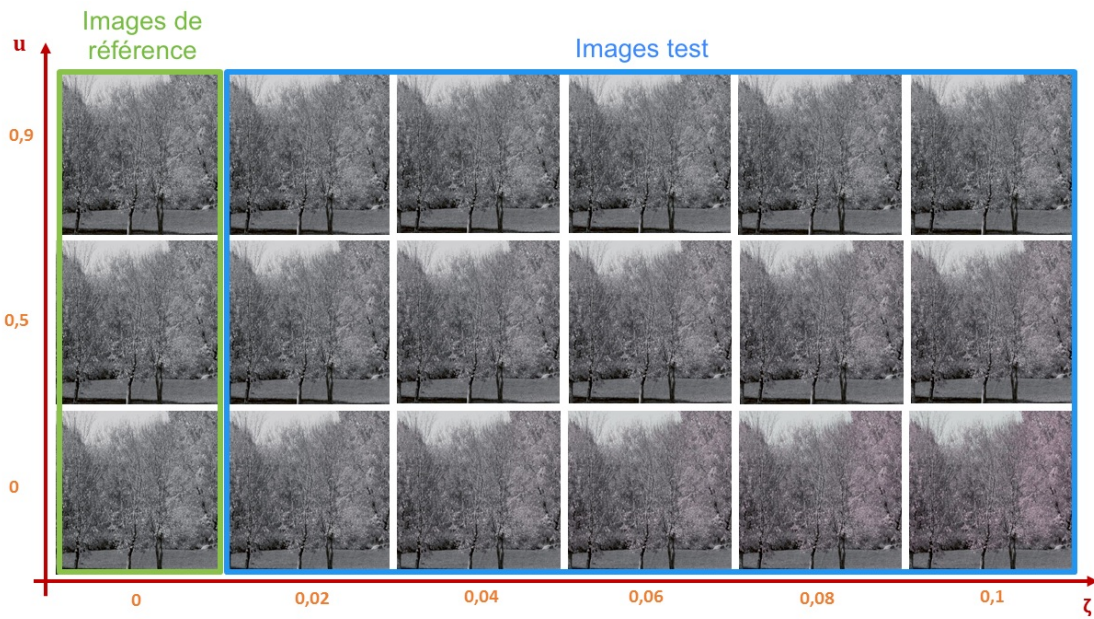


FIGURE 3.13 – Exemple de stimuli générés à partir d'une des 10 images originales pour l'expérience 3. Les images de référence ne contiennent, comme dans les expériences précédentes, qu'une proportion η de l'information de luminance initiale, et les images test correspondantes contiennent désormais cette même information de luminance, à laquelle s'ajoute un ratio ζ de l'information de chrominance absolue initiale, projetée cette fois en partie sur le canal de luminance de l'espace DKL et en partie sur le canal de chrominance, dans les proportions respectives u et $1 - u$.

3.5.3 Résultats

La Figure 3.14 montre, pour chacun des participants, la proportion de réponses correctes en fonction de la quantité d'information de chrominance sensorielle prédite par le mécanisme linéaire en tenant compte du décalage. Chaque ligne correspond à une valeur différente du paramètre a , qui exprime le décalage du canal de chrominance Rouge-Vert de l'observateur avec l'axe RG de l'espace DKL. Pour chacune des trois valeurs prises par le paramètre u , nous avons ajusté, par maximum de vraisemblance, les courbes psychométriques prédites par le

108

modèle aux données expérimentales. Nous obtenons invariablement 3 courbes, ce qui signifie qu'une seule valeur du paramètre σ est insuffisante pour expliquer toutes les données. Ce résultat suggère qu'il est donc peu probable qu'un observateur ait utilisé ce mécanisme dans la première expérience.

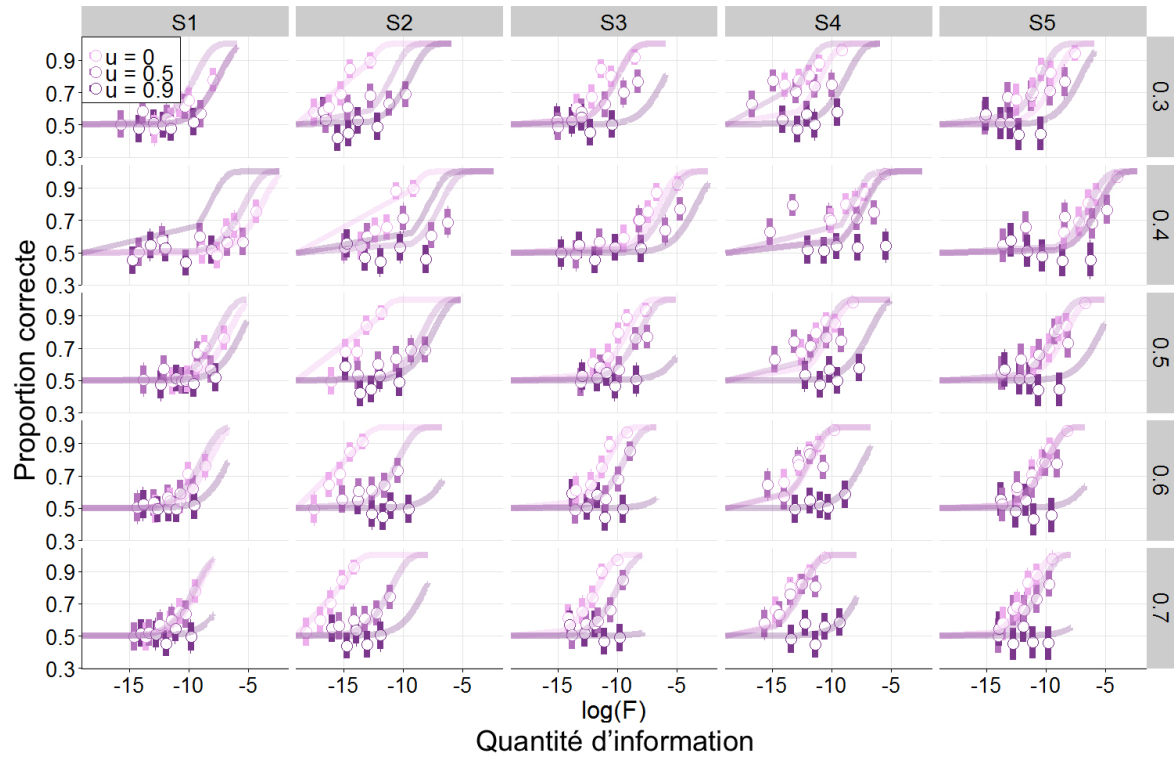


FIGURE 3.14 – Proportion d’images test correctement identifiées par les participants de l’expérience 3 en fonction du logarithme de la quantité d’information sensorielle, F , et sous l’hypothèse d’un mécanisme linéaire. Chaque colonne correspond à un participant, chaque ligne à une valeur du paramètre a , exprimant le décalage de l’axe du mécanisme de chrominance Rouge-Vert de l’observateur avec la direction cardinale RG de l’espace DKL, et chaque couleur correspond à une valeur du paramètre u . L’ajustement de fonctions psychométriques aux données obtenues pour chaque valeur du paramètre u par maximum de vraisemblance donne systématiquement 3 courbes différentes.

3.6 Discussion

3.6.1 Synthèse des résultats

Pour résumer, dans cette étude :

- Nous avons proposé un nouveau paradigme expérimental, qui nous a permis de montrer que les performances de détection de couleur dans des scènes naturelles dépendaient du

niveau de luminance moyen des images.

- Nous avons ensuite développé un modèle simple d'encodage de l'information chromatique et de prise de décision permettant d'expliquer nos données expérimentales.
- Nous avons par la suite utilisé le cadre général de ce modèle pour comparer deux hypothèses mécanistiques concernant l'encodage de l'information de chrominance dans l'espace sensoriel d'un observateur. Nous avons ainsi pu mettre en évidence un avantage du modèle divisif sur le modèle linéaire.
- Le mécanisme linéaire s'est cependant avéré plus compliqué à exclure que nous aurions pu l'espérer, en raison notamment des variations inter-individuelles dans la pondérations des différents types de cônes dans les canaux post-récepteurs. Deux expériences supplémentaires nous ont permis de confirmer l'avantage du modèle divisif, suggérant ainsi l'existence de mécanismes de normalisation au niveau des canaux post-récepteurs.

3.6.2 Cas du mécanisme logarithmique - Adaptation des photorécepteurs

Dans le chapitre précédent (Chapitre 2), nous mentionnions un autre modèle mécanistique des canaux post-récepteurs : le modèle logarithmique. Ce modèle présente un processus d'adaptation non-linéaire au niveau des photorécepteurs et non des canaux post-récepteurs. Afin d'évaluer la capacité de ce modèle à expliquer nos résultats empiriques, nous avons suivi la même procédure que pour nos deux autres hypothèses mécanistiques et ajusté les courbes psychométriques aux données expérimentales. Nous avons ensuite calculé les AIC pour chaque participant, et comparé ces valeurs à celles obtenues avec le modèle associé au mécanisme divisif (Figure 3.15). Nous constatons que les valeurs d'AIC sont très proches pour ces deux mécanismes, avec cependant un léger avantage pour le modèle logarithmique. Nos résultats ne nous permettent pas de statuer catégoriquement sur la supériorité de l'un ou l'autre de ces modèles. Nous constatons cependant, que les deux modèles intégrant une non-linéarité précoce (log ou div), sont plus performants qu'un modèle linéaire (lin) pour expliquer nos données de détection.

Il est intéressant de noter que, les deux modèles non-linéaires que nous avons considérés ici, ont été proposés à partir de considérations théoriques uniquement, basées sur des analyses statistiques sur les images de scènes naturelles. Nous avons donc ici apporté des données psychophysiques expérimentales soutenant l'existence d'une non-linéarité précoce dans traitement des informations de luminance et de chrominance des scènes naturelles.

3.6.3 Perspectives

Comprendre la nature exacte des mécanismes qui sous-tendent la perception de la couleur et de la luminance est le sujet de nombreuses études en sciences de la vision. Peu d'études se sont cependant intéressées à ces questions en travaillant sur des stimuli naturels. Beaucoup de travaux se sont focalisés sur l'analyse de stimuli simples, et ont permis de comprendre certaines

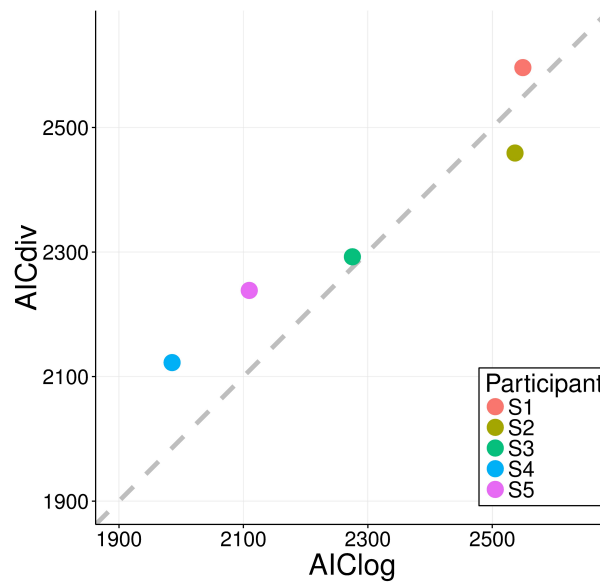


FIGURE 3.15 – Comparaison des AIC calculés pour chacun des participants sous l’hypothèse mécanistique logarithmique (AIClog) et sous l’hypothèse mécanistique divisive (AICdiv). La ligne en pointillés correspond au cas où les valeurs seraient identiques avec les deux modèles.

propriétés fondamentales de la vision des couleurs, mais il reste de très nombreuses questions irrésolues au sujet du traitement de l’information chromatique dans un contexte naturel. Nous proposons ici un nouveau paradigme expérimental pour l’étude de la perception de la couleur dans un contexte plus écologique. Utiliser des images de scènes naturelles et modifier leurs contenus de luminance et de chrominance dans un espace tri-dimensionnel, tel que l’espace DKL, présente un cadre pertinent, et fournit de nombreuses perspectives pour l’étude des mécanismes de perception des couleurs. Cependant, comme nous avons pu le constater dans cette étude, plusieurs paramètres doivent être scrupuleusement ajustés, et des expériences contrôles peuvent être nécessaires pour pouvoir formuler des conclusions sur les résultats obtenus. Il est notamment important de rester vigilants concernant les effets que peuvent avoir les différences entre les mécanismes de traitement post-récepteurs des observateurs humains et ceux d’un observateur standard, modélisés par l’espace DKL, pour l’interprétation des résultats.

Le modèle d’observateur bayésien idéal que nous avons ici proposé est très simple, et n’intègre pas de nombreuses propriétés connues du fonctionnement du système visuel humain. Premièrement, nous avons supposé que l’information de chaque pixel était encodée indépendamment, or la résolution spatiale du système visuel, initialement déterminée par celle de la mosaïque des photorécepteurs de la rétine, ne permet pas une telle précision. Nous avons également considéré une résolution spatiale identique pour l’ensemble du champ visuel, mais celle-ci décroît en réalité rapidement avec l’excentricité. Enfin, nous n’avons pas non plus tenu compte des différences connues de sensibilité spatiale entre les canaux de luminance et de chrominance. Notre modèle peut donc être considéré comme idyllique et approximatif ; mais il présente cependant l’intérêt d’expliquer efficacement nos données expérimentales compte tenu

de sa grande simplicité. Affiner ce modèle pour y intégrer, par exemple, des considérations spatiales, constitue une perspective intéressante.

Nous n'avons pas non plus eu le temps de creuser la question de l'efficacité du modèle à prédire les performances de détection des observateurs en fonction des images. Le modèle est-il plus efficace dans le cas de certaines images que pour d'autres ? Si oui, quels paramètres pourraient l'expliquer ? Autant de questions intéressantes qui restent ouvertes.

Nous avons ici volontairement laissé de côté le cas du canal d'information de chrominance Bleu-Jaune, mais étendre ces travaux à l'étude de celui-ci est une piste qui mérite d'être considérée.

Notre étude semble mettre en évidence l'existence de processus d'adaptation, mais ne nous permet pas de trancher entre adaptation précoce au niveau des photorécepteurs ou un peu plus tardive, sous forme notamment d'interactions entre canaux post-récepteurs. Il serait intéressant de proposer un paradigme expérimental permettant de confronter ces deux hypothèses, afin de mieux comprendre les processus d'adaptation entrant en jeu dans la perception des couleurs et dans quelle mesure le traitement réalisé par chacun des mécanismes post-récepteurs peut être considéré de façon isolée. Comme souligné par Vorobyev et Osorio (1998), l'interprétation de résultats psychophysiques en termes de mécanismes post-récepteurs présente quelques limites, à cause notamment du bruit déjà présent au niveau du signal d'excitation des cônes, et il est peut-être nécessaire d'envisager d'autres approches pour étudier plus précisément ces processus d'adaptation.

3.7 Conclusion

Finalement, nous avons malgré tout pu montrer par cette étude, que fournir une description mécanistique de la chrominance adaptée à la perception des scènes naturelles n'était pas une tâche triviale et nécessitait des investigations soignées. Notre modèle d'encodage-décodage de l'information par un observateur bayésien idéal, bien que très simple, explique en grande partie nos données expérimentales dans le cas d'un mécanisme divisif. Dans le chapitre suivant, nous admettrons cette hypothèse mécanistique et nous intéresserons au rôle de l'information de chrominance ainsi définie dans le cas d'une tâche spécifique : la classification de contours.

4

Rôle de la couleur dans la classification de contours

All vision is colour vision, for it is only by observing differences of colour that we distinguish the forms of objects.

James Clerk Maxwell, 1871

Sommaire

4.1	Introduction	114
4.2	Hypothèses	115
4.3	Expérience de classification de contours	116
4.3.1	Méthode	116
4.3.2	Résultats	121
4.4	Classifieurs	123
4.4.1	Modèle de classifieur bayésien	123
4.4.2	Propriétés des images	123
4.4.3	Analyse discriminante linéaire	127
4.4.4	Résultats	128
4.5	Comparaison des performances des observateurs humains et des classifieurs	130
4.6	Discussion	131
4.6.1	Synthèse des résultats	131
4.6.2	Perspectives	131
4.7	Conclusion	132

L'étude présentée dans ce chapitre a été menée dans le cadre d'une collaboration avec Frederick Kingdom et Ben Jennings (McGill Vision Research, Montreal) et a donné lieu à un article soumis dans le journal *Plos Computational Biology* : "*Color improves edge classification*", Camille Breuil, Ben Jennings, Simon Barthémé, Nathalie Guyader et Frederick Kingdom, 2018.

Dans ce chapitre, nous allons chercher à évaluer le rôle de la couleur dans une étape spécifique, précoce et importante de la perception des scènes naturelles : la classification de contours. Les contours sont omniprésents dans notre environnement visuel et peuvent résulter de changements de propriétés matérielles des objets et/ou de variations d'illuminants. Pouvoir les classer rapidement comme appartenant à l'une et/ou l'autre de ces catégories est indispensable pour acquérir une représentation cohérente des scènes qui nous entourent. Afin de mesurer l'utilité de la couleur pour cette tâche, nous avons dans un premier temps conduit une expérience psychophysique de classification de contours extraits d'images de scènes naturelles pour comparer les performances de participants dans le cas de stimuli en couleur ou en niveaux de gris. Nous avons ensuite construit des classifieurs artificiels, utilisant des propriétés simples des images, pour tester si des propriétés de chrominance amélioreraient ou non leurs performances de classification.

4.1 Introduction

En dépit de la complexité de notre environnement visuel, nous confondons rarement les changements d'illuminants, résultant par exemple de la présence d'ombres, et les changements matériels, tels qu'une tâche sur une surface ou une variation de texture d'un objet. Cette capacité à distinguer ces types de contours est cruciale pour assurer un traitement visuel efficace des scènes naturelles, et permettre notamment de segmenter les objets. Nous nous intéressons ici au rôle que pourrait jouer l'information de chrominance dans ce processus spécifique qu'est la classification de ces différents types de contours.

Les contours rencontrés dans notre environnement visuel peuvent avoir différentes origines : les bords d'un objet, des changements de réflectance ou de texture, des ombres,... Ces trois premières consistent en des changements des propriétés matérielles, tandis que les ombres sont des variations des propriétés d'illumination. Gilchrist et ses collègues furent l'un des premiers groupes à souligner l'importance de la classification des contours en ces deux catégories, dans leur cas pour estimer les propriétés de réflectance des surfaces (Gilchrist, 1979 ; Gilchrist et al., 1983). Distinguer les variations d'illuminants des variations des propriétés matérielles peut se faire en exploitant une multitude d'indices (Kingdom, 2008). On peut, *a priori*, supposer que la couleur en est un : dans notre environnement visuel, les variations de chrominance correspondent plus souvent à des changements matériels, tandis que, les changements de luminance peuvent être associés à des ombres comme à des changements matériels (Kingdom, Beauce et al., 2004 ; Rubin et Richards, 1982). On peut donc s'attendre à ce que les indices

de chrominance soient exploités par le système visuel pour réaliser cette classification.

Cette conjecture selon laquelle la couleur est préférentiellement associée à des changements de propriétés matérielles a d'ailleurs été utilisée avec succès par des algorithmes computationnels visant à segmenter les objets dans des images naturelles (Finlayson et al., 2002 ; Tappen et al., 2003 ; Olmos et Kingdom, 2004b). Des expériences psychophysiques, utilisant des stimuli simples et contrôlés, ont également montré que l'information de chrominance pouvait faciliter l'identification des ombres et des formes, en exploitant la supposition qu'une variation de chrominance est plus susceptible d'être associée à un changement matériel qu'à un changement d'illuminant (Kingdom, Beauce et al., 2004 ; Kingdom, 2003). Pour le cas des scènes naturelles, des preuves d'un bénéfice de l'information de chrominance pour l'*identification* des contours ont été apportées (Hansen et Gegenfurtner, 2017), mais, il n'existe pas, à ce jour, de données psychophysiques montrant que la chrominance faciliterait la *classification* des contours.

4.2 Hypothèses

Nous faisons l'hypothèse que, si la couleur est un indice utile, alors les performances d'observateurs dans des tâches de classification de contours devraient être améliorées pour des images couleur relativement à des images en niveaux de gris (i.e. ne contenant pas d'information de chrominance). On peut également supposer que les différences de performance entre ces deux conditions tendront à se réduire lorsque l'observateur possèdera plus d'information contextuelle sur le contour, pouvant ainsi exploiter d'autres indices pour réaliser la classification.

Afin de tester nos hypothèses, nous avons conduit une expérience psychophysique visant à mesurer la capacité d'observateurs humains à classifier des contours dans les catégories *ombre* ou *autre changement* pour des images couleur et en niveaux de gris, de trois tailles différentes. Pour vérifier si des attributs de chrominance des images pouvaient être exploités pour effectuer la classification, nous avons également construit des "observateurs artificiels", sous forme de classifieurs bayésiens linéaires intégrant des propriétés simples des images. Nous avons ensuite évalué les performances de ces classifieurs sur les images utilisées lors de l'expérience psychophysique et comparé les résultats à ceux de nos observateurs humains, afin de mesurer si un modèle très simple exploitant certaines propriétés basiques des images permettait d'expliquer facilement nos données expérimentales.

4.3 Expérience de classification de contours

4.3.1 Méthode

4.3.1.1 Stimuli

Les stimuli que nous avons utilisés pour cette expérience proviennent de la base d'images couleur calibrées McGill (<http://tabby.vision.mcgill.ca>, Olmos et Kingdom, 2004a). Des contours de type *ombre* ou *autre changement* ont été extraits des images par inspection visuelle. Les *ombres* sont définies comme des contours présentant uniquement une variation d'illumination tandis que, les *autres changements* réunissent tous les autres contours liés à des variations des propriétés matérielles (et n'étant pas associés à un changement d'illumination). Des carrés de différentes dimensions (petits : 72×72 pixels, moyens : 144×144 pixels et grands : 288×288 pixels, soit approximativement 2.5×2.5 , 5×5 et 10×10 degrés d'angle visuel dans les conditions d'observation de l'expérience) centrés sur des contours ont été sélectionnés parmi les images de la base à l'aide d'un algorithme d'extraction conçu sur mesure. Des images représentatives de nos stimuli sont présentées sur les Figures 4.1 et 4.2 (respectivement, avec et sans leurs contextes d'origine). Nous appliquons à chacune des images de contour un masque circulaire, lissé par une gaussienne (de dimensions 12×12 pixels et d'écart-type 30 pixels).

4.3.1.2 Espace de couleur

Pour déterminer si la couleur améliore les performances des observateurs, nous allons comparer leurs capacités à classer les contours dans le cas d'images couleur et dans celui d'images en niveaux de gris, ne contenant que l'information de luminance. Pour simplifier, nous nous référerons à ces deux conditions comme images "couleur" (Col) et images "pure luminance" (Lum). Tous les stimuli de l'expérience ont été générés en utilisant la décomposition dans l'espace d'opposition des cônes DKL (Derrington et al., 1984), constitué d'un canal de luminance (Lum) et de deux canaux de chrominance (RG et BY). Les excitations respectives de chacun des types de cônes L, M et S résultant de la présentation de l'image sur notre écran ont été calculées à partir des fonctions fondamentales d'absorption de Smith et Pokorny (1975). La procédure de conversion dans l'espace DKL est identique à celle utilisée dans les chapitres précédents et décrite dans l'Annexe A. Les stimuli de la condition "pure luminance" ont ensuite été obtenus par une méthode similaire à celle présentée dans le Chapitre 3 : en projetant la représentation de chacun des pixels dans l'espace DKL sur l'axe de luminance, conservant ainsi l'information de luminance et retirant le contenu de chrominance ($RG = BY = 0$). La Figure 4.3 fournit une illustration schématique des étapes de traitement effectuées sur les images de contours pour obtenir les stimuli présentés dans l'expérience.



FIGURE 4.1 – Exemple de stimuli (encadrés en rouge) extraits d’images de scènes naturelles pour notre expérience dans leurs contextes d’origine. À gauche : Contours de la catégorie *autre changement*, correspondant à des variations de propriétés matérielles. À droite : Contours de la catégorie *ombre*. De haut en bas : Stimuli de tailles petites, moyennes et grandes.

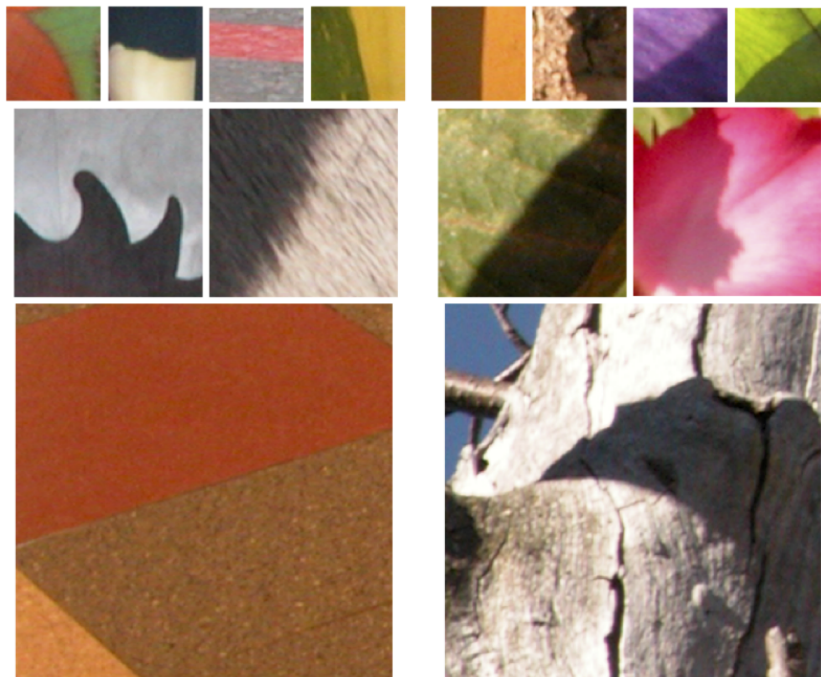


FIGURE 4.2 – Exemple de stimuli obtenus après extraction des contours des images. À gauche : Contours de la catégorie *autre changement*, correspondant à des variations de propriétés matérielles. À droite : Contours de la catégorie *ombre*. De haut en bas : Stimuli de tailles petites (8 exemples), moyennes (4 exemples) et grandes (2 exemples).

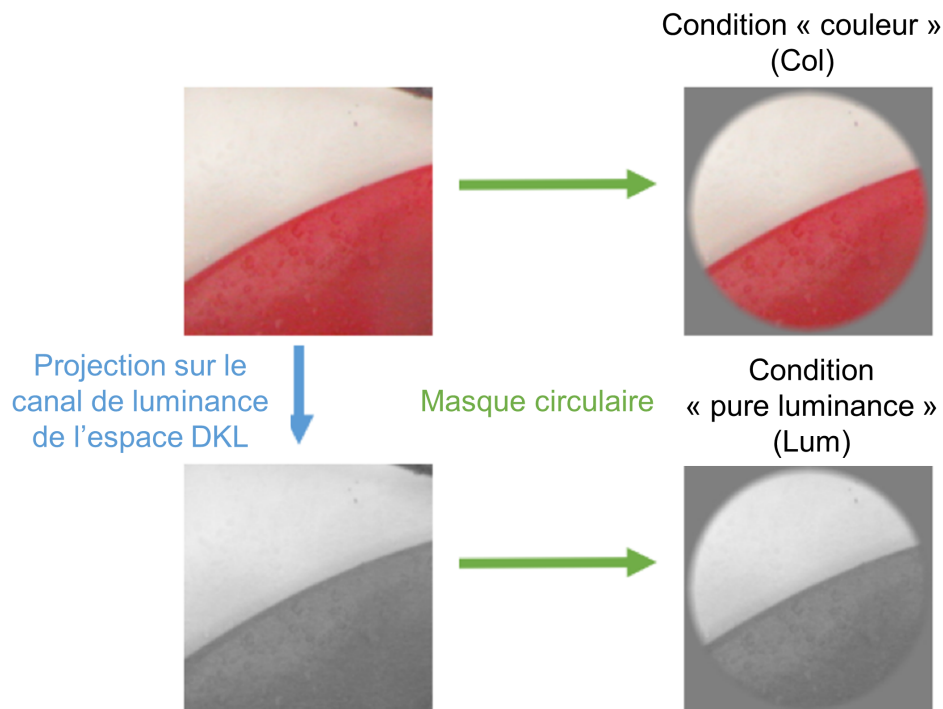


FIGURE 4.3 – Représentation schématique des étapes de traitement des images de contours pour obtenir les stimuli dans les conditions "couleur" (Col) et "pure luminance" (Lum).

4.3.1.3 Participants

10 participants (2 femmes, âges : 20 à 40 ans) ont pris part à l'expérience. Ils avaient tous une acuité visuelle normale ou corrigée à la normale, et leur vision des couleurs était normale. Chacun des participants a donné son consentement éclairé à prendre part à l'expérience.

4.3.1.4 Dispositif

Tous les stimuli ont été présentés aux participants sur un écran CRT de 21 pouces (Philips 201B4), de résolution spatiale 1024×768 pixels, avec une fréquence de rafraîchissement de 100 Hz. Les stimuli étaient affichés sur un fond gris neutre ($RGB = [127,127,127]$). La description détaillée de la procédure de calibration ainsi que les valeurs des paramètres de l'écran mesurés sont fournis en annexe (Annexe B). Durant toute l'expérience, les participants étaient assis dans une pièce sombre à une distance de 57 cm de l'écran. Leurs têtes étaient stabilisées au niveau du front et du menton par une mentonnière.

4.3.1.5 Protocole

À chacun des essais, un stimulus était brièvement affiché (500 ms) au centre de l'écran. Les participants devaient alors choisir si le contour présenté était une *ombre*, ou un *autre changement* en pressant une touche du clavier pour donner leur réponse.

L'expérience était divisée en 12 blocs de 50 essais. Au sein d'un bloc, tous les stimuli étaient de la même taille, mais aléatoirement présentés dans la condition couleur ou pure luminance. Pour chacune des 3 tailles de stimuli, et chacune des deux catégories de contours, 50 images couleur et 50 images pure luminance ont été présentées. Après une session d'entraînement de 8 essais avec feedback sonore, les participants ne recevaient plus aucun feedback durant la phase de test de l'expérience.

4.3.1.6 Analyse des données

Étant donnée la procédure expérimentale que nous employons, les participants sont susceptibles d'avoir un biais de réponse, c'est-à-dire une tendance à répondre plus systématiquement *ombre* que *autre changement* (ou l'inverse). Nous convertissons donc nos données en termes de d' , un indice de sensibilité défini dans le cadre de la théorie de la détection du signal permettant d'évaluer les capacités de discrimination des observateurs, indépendamment de la stratégie qu'ils adoptent pour prendre leur décision (caractérisée par le biais ou critère de réponse, c). Pour ce faire, les données sont, dans un premier temps, converties en proportions d'identifications correctes (pH) et de fausses alarmes (pFA). Une identification correcte correspond ici au cas où l'observateur répond *autre changement* quand le contour présenté est effectivement de la catégorie *autre changement* et une fausse alarme correspond à la situation où l'observateur répond *autre changement* alors que le contour présenté était une *ombre*. Le d' est ensuite calculé en faisant la différence des valeurs pH et pFA après les avoir transformés par la fonction $zscore$, z , qui exprime l'écart relatif des valeurs par rapport à la moyenne de la distribution, en termes d'écart-type :

$$d' = z(pH) - z(pFA) \quad (4.1)$$

avec $z(pX) = \frac{pX - \overline{pX}}{\sigma_{pX}}$ où \overline{pX} est la valeur moyenne de la variable pX et σ_{pX} son écart-type.

Pour chacun de nos participants nous avons également calculé le biais de réponse, c , défini comme l'opposé du $zscore$ de la proportion de fausses alarmes ($-z(pFA)$) (DeCarlo, 1998). Une valeur de c inférieure à 1 indique une tendance de l'observateur à répondre *ombre*, alors qu'une valeur supérieure 1 montre un biais en faveur de la réponse *autre changement*.

4.3.2 Résultats

Sur l'ensemble des essais de tous les participants le taux de réussite pour l'expérience est de 72% de bonnes réponses. Les participants ont donc globalement bien réussi la tâche.

Nous présentons les valeurs de d' obtenues en séparant les conditions images couleur (Col) et images pure luminance (Lum), en fonction de la taille des stimuli, pour chacun des participants individuellement (Figure 4.4) puis en moyennant sur l'ensemble des participants (Figure 4.5).

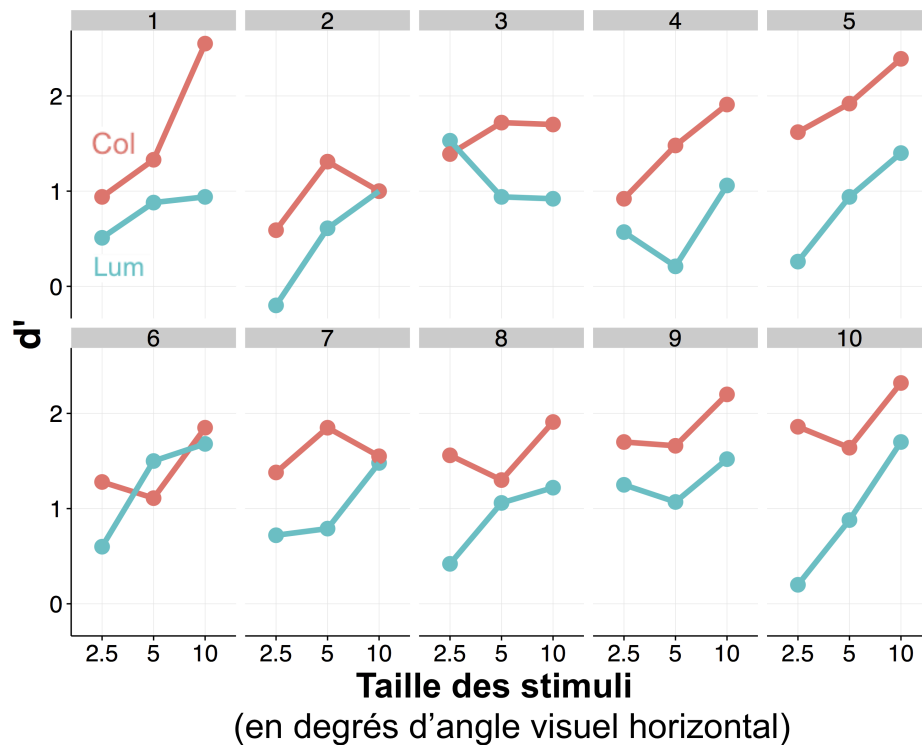


FIGURE 4.4 – Performances de classification des 10 participants. Les résultats sont présentés en termes de d' en fonction de la taille des stimuli et de la condition (Col ou Lum). Même s'il y a des variations entre les participants, les d' sont globalement plus élevés dans la condition Col que dans la condition Lum et augmentent avec la taille des stimuli.

La Figure 4.5 illustre l'effet, à la fois de la taille des stimuli et de la présence d'information de chrominance, sur les performances de l'ensemble des observateurs. Nous pouvons constater que les performances de classification sont améliorées par l'augmentation de la taille des stimuli, ainsi que par la présence d'information de chrominance. Ce résultat suggère que la couleur pourrait constituer un indice utile pour distinguer un contour correspondant à une ombre d'un autre type de changement. La différence entre les performances des observateurs dans la condition couleur et la condition pure luminance semble, en revanche, à peu près constante à travers les différentes tailles de nos stimuli. L'écart entre les deux conditions ne semble donc pas beaucoup dépendre de l'importance du contexte dans le stimulus présenté.

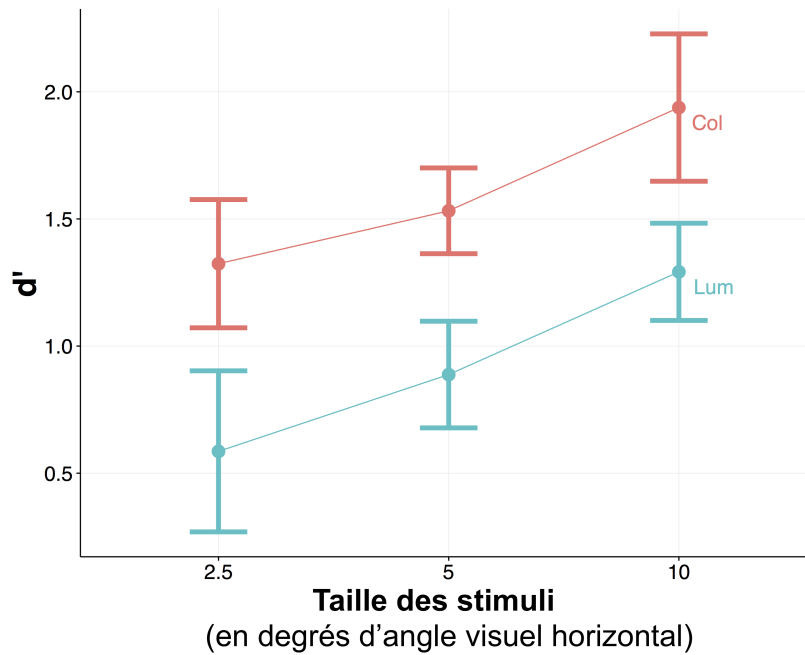


FIGURE 4.5 – Performances de classification moyennes des 10 participants. Les résultats sont présentés en termes de d' en fonction de la taille des stimuli et de la condition (Col ou Lum). Les barres d'erreur correspondent à l'erreur-type de la moyenne des d' sur l'ensemble des participants.

La Figure 4.6 répertorie les valeurs de biais, c , calculées pour chacun des participants, dans chacune des conditions (couleur ou pure luminance). Nous pouvons observer que la majorité de nos observateurs ont un biais vers la réponse *ombre*, à l'exception du participant 5, le biais pour l'une ou l'autre des réponses reste identique dans les deux conditions.

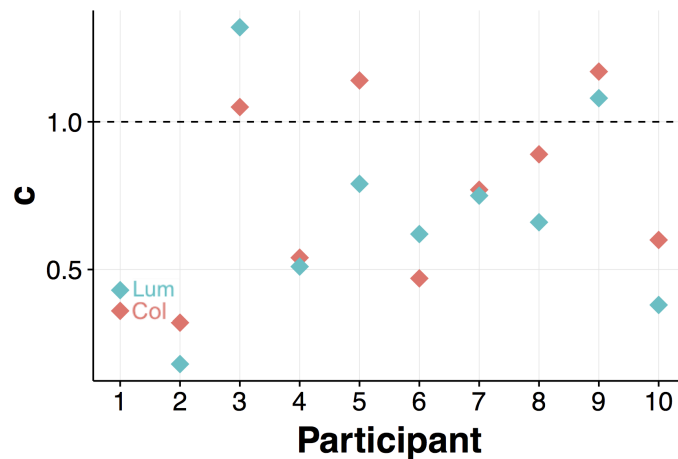


FIGURE 4.6 – Biais pour chaque participant, c , en fonction de la condition (Col ou Lum). On constate chez la majorité des participants un biais en faveur de la réponse *ombre* ($c > 1$), quelle que soit la condition.

4.4 Classifieurs

4.4.1 Modèle de classifieur bayésien

Nous considérons ici un modèle pour une tâche de classification binaire, dans laquelle un contour peut appartenir à l'une ou l'autre de deux catégories, représentées par la variable $e \in \{\text{ombre}, \text{autre changement}\}$. Pour prendre une décision concernant la valeur de e , le classifieur va tenir compte d'un certain nombre de propriétés du stimulus, que nous noterons s . Nous nous plaçons dans le cas d'un classifieur bayésien (Geisler, 2003; Green et Swets, 1988) qui réalise la tâche de façon optimale compte tenu de l'information disponible. Prenons le cas d'un stimulus appartenant à la catégorie *ombre*, la probabilité que la classification soit effectuée correctement est donnée par $P(\mu > 0)$, où $\mu = \frac{P(s | e = \text{ombre})}{P(s | e = \text{autre changement})}$ est la variable de décision. Nous supposons que la décision repose sur une combinaison linéaire des propriétés de l'image et nous allons utiliser une analyse discriminante linéaire pour tester si l'ajout de propriétés concernant l'information de chrominance des stimuli permet de prédire une amélioration des performances de classification.

4.4.2 Propriétés des images

4.4.2.1 Pré-traitement et compartimentation

Nous utilisons pour notre classifieur les mêmes stimuli que pour notre expérience psychophysique. Pour chacun, nous avons dans un premier temps manuellement repéré la position du contour, afin de positionner celui-ci au centre et l'orienter horizontalement dans chacune

de nos images. Toutes nos images sont ensuite compartimentées en trois régions : L_1 , L_2 et R , où R correspond aux dix lignes de pixels au centre de l'image, L_1 la région située au-dessus et L_2 la région en-dessous.

4.4.2.2 Propriétés de luminance

Pour chacun des pixels, nous calculons dans un premier temps une valeur de luminance, définie comme la somme des réponses des cônes de types L et M, soit la coordonnée Lum du pixel dans l'espace DKL. Nous normalisons ensuite les valeurs sur chacune des images individuellement, de sorte à couvrir une gamme comprise entre 0 et 1. Nous avons calculé, à partir de ces valeurs, trois mesures correspondant à des propriétés des images issues de la littérature de la vision par ordinateur et utilisées dans de précédentes études (DiMattina et al., 2012 ; Fine et al., 2003 ; Ing, 2010).

La première, C_m , est basée sur le contraste de Michelson et est calculée de la façon suivante :

$$C_m = \frac{|Lum_1 - Lum_2|}{Lum_1 + Lum_2} \quad (4.2)$$

où Lum_1 et Lum_2 sont les moyennes des valeurs de luminance des pixels des régions L_1 et L_2 . Nous avons, arbitrairement et sans que ceci n'ait d'impact sur nos résultats, orienté nos images de sorte à ce que $Lum_1 > Lum_2$, c'est-à-dire que L_1 soit systématiquement la région possédant la luminance moyenne la plus élevée.

La seconde mesure, L_{mean} , caractérise la luminance moyenne :

$$L_{mean} = \frac{Lum_1 + Lum_2}{2} \quad (4.3)$$

La troisième, $\Delta\sigma$, est une différence de contrastes :

$$\Delta\sigma = |\sigma(L_1) - \sigma(L_2)| \quad (4.4)$$

où $\sigma(L_1)$ et $\sigma(L_2)$ sont les écart-types des distributions des valeurs de luminance des pixels des régions L_1 et L_2 .

Enfin, pour quantifier la netteté du contour, nous effectuons une quatrième mesure correspondant à la pente de la transition des valeurs de luminance associée au contour. Nous suivons pour cela la méthode proposée par Vilankar et Golden (2014) afin de convertir chaque image en profil uni-dimensionnel de luminance et en extraire la valeur de la pente de la transition qui concorde avec la présence du contour. La pente, p , est calculée comme le changement moyen entre les valeurs de luminance moyennes des deux lignes de pixels situées aux extrémités de la région R :

$$p = \frac{\overline{Lum}(R(1), :) - \overline{Lum}(R(10), :)}{10} \quad (4.5)$$

où $R(1)$ est, en partant du haut de l'image, la première ligne de pixels de la région R et $R(10)$ la dernière. $\overline{Lum}(R(1), :)$ (resp. $\overline{Lum}(R(10), :)$) correspond à la moyenne des valeurs de luminance des pixels de la ligne $R(1)$ (resp. $R(10)$).

L'exemple d'une image compartimentée en régions L_1 , L_2 et R , ainsi que son profil uni-dimensionnel de luminance sont présentés sur la Figure 4.7. Nous montrons également, sur la Figure 4.8, les profils de luminance moyens obtenus pour l'ensemble des images de chacune des deux catégories et pour chaque taille indépendamment. Les valeurs des pentes moyennes (p) sont très légèrement différentes d'une catégorie à l'autre et décroissent avec la taille des stimuli.

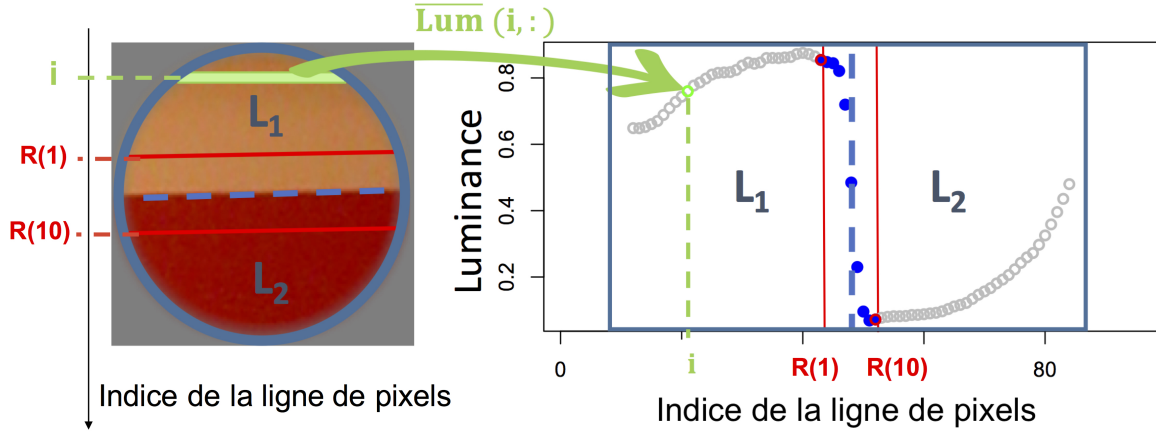


FIGURE 4.7 – A : Exemple d'une image segmentée en trois régions. Pour chacun des pixels de l'image localisé dans la ligne i et la colonne j nous calculons une valeur de luminance ($Lum(i, j)$). B : Profil uni-dimensionnel de luminance de l'image, obtenu en calculant la moyenne des valeurs de luminance de chaque ligne i de pixels de l'image ($\overline{Lum}(i, :)$).

4.4.2.3 Propriétés de chrominance

Nous utilisons les définitions proposées par Párraga, Brelstaff et al. (1998), incluant une normalisation divisive, pour calculer les valeurs de chrominance Rouge-Verte (RG) et Bleu-Jaune (BY) associées à chacun des pixels de nos images :

$$RG = \frac{L - \alpha M}{L + M} \text{ et } BY = \frac{2S - \beta(L + M)}{2S + \beta(L + M)} \quad (4.6)$$

avec (α, β) un couple de paramètre associé à l'écran utilisé pour l'affichage des images.

Comme nous l'avons mentionné précédemment (Chapitre 3), ces définitions présentent plusieurs avantages et semblent particulièrement pertinentes lorsque l'on s'intéresse à l'information contenue dans les pixels d'images de scènes naturelles (Olmos et Kingdom, 2004a). Comme proposé par DiMattina et al. (2012), nous employons deux mesures quantifiant les différences de contenu d'information de chrominance dans les régions situées de part et d'autre du contour. La première, ΔRG , concerne l'information de chrominance Rouge-Vert :

$$\Delta RG = |RG_1 - RG_2| \quad (4.7)$$

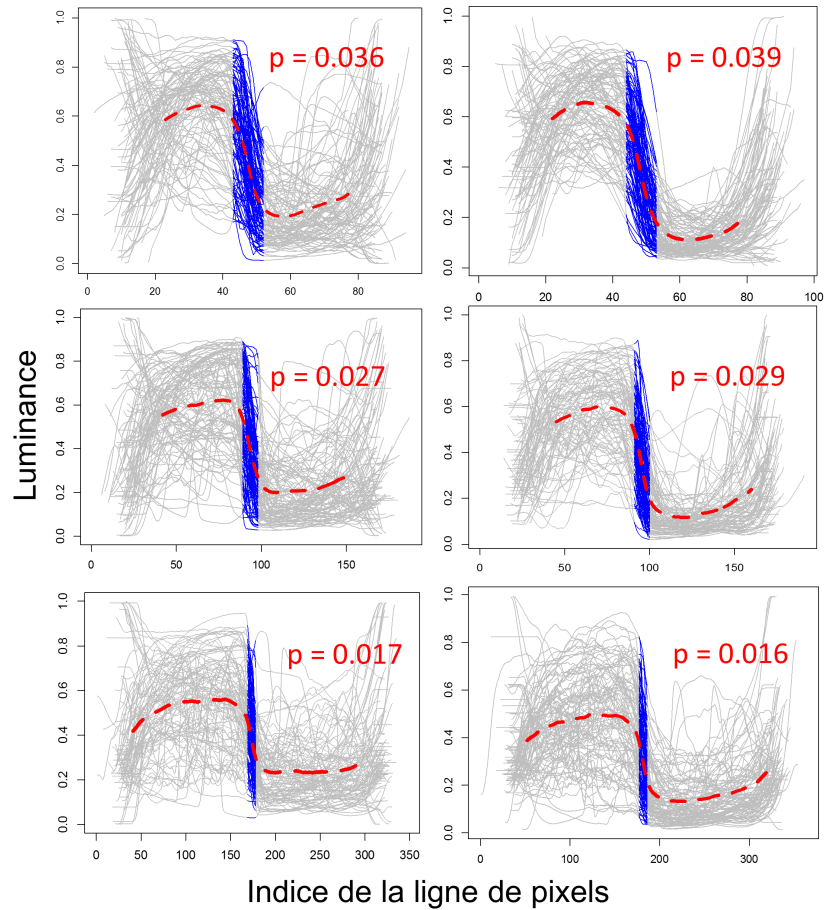


FIGURE 4.8 – Profils de luminance uni-dimensionnels de chacune des images (en gris) et profils moyens (en rouge) sur l'ensemble des images de chacune des catégories et de chacune des tailles. À gauche : images de la catégorie *autre changement*. À droite : images de la catégorie *ombre*. De haut en bas : images de tailles petites (72 lignes de pixels), moyennes (144 lignes de pixels) et grandes (288 lignes de pixels).

où RG_1 et RG_2 sont les valeurs moyennes de chrominance Rouge-Vert des pixels situés dans les régions L_1 et L_2 .

La seconde, ΔBY , est l'équivalent pour la chrominance Bleu-Jaune :

$$\Delta BY = | BY_1 - BY_2 | \quad (4.8)$$

Afin de quantifier l'utilité potentielle de ces propriétés de luminance et de chrominance pour la tâche de classification de contours, nous avons mesuré les performances de différents classifieurs. Chacun d'entre eux a été défini par l'utilisation d'une combinaison spécifique de certaines de ces propriétés, intégrant ou non les propriétés de chrominance. Nous souhaitons, en effet, mesurer en particulier si l'ajout de l'information de couleur, à travers les propriétés de chrominance, permet d'améliorer les performances des classifieurs.

4.4.3 Analyse discriminante linéaire

Nous utilisons une analyse discriminante linéaire (ADL) (Fisher, 1936) pour trouver les combinaisons linéaires de propriétés des images permettant de séparer au mieux nos catégories de contours (*ombre* ou *autre changement*). L'ADL permet, dans notre cas, d'expliquer et de prédire l'appartenance d'une image à l'une ou l'autre des catégories, à partir de ses caractéristiques mesurées à l'aide de variables prédictives (les propriétés de l'image).

Considérons un groupe de n images de contours et les observations des propriétés $s = \{s_1, \dots, s_k\}$ pour chacune d'entre elles. Le problème de la classification en deux catégories consiste à trouver un bon prédicteur de la classe, e , pour n'importe quelle image (n'appartenant pas nécessairement au groupe initial) en s'appuyant uniquement sur les observations s du groupe de n images. Pour utiliser cette approche, nous faisons l'hypothèse que les fonctions de densité de probabilité conditionnelles d'appartenance à l'une ou l'autre des classes, $P(s | e = \textit{ombre})$ et $P(s | e = \textit{autre changement})$, sont normalement distribuées, de moyennes et covariances respectives (μ_0, Σ_0) et (μ_1, Σ_1) . La solution optimale donnée par l'approche bayésienne est, de prédire que les contours appartiendront à la catégorie *ombre* lorsque le critère de décision, μ , sera supérieur à 1, ou de façon équivalente, lorsque $\log \mu > 0$ avec

$$\begin{aligned} \log(\mu) &= \log \left(\frac{P(s | e = \textit{ombre})}{P(s | e = \textit{autre changement})} \right) = \log \left(\frac{\det(\Sigma_1)^{\frac{1}{2}}}{\det(\Sigma_0)^{\frac{1}{2}}} \right) \\ &= \frac{1}{2} [(s - \mu_1^T) \Sigma_1^{-1} (s - \mu_1) - (s - \mu_0^T) \Sigma_0^{-1} (s - \mu_0)] \end{aligned} \quad (4.9)$$

Une hypothèse supplémentaire est requise pour l'utilisation de l'ADL : que les covariances de chacune des classes soient égales ($\Sigma_0 = \Sigma_1 = \Sigma$). En conséquence, $s \Sigma_0^{-1} s = s \Sigma_1^{-1} s$ et $s \Sigma_i^{-1} \mu_i = \mu_i \Sigma_i^{-1} s$ car Σ_i est symétrique. L'expression du critère de décision, $\log(\mu)$ donnée par l'équation 4.9 se simplifie alors :

$$\log(\mu) = s \Sigma^{-1} (\mu_1 - \mu_0) - \frac{1}{2} (\mu_0 \Sigma^{-1} \mu_0 - \mu_1 \Sigma^{-1} \mu_1) \quad (4.10)$$

En notant $w = \Sigma^{-1}(\mu_1 - \mu_0)$ et $c = \frac{1}{2}(\mu_0 \Sigma^{-1} \mu_0 - \mu_1 \Sigma^{-1} \mu_1)$, une expression simple du critère de décision devient donc $w \cdot s > c$, où $w \cdot s$ est le produit scalaire du vecteur w avec les observations s . Autrement dit, le critère pour déterminer à quelle catégorie, e , un contour appartient consiste en une simple combinaison linéaire des observations, s .

Nous allons donc utiliser cette méthode pour classifier nos images de contours et tester si les propriétés de chrominance améliorent les performances des classifieurs. Si les performances des participants à notre expérience psychophysique surpassent celles de nos classifieurs, nous pourrions en conclure que les observateurs humains ont, soit utilisé d'autres informations que celles incluses parmi nos propriétés, soit combiné ces propriétés par des mécanismes plus complexes que de simples transformations linéaires. Un tel résultat suggérerait donc que les performances des classifieurs pourraient être améliorées en tenant compte d'autres propriétés des images, ou en utilisant des classifieurs qui ne soient pas restreints à combiner linéairement ces propriétés.

4.4.4 Résultats

Nous avons testé un grand nombre de classifieurs, chacun d'entre eux étant défini à partir d'un sous-ensemble des propriétés de luminance présentées précédemment (C_m , L_{mean} , $\Delta\sigma$ et p), et en ajoutant, ou non, une, ou des, propriétés de chrominance (ΔRG et ΔBY). Nous avons mesuré les performances de nos classifieurs, à la fois pour l'intégralité de nos images et en séparant celles-ci par groupes de taille. Nous avons ensuite estimé les performances de nos classifieurs en termes de d' , à partir des matrices de confusion, afin de pouvoir les comparer aux performances des observateurs humains dans notre expérience psychophysique.

Comme illustré sur la Figure 4.9, pour tous les classifieurs que nous avons testés sur l'intégralité des images, les performances de classification étaient améliorées par l'ajout d'information de chrominance (ΔRG et/ou ΔBY). Ceci montre qu'il semble exister une explication physiquement concrète à l'amélioration des performances de nos observateurs humains dans la condition couleur de l'expérience psychophysique. Les résultats des classifieurs intégrant à la fois les propriétés ΔRG et ΔBY sont sensiblement identiques (voire un peu en deçà) à ceux des classifieurs auxquels seule la mesure de chrominance ΔRG a été ajoutée. De plus, ajouter seulement l'information de chrominance ΔBY n'améliore que les performances des classifieurs basés uniquement sur les propriétés C_m et $C_m + L_{mean}$. Ces résultats suggèrent donc que, contrairement à l'information Rouge-Vert, l'information Bleu-Jaune ne présenterait pas une grande utilité pour la classification des contours.

Une synthèse des résultats de tous les classifieurs que nous avons testés sur chacun des groupes de taille des stimuli est présentée sur la Figure 4.10.

Afin de tester si notre modèle de classifieur bayésien linéaire permet d'expliquer nos données expérimentales, nous allons comparer ces performances à celles des participants à notre expérience.

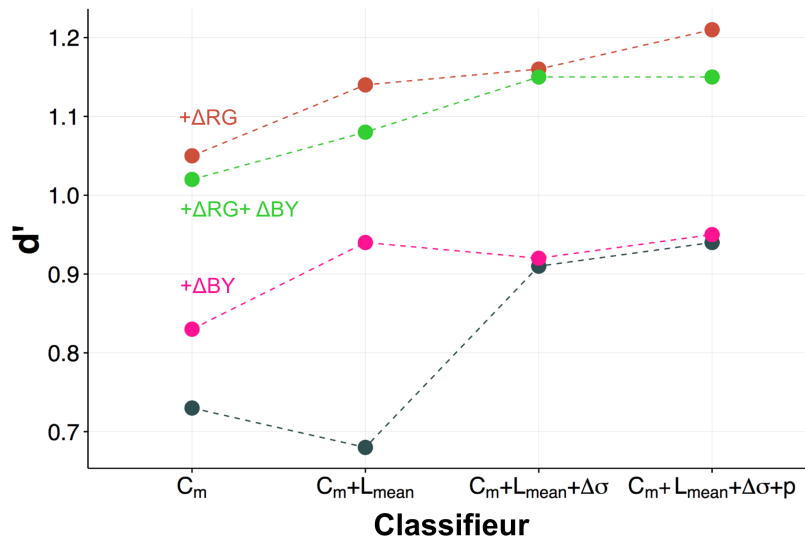


FIGURE 4.9 – Performances des classifieurs en fonction des propriétés utilisées. Les propriétés des images ont, dans un premier temps, été classées selon leur d' individuel, puis ajoutées par ordre décroissant. Le classifieur de référence (en bleu-gris) n'intègre que des propriétés de luminance, tandis que les classifieurs + ΔRG , + ΔBY et + $\Delta RG + \Delta BY$ intègrent, en plus, les propriétés ΔRG et/ou ΔBY des images.

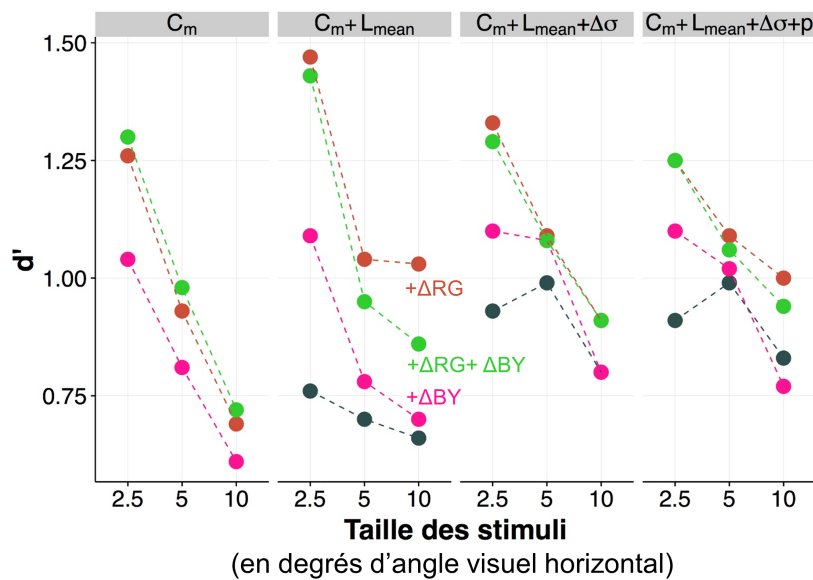


FIGURE 4.10 – Performances des différents classifieurs en fonction de la taille des stimuli.

4.5 Comparaison des performances des observateurs humains et des classifieurs

La Figure 4.11 présente à la fois les performances des participants à l'expérience psychophysique et celles des classifieurs. Dans le cas des données de la condition expérimentale pure luminance (Lum), les d' moyens des participants sont comparés à ceux du classifieur de référence intégrant toutes les propriétés de luminance des images ($C_m + L_{mean} + \Delta\sigma + p$) tandis que, les performances dans la condition expérimentale couleur (Col), sont comparées à celles du classifieur intégrant, en plus de toutes les propriétés de luminance, les mesures de chrominance ΔRG et ΔBY ($C_m + L_{mean} + \Delta\sigma + p + \Delta RG + \Delta BY$). Nous présentons également les performances du classifieur $C_m + L_{mean} + \Delta\sigma + p + \Delta RG$.

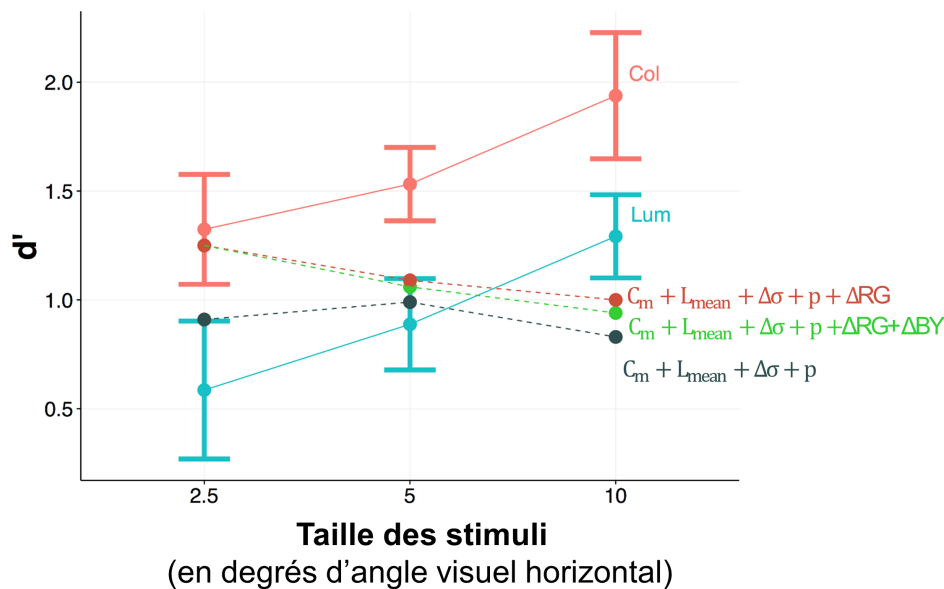


FIGURE 4.11 – Comparaison des performances des observateurs humains pour les conditions pure luminance (Lum) et couleur (Col) avec celles du classifieur intégrant toutes les propriétés de luminance $C_m + L_{mean} + \Delta\sigma + p$ avec ou sans les propriétés de chrominance ($+\Delta RG$ et $+\Delta RG + \Delta BY$).

Nous pouvons voir que les classifieurs capturent bien l'amélioration globale des performances avec l'ajout de l'information de chrominance observée expérimentalement. En revanche, ils échouent à reproduire l'effet de la taille des stimuli mesurés chez les observateurs humains, en particulier, pour les stimuli les plus grands pour lesquels les performances des classifieurs sont assez médiocres comparativement à celles des participants. Ceci suggère que les observateurs humains doivent utiliser des propriétés supplémentaires, comme par exemple des informations spatiales, que nous n'avons pas intégrées dans nos classifieurs. Les propriétés des images que nous avons utilisées pour nos classifieurs sont, en effet, très simples et ne tiennent pas compte de nombreuses connaissances sur le fonctionnement du système visuel, telles que la résolution spatiale des différents canaux. Il est cependant intéressant de noter qu'elles per-

mettent, malgré tout, de prédire l'amélioration des performances avec l'ajout d'information de chrominance observée empiriquement.

4.6 Discussion

4.6.1 Synthèse des résultats

Pour résumer, dans cette étude :

- Nous avons proposé un nouveau paradigme expérimental afin d'acquérir des données concernant l'utilité de l'information de chrominance pour la classification de contours extraits de scènes naturelles.
- Nos résultats révèlent une amélioration systématique des performances de classification de nos participants lorsque l'information de chrominance est disponible. Ceci suggère que la couleur pourrait effectivement être un indice utilisé pour réaliser cette tâche, vraisemblablement à une étape assez précoce du traitement visuel.
- Par ailleurs, les performances des observateurs augmentent également avec la taille des stimuli, ce qui semble insinuer l'utilisation d'indices contextuels. Contrairement à nos prédictions, nous n'observons cependant pas d'interaction entre les effets de taille et de couleur, l'écart entre les performances dans la condition pure luminance et couleur restant à peu près constant quelle que soient les dimensions des stimuli. Comme illustré sur la figure 4.5, la couleur ne semble pas être beaucoup plus utile pour classer les plus petits de nos stimuli que les plus grands.

La question de la classification des contours a été abondamment examinée dans le domaine de la vision par ordinateur, mais, à notre connaissance, les modèles qui en ont émergés n'ont jamais été directement comparés à des données psychophysiques. Nous avons ici proposé des modèles de classifieurs bayésiens linéaires très simples afin de comparer leurs performances à celles d'observateurs humains. Tous nos classifieurs permettaient d'expliquer raisonnablement bien l'accroissement des performances de nos participants dans la condition couleur. En revanche, comme illustré sur la figure 4.11, ils ont échoué à capturer l'effet de la taille des stimuli observé pour les participants de l'expérience.

Ce résultat suggère que les observateurs humains doivent être sensibles à des indices contextuels supplémentaires, qui n'étaient pas exploités par nos classifieurs.

4.6.2 Perspectives

Le modèle de classifieur bayésien que nous avons proposé ici est très grossier, et il est notamment très sensible au bruit qui augmente avec la taille des stimuli. Il serait intéressant d'affiner ce modèle, en tenant compte, par exemple, de caractéristiques spatiales dans la défi-

inition des propriétés des images pour parvenir, notamment, à répliquer l'effet de taille obtenu pour les observateurs humains.

Le fait que nous n'ayons, contrairement à nos prédictions, pas observé de véritable interaction des effets de taille et de couleur est peut-être dû à la taille des stimuli utilisés, qui contiendraient dans tous les cas trop d'information contextuelle pour que l'effet puisse être notable. Il serait donc intéressant de réaliser à nouveau l'expérience en utilisant des stimuli plus petits, pour voir si les conclusions seraient différentes.

Nous n'avons pas non plus testé le cas de fausses couleurs : il serait intéressant, en guise de contrôle, de modifier la chrominance des pixels de l'image, sans pour autant que les valeurs de nos propriétés de chrominance en soient affectées et voir si les performances des observateurs humains seraient modifiées.

4.7 Conclusion

Finalement, nos résultats mettent en évidence une réelle utilité de l'information de chrominance Rouge-Vert pour la classification des contours, à la fois dans le cas de nos observateurs humains et de nos classifieurs bayésiens linéaires, intégrant uniquement quelques propriétés simples des images. Il semble pertinent de s'inspirer des outils présentés ici pour étendre l'étude du rôle de la couleur à d'autres tâches visuelles.

5

Synthèse et perspectives

If the doors of perception were cleansed
everything would appear to man as it is,
infinite.

William Blake, *The Marriage of Heaven
and Hell*, 1793

Pour clore ce manuscrit, nous synthétisons les principaux résultats de ces travaux et présentons quelques perspectives pour la compréhension du rôle de la couleur dans la perception visuelle des scènes naturelles.

5.1 Synthèse des principaux résultats

Ce travail de thèse avait pour objectif de mieux comprendre l'information de chrominance et comment celle-ci était exploitée par le système visuel dans le cadre de la perception des images de scènes naturelles. Comme nous l'avons vu dans la partie introductive de ce manuscrit, la perception des couleurs a été extensivement étudiée à l'aide de stimuli simples et artificiels (comme, par exemple, des *patches* colorés) mais, finalement, assez peu dans le cas de stimuli plus complexes, tels que les scènes naturelles. Dans ces travaux, nous nous sommes donc attachés à aborder la question du traitement de l'information de chrominance pour la perception des scènes naturelles, afin de nous rapprocher de situations écologiques. Les trois principaux objectifs qui ont guidé notre démarche ont été de (1) quantifier l'apport de l'information de chrominance par rapport à l'information de luminance dans le contexte général de

la perception des scènes naturelles, (2) formuler une définition mécanistique de l'information de chrominance adaptée à ce cadre et (3) évaluer le rôle de cette information de chrominance dans des tâches visuelles. Afin de répondre à ces trois objectifs, nous avons réalisé une étude théorique sur les statistiques des scènes naturelles, mené une série d'expériences psychophysiques utilisant des images de scènes naturelles et proposé plusieurs modèles computationnels en vue de mieux comprendre comment l'information de chrominance était encodée par le système visuel et utilisée pour différentes tâches.

5.1.1 Analyse statistique des scènes naturelles

Le **Chapitre 2** présente une étude visant à quantifier l'indépendance statistique des informations de luminance et de chrominance dans les scènes naturelles. Cette analyse, en partie inspirée de celle de Hansen et Gegenfurtner (2009), nous a permis de proposer un indice pertinent pour classer les images en fonction de la redondance de leurs contenus de luminance et de chrominance, et a également mis en évidence d'importantes variations inter-images. À notre connaissance, notre analyse est la première à s'intéresser à la mesure de cette information à l'échelle d'une image (à la différence de l'étude de Hansen et Gegenfurtner (2009), qui s'intéressait à un ensemble d'images), et nous permet, après avoir fixé un certain nombre de paramètres et quantifié leurs effets, d'établir un classement des images en fonction de l'indépendance de leurs contenus de luminance et de chrominance. Nous avons ainsi pu retrouver les principaux résultats de Hansen et Gegenfurtner (2009) montrant que les contenus de luminance et de chrominance des images de scènes naturelles étaient en partie indépendants, et ce d'autant plus au cours des étapes du traitement visuel. Notre étude souligne également l'ampleur de la variabilité inter-images et pointe la nécessité de disposer d'une description mécanistique de la chrominance pertinente pour estimer correctement l'information mutuelle.

5.1.2 Description mécanistique de la vision des couleurs

Dans le **Chapitre 3**, nous nous sommes attachés à fournir une description mécanistique de la chrominance adaptée à la perception des scènes naturelles. Nous avons pour cela mené une série d'expériences de détection de couleur et de luminance dans des images de scènes naturelles et proposé un modèle pour l'encodage de l'information chromatique et la prise de décision dans cette tâche. Nos résultats ont montré un avantage des définitions intégrant une étape d'adaptation, que ce soit au niveau des réponses des photorécepteurs ($RG = L_{log} - M_{log}$ avec $L_{log} = \log(L) - \overline{\log(L)}$ et $M_{log} = \log(M) - \overline{\log(M)}$) ou des canaux post-récepteurs ($RG = \frac{L - M}{L + M}$), par rapport à des mécanismes strictement linéaires ($RG = L + M$) (Breuil, Guyader et al., soumis). Des expériences contrôlées nous ont permis de vérifier que les variations inter-individuelles dans la pondération des différents types de cônes dans les canaux post-récepteurs ne suffisaient pas, à elles seules, à expliquer nos données expérimentales, soutenant donc l'existence de processus d'adaptation dans les mécanismes de traitement de la couleur. Si cette étude ne nous a pas permis de trancher quant à la supériorité de l'un ou l'autre des processus d'adaptation considérés pour expliquer les données expérimentales, elle a cependant

été l'occasion de développer un paradigme expérimental et un cadre théorique qui semblent pertinents pour l'étude de la couleur dans la perception des scènes naturelles.

5.1.3 Rôle de la couleur pour la classification de contours

Dans le **Chapitre 4**, nous avons testé le rôle de l'information de chrominance dans une tâche spécifique : la classification de contours extraits d'images de scènes naturelles dans les catégories ombre ou autre changement. Nos données psychophysiques, tout comme les performances des classifieurs bayésiens basés sur des propriétés simples des images (contraste de luminance, luminance moyenne, différence de contrastes, netteté du contour, variation de la chrominance moyenne) que nous avons construits pour réaliser la catégorisation, montrent que la chrominance constitue un indice utile pour cette tâche (Breuil, Jennings et al., soumis).

L'ensemble de ces résultats nous a permis d'avancer dans la compréhension du rôle de la couleur dans la perception de scènes et de mieux apprécier les mécanismes à l'oeuvre dans l'extraction des informations de luminance et de chrominance dans un contexte écologique. Afin de saisir pleinement le rôle de ces informations dans la perception visuelle, il semble indispensable de mener des expériences sur des images de scènes naturelles, et c'est les perspectives que nous proposons.

5.2 Perspectives

Dans cette thèse, nous avons proposé des outils méthodologiques, des modèles théoriques et des paradigmes expérimentaux qui nous semblent particulièrement pertinents pour l'étude du rôle de la couleur dans la perception des scènes naturelles et pourraient être réutilisés et améliorés à cette fin. Les outils que nous avons développés peuvent, notamment, être appliqués à l'analyse d'un plus grand nombre d'images. Ils peuvent également permettre de développer de nouveaux paradigmes expérimentaux et de nouveaux modèles pour étendre l'étude du rôle de l'information de chrominance à d'autres tâches visuelles, telles que la catégorisation de scènes, la segmentation d'objets, l'attention ou encore la mémorabilité.

5.2.1 Lien entre information mutuelle et rôle de la couleur

L'analyse que nous avons réalisée sur les statistiques des images, montre que chrominance et luminance constituent deux sources d'information partiellement indépendantes (dont le degré d'indépendance varie en fonction des images). Comme Hansen et Gegenfurtner (2009) l'avaient déjà souligné, ce résultat implique que chacune de ces informations peut donc être exploitée par le système visuel de diverses façons et pour différents objectifs. Une question particulièrement importante est de savoir quand ces informations, relayées en partie par des canaux indépendants, sont combinées par le système visuel. Les processus d'extraction des contours de luminance et de chrominance dans les scènes naturelles sont-ils complètement

indépendants ? Si oui, les indices qu'ils fournissent sont-ils qualitativement différents ? C'est-à-dire, les contours de chrominance ont-ils, par exemple, une importance particulière pour distinguer les objets ? Il semble nécessaire d'étudier le rôle de l'information de chrominance dans des tâches spécifiques pour pouvoir répondre à ces questions. Nous avons pour cela proposé, entre autres, l'expérience de classification des contours présentée au Chapitre 4 et soulignant un rôle des contrastes de chrominance pour la réalisation de cette tâche. Nous avons, par ailleurs, mené deux autres séries d'expériences que nous n'avons pas présentées dans ce manuscrit, mais qui constituent des pistes intéressantes et que nous décrivons succinctement ici : des expériences d'exploration libre de scènes naturelles en oculométrie et des expériences sur le phénomène de remplissage des couleurs.

5.2.1.1 Étude du rôle de la couleur dans l'attention visuelle

Notre système visuel est submergé à chaque instant par un flot d'information trop dense et complexe pour pouvoir être saisi dans sa globalité. Ne pouvant allouer les mêmes ressources cognitives à chacun des détails de notre environnement, le système visuel doit donc sélectionner et filtrer l'information la plus pertinente dans une situation donnée afin d'optimiser le traitement perceptif. C'est ce processus que nous nommons *attention*. L'étude des mouvements oculaires est, à ce titre, particulièrement intéressante car, dans les situations écologiques, le mouvement des yeux suit celui de l'attention (Rizzolatti et al., 1987), offrant ainsi une fenêtre d'observation directe des mécanismes cognitifs impliqués dans les processus attentionnels. Depuis les travaux de Yarbus (1967), de nombreuses études ont pu mettre en évidence différents facteurs influençant l'exploration visuelle de scènes naturelles. La couleur est souvent considérée comme l'un de ces attributs et joue un rôle important dans les modèles d'attention visuelle qui ont été développés (Itti et al., 1998), au même titre que les propriétés de luminance ou d'orientation. Pour comprendre quels attributs attirent le plus notre attention dans une image, plusieurs études ont cherché à établir un lien entre les régions fixées par des observateurs lors d'explorations libres des images et les propriétés physiques de ces régions (luminance, couleur, contrastes,...). Afin d'évaluer, en particulier, le rôle de la couleur dans l'attention visuelle, des travaux se sont intéressés à la comparaison des mouvements oculaires sur des images en couleur et des images en niveaux de gris (Frey, Honey et al., 2008 ; Ho-Phuoc et al., 2012). Ces études ont montré que les différences entre les positions des fixations pour ces deux conditions étaient relativement faibles et dépendaient plus du contenu sémantique des images que de l'information de chrominance, mais que les durées de fixation étaient en moyenne plus élevées pour les images en niveaux de gris que pour les images en couleur (Ho-Phuoc et al., 2012). Une des problématiques rencontrées dans l'interprétation des résultats de ces études est qu'elles utilisent des espaces de représentation des couleurs différents et des méthodes différentes pour réaliser la conversion des images couleur en niveaux de gris. Il est ainsi difficile de comparer les résultats des différentes études et d'estimer l'effet véritable de l'information de chrominance.

Afin d'évaluer le rôle de la couleur dans l'attention visuelle, nous avons également, dans un premier temps, conduit une expérience d'exploration libre d'images de scènes naturelles comparant le cas d'images en couleurs et en niveaux de gris. Nous nous sommes, en particulier, intéressés à l'effet de la méthode de conversion en niveaux de gris en comparant trois

techniques différentes. Cette étude a été réalisée en collaboration avec David Alleysson et Alan Chauvin, du Laboratoire de Psychologie et NeuroCognition (Université Grenoble Alpes), et mise en place dans le cadre du stage de deuxième année d'école d'ingénieur d'Audrey Brouard (Grenoble INP). Nous avons, pour cette expérience, sélectionné 90 images de scènes naturelles. 41 participants ont pris part à l'expérience qui se déroulait en deux sessions au cours desquelles il leur était demandé d'observer des images de scènes naturelles, présentées chacune durant 4 secondes dans un ordre aléatoire. Lors de la première session, les images étaient en couleurs. Environ 15 jours plus tard, les mêmes images leur étaient à nouveau présentées, mais cette fois en niveaux de gris.

Pour la conversion en niveaux de gris, nous avons utilisé trois méthodes différentes :

- La première moyennant les valeurs des pixels R, G et B
- La seconde correspondant à la définition de $V(\lambda)$ et utilisée pour les autres expériences décrites dans le manuscrit
- La dernière adaptée à chaque participant, basée sur une évaluation des différences inter-individuelles à l'aide d'une technique de *minimum motion* (Anstis et al., 1983).

Des exemples d'images utilisées pour cette expérience sont présentées sur la Figure 5.1.

Nous avons ensuite analysé les saccades et les fixations réalisées par chacun des participants au cours des deux sessions (Figure 5.2) et comparé les corrélations de celles-ci entre images en couleurs et images en niveaux de gris pour les trois conversions utilisées.

Nos résultats ont mis en évidence un léger avantage pour la conversion adaptée au participant, pour laquelle les trajectoires d'exploration étaient plus proches de celles observées dans la condition couleur relativement aux autres méthodes de conversion en niveaux de gris. Plusieurs biais potentiellement présents dans notre expérience ne nous permettent pas de tirer de conclusion définitive de ces résultats, mais cette étude nous a permis de mettre en lumière deux problématiques importantes qui ont, en partie, orienté la suite de nos travaux. Premièrement, nous avons utilisé trois groupes d'images différents et constants pour chacune des conversions. Il se pourrait donc que ce choix soit, en partie, ou totalement, responsable de l'effet de groupe observé dans nos résultats. Il nous a donc paru indispensable de disposer d'un indice pertinent pour la constitution des groupes et pour la formulation d'hypothèses plus précises concernant le rôle de la couleur dans l'attention visuelle. Il semble notamment nécessaire de pouvoir quantifier l'information de couleur présente dans une image et, en particulier, l'indépendance de celle-ci avec le contenu de luminance pour prédire si la suppression de la couleur est susceptible de modifier l'exploration visuelle ou non. C'est cette conclusion qui a motivé notre étude sur les analyses statistiques des scènes naturelles et la proposition de l'information mutuelle comme estimation de l'indépendance statistique des contenus de luminance et de chrominance des images. Deuxièmement, cette expérience a également souligné la problématique du choix de définitions mécanistiques adaptées de la luminance et de la chrominance, et l'importance des variations inter-individuelles dans ces mécanismes.

À la suite de l'étude réalisée sur la mesure de l'indépendance des contenus de luminance et de chrominance des images de scènes naturelles (Chapitre 2), nous avons mené une nouvelle expérience d'exploration libre, basée sur les valeurs d'information mutuelle obtenues. La mise



FIGURE 5.1 – Exemple d’images en couleurs utilisées pour notre première expérience d’exploration libre. (Remarque : le rendu peu naturel des images est dû au fait que nous avons supprimé l’information initialement contenue dans les pixels bleus, en fixant leurs valeurs par rapport à celles des deux autres pixels ($B = \frac{R + G}{2}$).)

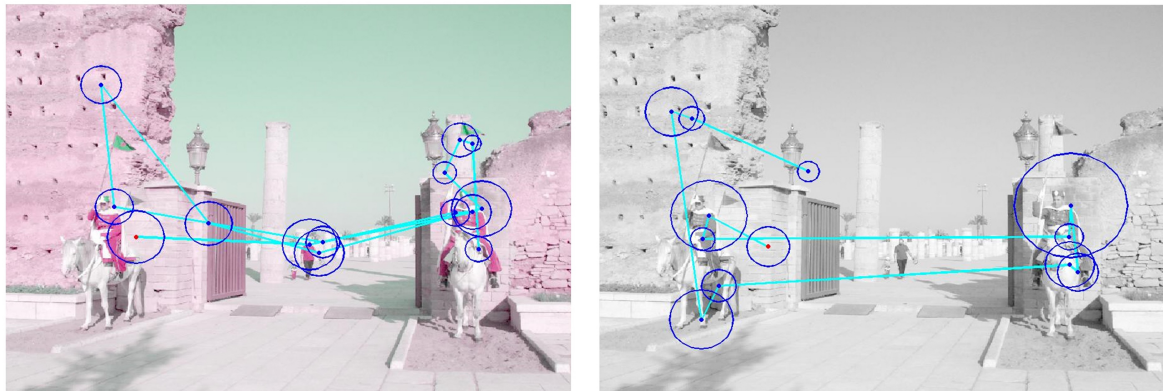


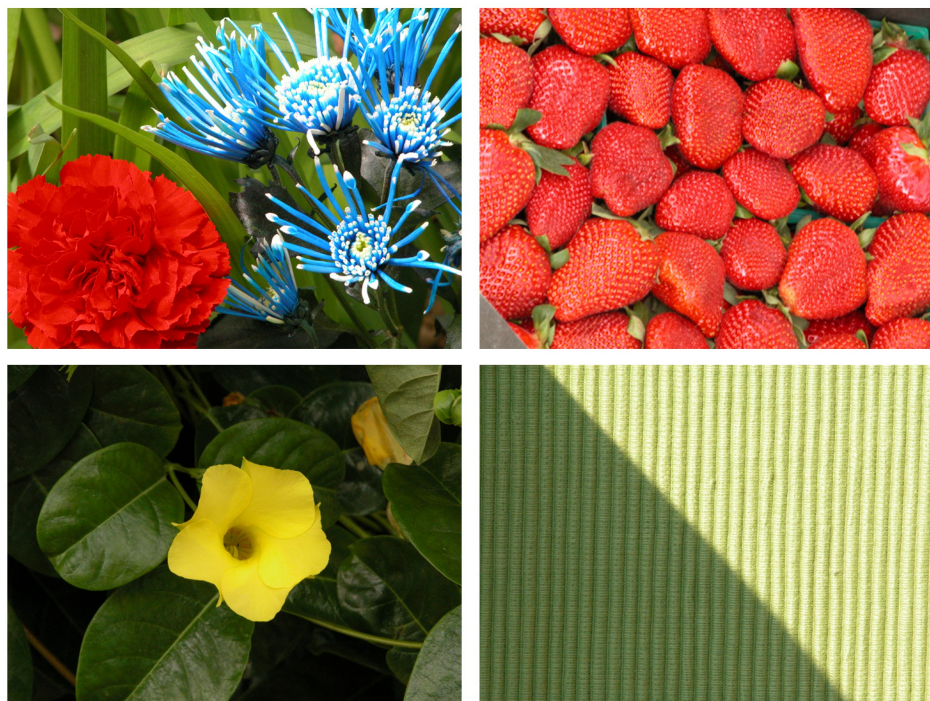
FIGURE 5.2 – Mouvements oculaires enregistrés pour un participant de l'expérience sur la même image présentée en couleurs (à gauche) et en niveaux de gris (à droite). Les cercles correspondent aux fixations et leur taille est proportionnelle à leur durée. Les trajectoires d'exploration sont légèrement différentes dans les deux conditions : en particulier, certains détails semblent avoir attiré le regard du participant dans le cas de l'image en couleurs, mais pas dans celle en niveaux de gris.

en place du protocole expérimental et les passations ont été effectuées par Léo Martin et Bastien Faure-Brac, dans le cadre de leur projet de 3^{ème} année (Polytech' Grenoble). Nous avons sélectionné 80 images de scènes naturelles, réparties en deux groupes : l'un pour lequel l'information mutuelle entre contours de luminance et de chrominance des images était élevée (IM forte) et l'autre pour lequel l'information mutuelle était faible (IM faible). Nous avons ensuite demandé à 14 participants d'explorer ces images durant 5 secondes, pendant que nous enregistrions leurs mouvements oculaires. Deux hypothèses peuvent être formulées concernant le lien entre information mutuelle et attention visuelle. La première, volontairement vague, est que la redondance entre information de luminance et de chrominance influence l'exploration visuelle de manière générale. On peut ainsi analyser différentes variables, telles que la durée des fixations ou l'amplitude des saccades et comparer leurs valeurs moyennes pour les différentes images aux valeurs d'information mutuelle. La seconde hypothèse que nous formulons, est que plus l'information de chrominance est différente de l'information de luminance, plus les durées de fixation seront élevées (car la quantité d'information globale à extraire sera plus importante) et plus la suppression de la couleur modifiera les trajectoires d'exploration si cette information est effectivement utilisée pour orienter l'attention.

Nous avons également analysé les propriétés locales de luminance et de chrominance aux positions des fixations des observateurs. Cette étude ne nous a pas permis, à ce stade, de mettre en évidence un lien trivial entre information mutuelle et rôle de la couleur dans l'attention visuelle dans le cas de l'exploration libre de scènes naturelles. Nous expliquons ce résultat en partie par la présence d'un trop grand nombre de variables non maîtrisées dans la sélection de nos stimuli et par les limites évoquées précédemment que comporte l'information mutuelle en tant que seul indice d'indépendance des contenus de luminance et de chrominance à l'échelle d'une image. Nous pensons, cependant, que cette étude mériterait d'être approfondie. Une piste intéressante serait également d'utiliser dans un premier temps des images plus artificielles pour



(a) Quatre images du groupe de faible information mutuelle entre contours de luminance et contours de chrominance



(b) Quatre images du groupe de forte information mutuelle entre contours de luminance et contours de chrominance

FIGURE 5.3 – Exemple d’images de chacun des deux groupes (IM faible et IM forte) utilisées pour la seconde expérience d’exploration libre.

étudier les cas d'images nulles, c'est-à-dire ne présentant aucune redondance d'information, et d'images parfaitement redondantes afin de préciser les hypothèses concernant le lien entre information mutuelle et rôle de la couleur dans l'attention visuelle.

5.2.1.2 Étude du rôle de la couleur dans les processus de remplissage

Une autre hypothèse concernant le traitement de l'information de couleur par le système visuel est que celle-ci interviendrait dans des processus de remplissage (Grossberg, 1987 ; Grossberg, 2000 ; Neumann et al., 2001). Selon cette hypothèse, le système visuel exploiterait dans un premier temps l'information de luminance afin d'estimer la structure globale d'une scène puis, dans un second temps, l'information de couleur serait utilisée pour déterminer les propriétés de surface des objets. Une question est notamment de savoir dans quelle mesure l'information de couleur que nous utilisons pour ces processus est directement extraite de l'environnement visuel ou en partie déterminée par les connaissances préalables que nous avons de celui-ci. Plusieurs études ont ainsi montré que la connaissance que nous avons de la couleur des objets en modulait notre perception (Hansen, Olkkonen et al., 2006 ; Olkkonen et al., 2008). D'autres travaux ont mis en évidence l'existence de phénomènes dits de complétion ou de remplissage des couleurs. En effet, l'information de couleur à laquelle a réellement accès le système visuel est, dans un contexte naturel, toujours incomplète (faible densité des cônes en périphérie de la rétine, tâche aveugle, caractère stochastique de l'absorption des photons et de la transduction,...) et il doit donc estimer et reconstruire l'information manquante pour obtenir un percept stable et cohérent. Ce phénomène de remplissage des couleurs a été mis en évidence pour des stimuli simples, dans le cas, par exemple, de la célèbre illusion *watercolor* (Pinna, 2008) (Fig. 5.4), mais également pour des stimuli plus complexes tels que des scènes naturelles (Balas et Sinha, 2007). Il serait intéressant d'évaluer si la quantité d'information de chrominance effectivement présente dans une image influence ces phénomènes de remplissage et d'étudier le rôle des traitements de bas-niveau (extraction de l'information de chrominance) et de haut-niveau (par exemple, l'influence de nos connaissances préalables sur l'apparence colorée des objets) pour la perception des couleurs dans les scènes naturelles.

Nous avons cherché à explorer cette piste en menant une expérience de catégorisation, inspirée de celle de Balas et Sinha (2007), dans laquelle nous présentions aux participants des images de scènes naturelles auxquelles nous avons localement retiré le contenu de chrominance initial, au centre ou en périphérie, pour des régions de trois tailles différentes. Cette étude a été réalisée dans le cadre du stage de Master 1 de Léa Entzmann (Sciences Cognitives). Nous désignons sous le nom de *chimères* ces images auxquelles nous avons retiré une partie de l'information de chrominance et nous appelons *scotome* la région circulaire située au centre de l'image qui est la seule où le contenu de chrominance est, soit supprimé, soit conservé. Nous avons sélectionné 50 images, réparties en deux groupes : scènes naturelles et scènes artificielles (Figure 5.5). Nous incluons ici dans la catégorie scènes naturelles les images de paysages (forêts, champs, rivages,...), tandis que la catégorie scènes artificielles regroupent les images présentant des objets manufacturés (immeubles, véhicules, scènes d'intérieur,...). Chaque image a ensuite été déclinée en six chimères : trois pour lesquelles l'information de chrominance était retirée au centre et trois pour lesquelles elle était retirée en périphérie,

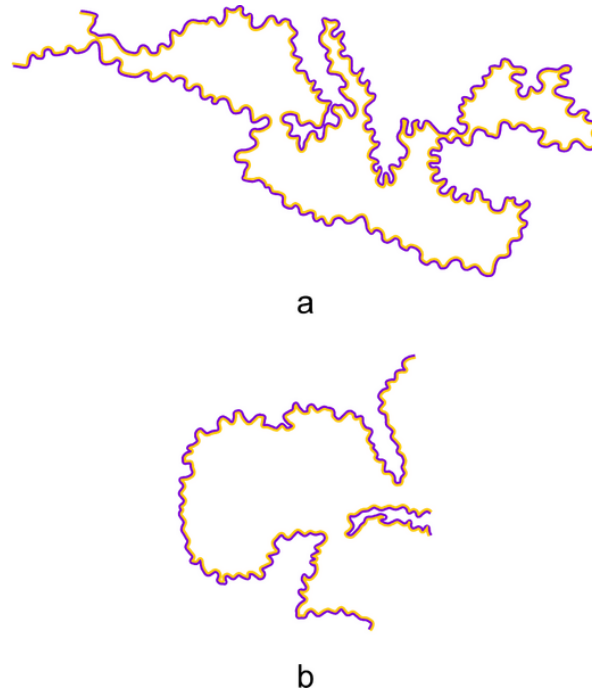


FIGURE 5.4 – Effet *watercolor* : En juxtaposant simplement des contours violets et oranges, on crée une illusion de remplissage des surfaces par une teinte orangée (d’après Pinna (2008)).

pour trois tailles de scotome différentes (Figure 5.6). Au cours de l’expérience, chacune des six chimères était présentée une fois au participant et l’image couleur correspondante était présentée six fois pour équilibrer les conditions. Au total, l’expérience consistait donc en 600 essais, une image étant présentée durant 80 ms à chacun d’entre eux. Les observateurs devaient déterminer, à chaque essai, si l’image qui leur était présentée était entièrement colorée (image couleur) ou seulement partiellement (image chimère).

Nous pouvons formuler différentes hypothèses concernant les résultats de cette expérience :

- S’il existe effectivement des phénomènes de remplissage, on s’attend à un nombre important de fausses alarmes (FAs), c’est-à-dire à ce que les participants répondent avoir perçu une image couleur lorsque celle-ci était en réalité une image chimère.
- Si la quantité totale d’information de chrominance présente dans l’image influence l’effet de remplissage, on s’attend à observer un effet de la taille des scotomes sur le nombre de fausses alarmes.
- Si ces phénomènes de remplissage dépendent de nos connaissances préalables sur notre environnement, alors on devrait observer plus de fausses alarmes pour les images de la catégorie scènes naturelles, où la couleur *a priori*, une valeur diagnostique plus élevée, que pour celles de la catégorie scènes artificielles.
- Enfin, on peut supposer que plus le contenu de chrominance de l’image initiale est indépendant du contenu de luminance, plus les participants détecteront une absence

partielle de celui-ci, et leurs taux de fausses alarmes devraient donc être plus faibles.



FIGURE 5.5 – Exemple d’images utilisées pour construire les stimuli de l’expérience de remplissage. En haut : deux images de la catégorie scènes naturelles. En bas : deux images de la catégorie scènes artificielles.

Nos résultats nous ont permis de vérifier nos deux premières hypothèses (Figure 5.7) : nous avons, en effet, mesuré un nombre important de fausses alarmes pour l’ensemble de nos participants et une influence de la taille du scotome sur celui-ci (plus la quantité d’information de chrominance présente est importante dans l’image chimère, plus la tâche est difficile). Ces résultats suggèrent donc l’existence de processus de remplissage des couleurs, qui semblent être d’autant plus efficaces que l’information manquante est faible. Cette interprétation est peut-être un peu directe et nous pensons que d’autres explications pourraient être proposées mais c’est l’hypothèse de l’étude Balas et Sinha (2007) et nous ne développerons pas plus ce point ici. Contrairement à nos hypothèses, nous avons, en revanche, observé un nombre de fausses alarmes plus important pour les images de la catégorie scènes artificielles que pour celles de la catégorie scènes naturelles. Ce résultat suggère que la valeur diagnostique de la couleur aurait donc finalement peu d’influence sur les phénomènes de remplissage. Nous n’avons également pas observé de corrélation entre l’indépendance des informations de luminance et de chrominance des images couleur (mesurée au travers des valeurs d’information mutuelle entre contours) et les performances des participants. Ces résultats sont malgré tout intéressants pour comprendre un peu mieux ces phénomènes de remplissage et le rôle des traitements de



(a) Chimères construites à partir d'une image de la catégorie scènes naturelles.



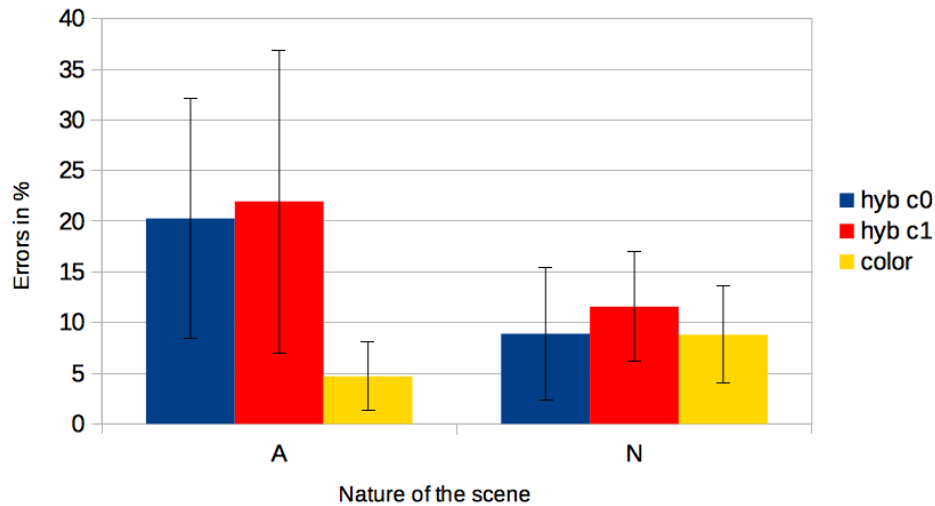
(b) Chimères construites à partir d'une image de la catégorie scènes artificielles.

FIGURE 5.6 – Exemple des six chimères construites pour une image de chacune des catégories. En haut : l'information de chrominance est retirée de la partie centrale. En bas : l'information de chrominance est retirée en périphérie. De gauche à droite : le diamètre du scotome augmente.

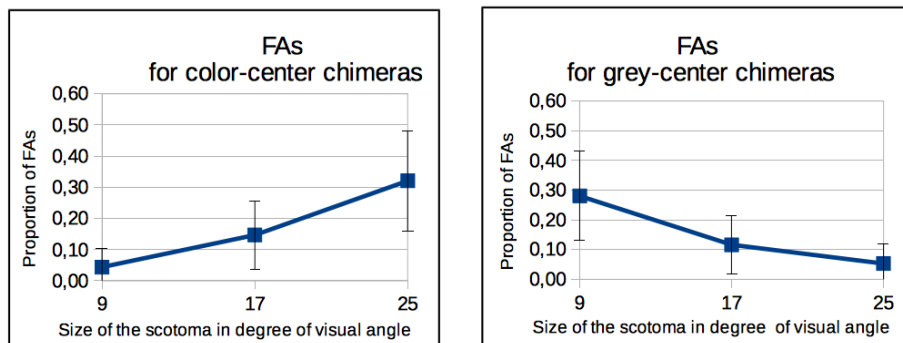
bas-niveau, réalisés sur l'information de chrominance extraite des images, et de processus de plus haut-niveau pour la construction des percepts de couleur.

Nous avons, par la suite, étendu notre étude au cas d'images où nous avons aléatoirement permuté la position de blocs de pixels pour déconstruire la structure spatiale des scènes et étudier plus précisément l'influence respectives des facteurs de bas-niveau et de haut-niveau (Figure 5.8). Ces études méritent de plus amples investigations mais constituent des pistes intéressantes pour la compréhension de ces phénomènes.

De façon plus générale, pour l'ensemble des travaux présentés dans ce manuscrit nous pouvons également séparer les stimuli utilisés en deux groupes (naturels et artificiels) et comparer les résultats obtenus pour chacun d'entre eux. Une analyse préliminaire ne nous a pas permis d'obtenir de cette façon des conclusions évidentes, mais il s'agit également d'une perspective pertinente pour étudier l'influence des facteurs de bas-niveau et de haut-niveau sur la perception des couleurs dans les scènes naturelles.

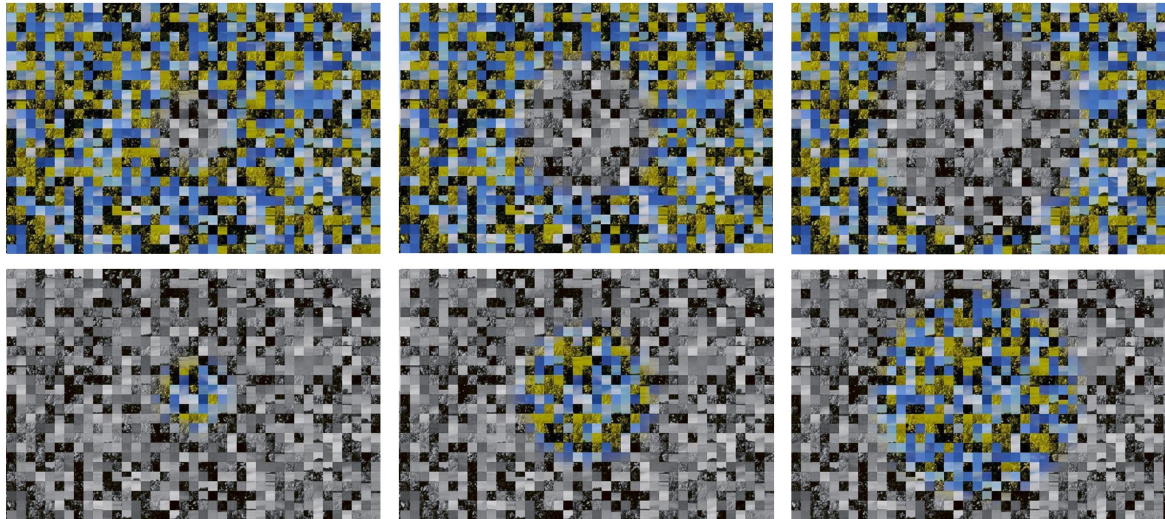


(a) Taux d'erreurs pour les images artificielles (à gauche) et naturelles (à droite). En bleu : les chimères pour lesquelles l'information de chrominance a été retirée de la partie centrale de l'image (hyb c0). En rouge : les chimères pour lesquelles l'information de chrominance a été retirée de la partie périphérique de l'image (hyb c1). En jaune : les images couleur (color).

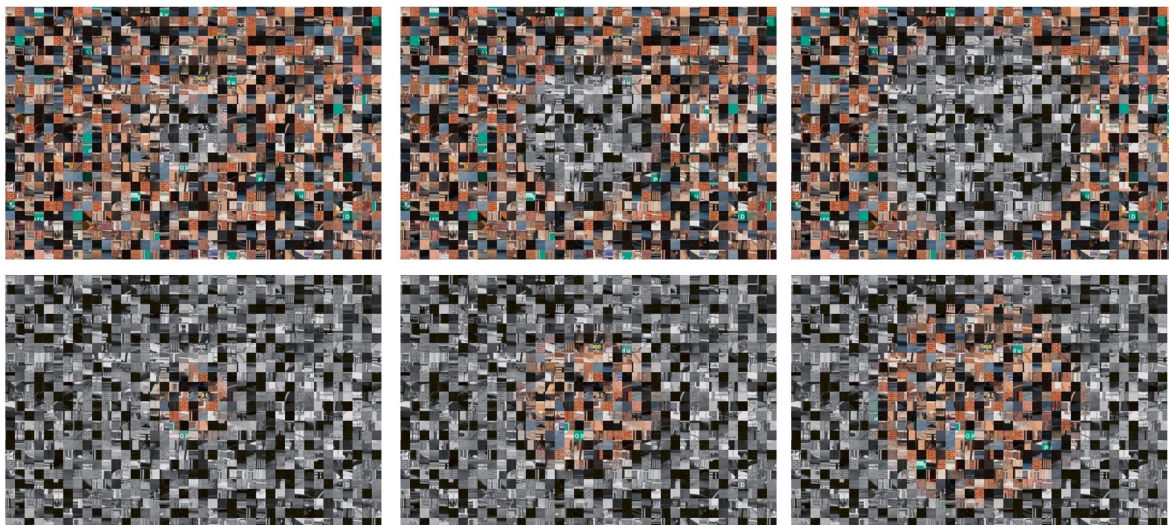


(b) Nombre de fausses alarmes (FAs) enregistrées en fonction de la taille du scotome, pour les images chimères où l'information a été retirée dans la partie centrale (à gauche) ou en périphérie (à droite).

FIGURE 5.7 – Résultats de l'expérience de catégorisation des images couleur et chimères. Figures extraites du rapport de stage de Léa Entzmann (2017).



(a) Chimères construites à partir d'une image de la catégorie scènes naturelles.



(b) Chimères construites à partir d'une image de la catégorie scènes artificielles.

FIGURE 5.8 – Exemple des six chimères construites à partir des deux même images que précédemment, dans le cadre de l'expérience avec permutation des blocs de pixels. En haut : l'information de chrominance est retirée de la partie centrale. En bas : l'information de chrominance est retirée en périphérie. De gauche à droite : le diamètre du scotome augmente.

5.2.2 Un modèle mécanistique pour la perception des couleurs dans les scènes naturelles

Dans les chapitres 3 et 4, nous avons proposé des modèles d'observateurs bayésiens utilisant des propriétés basiques des images pour réaliser nos tâches expérimentales. Bien que très simples, ces modèles se sont révélés assez efficaces pour expliquer nos données empiriques. Dans le chapitre 3, nous avons notamment pu montrer qu'une représentation simple des couleurs, basée sur l'espace tri-dimensionnel DKL (Derrington et al., 1984), et une description mécanistique du traitement visuel intégrant une étape d'adaptation au niveau des photorécepteurs ou des canaux post-récepteurs, permettait de bien expliquer les performances de détection des participants dans le cadre du modèle général que nous avons proposé. Cette approche semble donc pertinente pour décrire le traitement de l'information de couleur dans le cas des scènes naturelles, et nous avons ainsi pu mettre en évidence deux modèles mécanistiques de la vision des couleurs qui semblaient adaptés à la perception des scènes naturelles (Figure 5.9). Il serait cependant intéressant de parvenir à préciser les mécanismes d'adaptation entrant en jeu et les étapes du traitement au cours desquelles ils interviennent.

Le paradigme expérimental de détection de couleur présenté au Chapitre 3 pourrait également être exploité en vue de mesurer les variations inter-individuelles dans la pondération des différents types de cônes au niveau des canaux post-récepteurs. Les paramètres caractérisant le décalage des canaux de l'observateur avec les directions cardinales de l'espace DKL pourraient, en effet, être intégrés aux paramètres à ajuster du modèle afin d'évaluer leurs valeurs les plus vraisemblables pour chacun des observateurs.

Les différents modèles que nous avons proposés dans ces travaux pourraient également être affinés, en y intégrant, notamment, des propriétés spatio-temporelles pour expliquer plus efficacement les données expérimentales.

Nous nous sommes de plus limités ici à l'étude du rôle de la couleur pour quelques tâches spécifiques, mais il serait intéressant d'étendre ces analyses à d'autres tâches concernant le traitement des scènes complexes pour évaluer l'utilité de la couleur à différents niveaux du traitement visuel. Le cas de l'information de chrominance Bleu-Jaune a également été assez peu abordé afin de simplifier nos expériences et nos analyses, mais elles mériteraient d'être par la suite étendues à celui-ci.

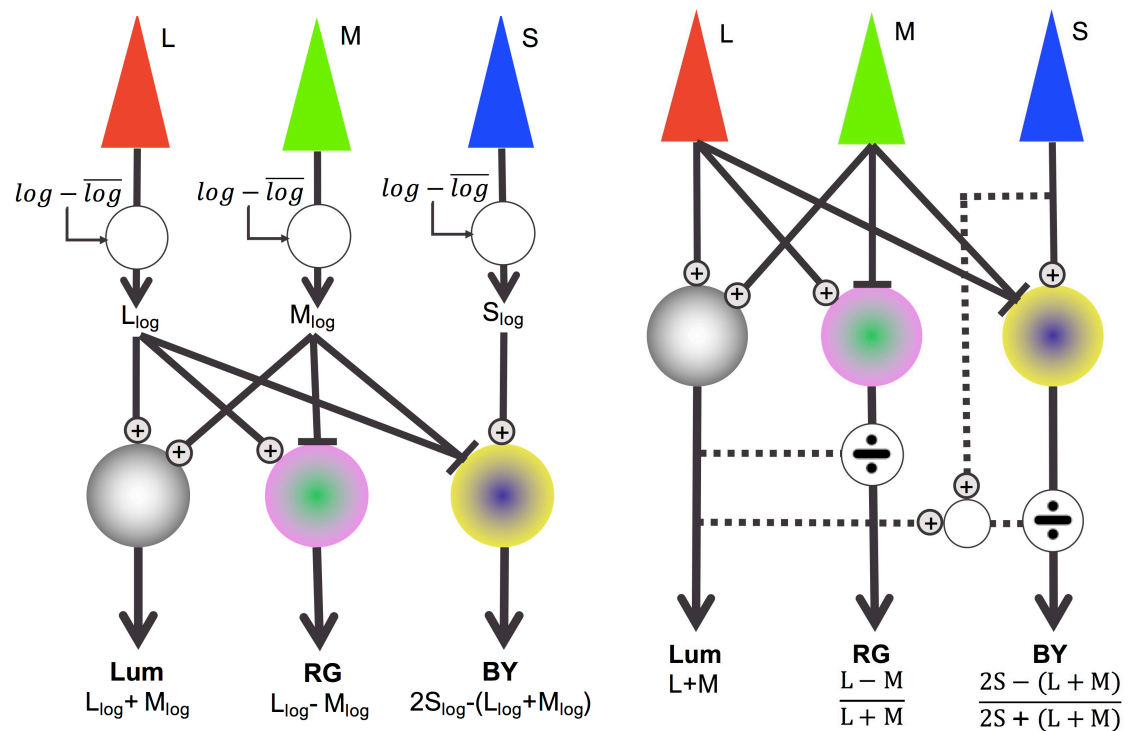


FIGURE 5.9 – Deux descriptions mécanistiques de la vision des couleurs adaptées à la perception des scènes naturelles.

Finalement, nos résultats suggèrent des interactions assez précoces entre information de luminance et de chrominance, et questionnent la pertinence d'une représentation modulaire de ces traitements pour la perception de scènes complexes. Il semble cependant que les modèles proposés et les définitions mécanistiques testées puissent se révéler adaptés pour caractériser le traitement de la couleur dans la perception des scènes naturelles. La nature exacte des mécanismes sous-tendant la perception des couleurs dans un contexte naturel demeure malgré tout en partie inexpliquée et la saisir pleinement constitue une entreprise vaste et ambitieuse. Si nous ne pouvons pas ici formuler de conclusions définitives concernant le traitement de l'information de chrominance opéré par le système visuel et le rôle de celle-ci dans la perception des scènes naturelles, nos travaux offrent néanmoins de nombreuses perspectives pour la compréhension de ces processus, l'étude des variations inter-individuelles et l'analyse des facteurs de bas et de haut niveaux impliqués dans la perception des couleurs dans un contexte écologique.

Bibliographie

- Akaike, H. (1974). “A New Look at the Statistical Model Identification”. *IEEE Transactions on Automatic Control* 19.6, p. 716–723 (cf. p. 95).
- Allen, G. (1879). *The colour-sense : Its origin and development. An essay in comparative psychology*. Boston, Houghton (cf. p. 5).
- Amano, K. et Foster, D. H. (2014). “Influence of local scene color on fixation position in visual search”. *JOSA A* 31.4, A254–A262 (cf. p. 6).
- Amano, K., Foster, D. H. et al. (2012). “Visual search in natural scenes explained by local color properties”. *JOSA A* 29.2, A194–A199 (cf. p. 6).
- Anstis, S. M. et al. (1983). “A minimum motion technique for judging equiluminance”. *Colour vision : Physiology and psychophysics*, p. 155–166 (cf. p. 137).
- Attneave, F. (1954). “Some informational aspects of visual perception.” *Psychological review* 61.3, p. 183 (cf. p. 34).
- Balas, B. et Sinha, P. (2007). ““Filling-in” colour in natural scenes”. *Visual Cognition* 15.7, p. 765–778 (cf. p. 141, 143).
- Barlow (1961). *Possible Principles Underlying the Transformations of Sensory Messages* (cf. p. 34).
- Barlow, H. (2001). “Redundancy reduction revisited”. 12, p. 241–253 (cf. p. 34).
- Barlow, H. B. (1980). “The Ferrier lecture. Critical limiting factors in the design of the eye and visual cortex.” *Proc. R. Soc. Lond. B* 212.1186, p. 1–34 (cf. p. 5).
- Barthelmé, S. et Mamassian, P. (2011). “Evaluation of Objective Uncertainty in the Visual System”. *PLoS Computational Biology* 101.11, p. 1435–1439 (cf. p. 37).
- Baylor, D. A. (1987). “Photoreceptor signals and vision. Proctor lecture.” *Investigative ophthalmology & visual science* 28.1, p. 34–49 (cf. p. 13).
- Biederman, I. et Gerhardstein, P. C. (1993). “Recognizing depth-rotated objects : evidence and conditions for three-dimensional viewpoint invariance.” *Journal of Experimental Psychology : Human perception and performance* 19.6, p. 1162 (cf. p. 6).
- Boynton, R. M. (1996). “Frederic Ives Medal paper. History and current status of a physiologically based system of photometry and colorimetry”. *Journal of the Optical Society of America A* 13.8, p. 1609–1621 (cf. p. 16).
- Bradley, A., Switkes, E. et De Valois, K. (1988). “Orientation and spatial frequency selectivity of adaptation to color and luminance gratings”. *Vision research* 28.7, p. 841–856 (cf. p. 22).
- Brady, N. et Field, D. J. (2000). “Local contrast in natural images : normalisation and coding efficiency”. *Perception* 29.9, p. 1041–1055 (cf. p. 31).
- Brainard, D. (1996). “Cone contrast and opponent modulation color spaces”. *Human color vision*, p. 563–579 (cf. p. 81, 169).

- Brainard, D. H. (1997). “The Psychophysics Toolbox”. *Spatial Vision*, 10(4), 433-436 (cf. p. 37, 85).
- Brainard, D. et Gazzaniga, M. S. (2009). “Bayesian approaches to color vision”. *The visual neurosciences* 4 (cf. p. 37).
- Brainard, D. H. (2015). “Color and the Cone Mosaic”. *Annual Review of Vision Science* 1.1, p. 519–546 (cf. p. 15).
- Brainard, D. H. et Wandell, B. A. (1986). “Analysis of the retinex theory of color vision”. *Journal of the Optical Society of America A* 3.10, p. 1651 (cf. p. 81).
- Bramão, I. et al. (2011). “The role of color information on object recognition : A review and meta-analysis”. *Acta psychologica* 138.1, p. 244–253 (cf. p. 5).
- Breuil, C., Guyader, N. et Barthelmé, S. (soumis). “Detecting color in natural scenes : evidence for early interactions between luminance and chrominance”. *PLoS Computational Biology* (cf. p. 134).
- Breuil, C., Jennings, B. J. et al. (soumis). “Color improves edge classification”. *PLoS Computational Biology* (cf. p. 135).
- Buchsbaum, G. (1980). “A spatial processor model for object colour perception”. *Journal of the Franklin Institute* 310.1, p. 1–26 (cf. p. 81).
- Buchsbaum, G. et Gottschalk, A. (1983). “Trichromacy, Opponent colours coding and optimum colour information transmission in retina”. *Proceedings of the Royal Society of London* 220.1218, p. 89–113 (cf. p. 42, 54, 71, 72).
- Burton, G. J. et Moorhead, I. R. (1987). “Color and spatial structure in natural scenes”. *Applied Optics* 26.1, p. 157–170 (cf. p. 42).
- Buzás, P. et al. (2006). “Specificity of M and L cone inputs to receptive fields in the parvocellular pathway : random wiring with functional bias”. *Journal of Neuroscience* 26.43, p. 11148–11161 (cf. p. 22).
- Carandini, M. et Heeger, D. (2012). “Normalization as a canonical neural computation.” *Nature Reviews Neuroscience* November, p. 1–12 (cf. p. 82).
- Carroll, J., Neitz, J. et Neitz, M. (2002). “Estimates of L : M cone ratio from ERG flicker photometry and genetics”. *Journal of Vision* 2.8, p. 1 (cf. p. 15).
- Castelhano, M. S. et Henderson, J. M. (2008). “The influence of color on the perception of scene gist.” *Journal of Experimental Psychology : Human Perception and Performance* 34.3, p. 660–675 (cf. p. 5).
- Cavanagh, P. (1987). “Reconstructing the third dimension : Interactions between color, texture, motion, binocular disparity, and shape”. *Computer Vision, Graphics, and Image Processing* 37.2, p. 171–195 (cf. p. 5).
- (1991). “Vision at equiluminance”. *Vision and visual dysfunction : Limits of vision* 5, p. 234–250 (cf. p. 5).
- Changizi, M. A., Zhang, Q. et Shimojo, S. (2006). “Bare skin, blood and the evolution of primate colour vision”. *Biology Letters* 2.2, p. 217–221 (cf. p. 5).
- Chaparro, A. et al. (1994). “Separable red-green and luminance detectors for small flashes”. *Vision Research* 34.6, p. 751–762 (cf. p. 99).
- Cheong, S. K. et al. (2011). “Slow intrinsic rhythm in the koniocellular visual pathway”. *Proceedings of the National Academy of Sciences* 108.35, p. 14659–14663 (cf. p. 22).

- Chichilnisky, E. J. et Baylor, D. A. (1999). "Receptive-field microstructure of blue-yellow ganglion cells in primate retina". *Nature neuroscience* 2.10, p. 889 (cf. p. 22).
- CIE (2006). "Fundamental Chromaticity Diagram with Physiological Axes-Part 1". *CIE*. T. 170 (cf. p. 20).
- Cole, G. R., Stromeyer III, C. F. et Kronauer, R. E. (1990). "Visual interactions with luminance and chromatic stimuli". *Journal of the Optical Society of America A* 7.1, p. 128–140 (cf. p. 81).
- Cole, G. R., Hine, T. et McIlhagga, W. (1993). "Detection mechanisms in L-, M-, and S-cone contrast space." *Journal of the Optical Society of America. A, Optics and image science* 10.1, p. 38–51 (cf. p. 22, 81, 99).
- Conway, B. R. (2001). "Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1)". *Journal of Neuroscience* 21.8, p. 2768–2783 (cf. p. 24).
- Conway, B. R. (2009). "Color vision, cones, and color-coding in the cortex". *Neuroscientist* 15.3, p. 274–290 (cf. p. 7).
- Conway, B. R., Chatterjee, S. et al. (2010). "Advances in color science : from retina to behavior." *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30.45, p. 14955–63 (cf. p. 24).
- Conway, B. R., Hubel, D. H. et Livingstone, M. S. (2002). "Color contrast in macaque V1". *Cerebral cortex* 12.9, p. 915–925 (cf. p. 24).
- Conway, B. R. et Livingstone, M. S. (2006). "Spatial and temporal properties of cone signals in alert macaque primary visual cortex." *Journal of Neuroscience* 26.42, p. 10826–10846 (cf. p. 24).
- Cottaris, N. P. et De Valois, R. L. (1998). "Temporal dynamics of chromatic tuning in macaque primary visual cortex". *Nature* 395.6705, p. 896 (cf. p. 24).
- Curcio, C. A. et al. (1991). "Distribution and morphology of human cone photoreceptors stained with anti-blue opsin". *Journal of Comparative Neurology* 312.4, p. 610–624 (cf. p. 15).
- Dacey, D. M. (2000). "Parallel pathways for spectral coding in primate retina". *Annual review of neuroscience* 23.1, p. 743–775 (cf. p. 32).
- Dacey, D. M. et Lee, B. B. (1994). "The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type". *Nature* 367.6465, p. 731 (cf. p. 22).
- Dartnall, H. J. A., Bowmaker, J. K. et Mollon, J. D. (1983). "Human visual pigments : microspectrophotometric results from the eyes of seven persons". *Proc. R. Soc. Lond. B* 220.1218, p. 115–130 (cf. p. 16).
- Davidoff, J., Walsh, V. et Wagemans, J. (1997). "Higher-level cortical processing of colour". *Acta Psychologica* 97, p. 25–35 (cf. p. 6).
- De Monasterio, F. M., Gouras, P. et Tolhurst, D. J. (1975). "Concealed colour opponency in ganglion cells of the rhesus monkey retina." *The Journal of Physiology* 251.1, p. 217–229 (cf. p. 21).
- De Valois, R. L., Smith, C. J. et al. (1958). "Response of single cells in monkey lateral geniculate nucleus to monochromatic light." *Science* (cf. p. 21).
- De Valois, R. L. (1965). "Analysis and coding of color vision in the primate visual system". *Cold Spring Harbor Symposia on Quantitative Biology*. T. 30. Cold Spring Harbor Laboratory Press, p. 567 (cf. p. 10).

- De Valois, R. L., Abramov, I. et Jacobs, G. H. (1966). "Analysis of response patterns of LGN cells". *JOSA* 56.7, p. 966–977 (cf. p. 21).
- De Valois, R. L., Cottaris, N. P. et al. (2000). "Some transformations of color information from lateral geniculate nucleus to striate cortex". *Proceedings of the National Academy of Sciences* 97.9, p. 4997–5002 (cf. p. 24).
- De Valois, R. L., Jacobs, G. H. et Abramov, I. (1964). "Responses of single cells in visual system to shifts in the wavelength of light". *Science* 146.3648, p. 1184–1186 (cf. p. 21).
- DeCarlo, L. T. (1998). "Signal detection theory and generalized linear models." *Psychol. Methods* 3.2, p. 186–205 (cf. p. 120).
- Delorme, A., Richard, G. et Fabre-Thorpe, M. (2000). "Ultra-rapid categorisation of natural scenes does not rely on colour cues : a study in monkeys and humans". *Vision Research* 40, p. 2187–2200 (cf. p. 6).
- Derrington, A. M., Lennie, P. et Krauskopf, J. (1984). "Chromatic mechanisms in lateral geniculate nucleus of macaque". *The Journal of physiology* 357.1, p. 241–265 (cf. p. 21–23, 86, 116, 148).
- DeValois, R. L. et DeValois, K. K. (1990). *Spatial Vision*. Oxford, UK : Oxford University Press (cf. p. 13).
- DiMattina, C., Fox, S. A. et Lewicki, M. S. (2012). "Detecting natural occlusion boundaries using local cues". *Journal of Vision* 12.13, p. 15–15 (cf. p. 124, 125).
- Doi, E. et al. (2003). "Spatiochromatic receptive field properties derived from information-theoretic analyses of cone mosaic responses to natural scenes." *Neural computation* 15.2, p. 397–417 (cf. p. 43).
- Dominy, N. J. et Lucas, P. W. (2001). "Ecological importance of trichromatic vision to primates". *Nature* 410.6826, p. 363–366 (cf. p. 5).
- D’Zmura, M. et Lennie, P. (1986). "Mechanisms of color constancy". *Journal of the Optical Society of America A* 3.10, p. 1662–1672 (cf. p. 81).
- Elliot, A. J. (2015). "Color and psychological functioning : a review of theoretical and empirical work". *Frontiers in Psychology* 6, p. 368 (cf. p. 7).
- Entzmann, L. (2017). *Color completion in natural and artificial scenes*. Grenoble INP, Cognitive Science Department (cf. p. 146).
- Eskew, R. T. (2009). "Higher order color mechanisms : A critical review". *Vision Research* 49.22, p. 2686–2704 (cf. p. 32).
- Eskew, R. T. et Boynton, R. M. (1987). "Effects of field area and configuration on chromatic and border discriminations". *Vision Research* 27.10, p. 1835–1844 (cf. p. 23).
- Fairchild, M. D. (1992). "Chromatic adaptation and color constancy". *Advances in Color Vision Technical Digest* 4, p. 112–114 (cf. p. 26).
- Fairchild, M. D. (1993). "Chromatic adaptation in hard copy/soft copy comparisons". *Color Hard Copy and Graphic Arts II*. T. 1912. International Society for Optics et Photonics, p. 47–62 (cf. p. 26).
- Fairchild, M. D. (2013). *Color Appearance Models*, p. 471 (cf. p. 26).
- Felsen, G. et Dan, Y. (2005). "A natural approach to studying vision." *Nature neuroscience* 8.12, p. 1643–6 (cf. p. 81).
- Field, D. J. (1987). "Relations between the statistics of natural images and the response properties of cortical cells". *Josa a* 4.12, p. 2379–2394 (cf. p. 45).

- (1994). “What is the goal of sensory coding?” *Neural computation* 6.4, p. 559–601 (cf. p. 45).
- Field, G. D. et al. (2010). “Functional connectivity in the retina at the resolution of photoreceptors”. *Nature* 467.7316, p. 673–677 (cf. p. 22).
- Fine, I., Macleod, D. I. A. et Boynton, G. M. (2003). “Surface segmentation based on the luminance and color statistics of natural scenes” (cf. p. 5, 45, 71, 124).
- Finlayson, G. D., Hordley, S. D. et Drew, M. S. (2002). “Removing Shadows from Images”. *Computer Science* 2353.1, p. 823–836 (cf. p. 115).
- Fisher, R. A. (1936). “the Use of Multiple Measurements in Taxonomic Problems”. *Annals of Eugenics* 7.2, p. 179–188 (cf. p. 127).
- Foster, D. H. (2018). “The Verriest Lecture : Color vision in an uncertain world”. *Journal of the Optical Society of America A* 35.4, p. 192–201 (cf. p. 31).
- Foster, D. H., Nascimento, S. M. C. et Amano, K. (2004). “Information limits on neural identification of colored surfaces in natural scenes.” *Visual neuroscience* 21.3, p. 331–336 (cf. p. 71).
- Frey, H.-P., Honey, C. et König, P. (2008). “What’s color got to do with it? The influence of color on visual attention in different categories”. *Journal of Vision* 8.14, p. 6 (cf. p. 6, 136).
- Frey, H.-P., Wirz, K. et al. (2011). “Beyond Correlation : Do Color Features Influence Attention in Rainforest?” *Frontiers Human Neuroscience* 5.April, p. 1–13 (cf. p. 6, 56).
- Frome, F. S., Buck, S. L. et Boynton, R. M. (1981). “Visibility of borders : separate and combined effects of color differences, luminance contrast, and luminance level”. *JOSA* 71.2, p. 145–150 (cf. p. 23).
- Gegenfurtner, K. R. et Kiper, D. C. (1992). “Contrast detection in luminance and chromatic noise”. *J Opt. Soc. Am. A* 9.11, p. 1880–1888 (cf. p. 22).
- Gegenfurtner, K. R. (2003). “Cortical mechanisms of colour vision”. *Nature Reviews Neuroscience* (cf. p. 32).
- Gegenfurtner, K. R. et Kiper, D. C. (2003). “Color vision.” *Annual review of neuroscience* (cf. p. 7).
- Gegenfurtner, K. R. et Rieger, J. (2000). “Sensory and cognitive contributions of color to the recognition of natural scenes”. *Current Biology* 10, p. 805–808 (cf. p. 5, 6).
- Geisler, W. S. et al. (2001). “Edge co-occurrence in natural images predicts contour grouping performance”. *Vision Research* 41.6, p. 711–724 (cf. p. 37).
- Geisler, W. S. (2003). “Ideal Observer Analysis”. *The visual neurosciences* 10.7, p. 12–12 (cf. p. 89, 123).
- Geisler, W. S. (2007). “Visual Perception and the Statistical Properties of Natural Scenes” (cf. p. 35).
- Gibson, J. (1979). “The ecological approach to visual perception”. *Dallas : Houghton Mifflin* (cf. p. 34).
- Gibson, K. S. et Tyndall, E. P. T. (1923). *Visibility of Radiant Energy*. Scientific (cf. p. 16).
- Gilchrist, A. L. (1979). *The perception of Surface Black and Whites* (cf. p. 114).
- Gilchrist, A., Delman, S. et Jacobsen, A. (1983). “The classification and integration of edges as critical to the perception of reflectance and illumination”. *Perception & Psychophysics* 33.5, p. 425–436 (cf. p. 114).

- Goffaux, V. et al. (2005). “Diagnostic colours contribute to the early stages of scene categorization : Behavioural and neurophysiological evidence”. *Visual Cognition* 12.6, p. 878–892 (cf. p. 5).
- Goldsmith, T. H. (1991). “The evolution of visual pigments and colour vision”. *Evolution of the Eye and Visual System*. Sous la dir. de J. C.-D. MacMillan. London (cf. p. 13).
- Gouras, P. (1968). “Identification of cone mechanisms in monkey ganglion cells”. *The Journal of physiology* 199.3, p. 533–547 (cf. p. 21).
- Grassmann, H. (1854). “On the theory of compound colors”. *Phil. Mag* 7, p. 254–264 (cf. p. 17).
- Green, D. M. et Swets, J. A. (1988). *Signal Detection Theory and Psychophysics* (cf. p. 89, 123).
- Grossberg, S. (1987). “Cortical dynamics of three-dimensional form, color, and brightness perception : I. Monocular theory”. *Perception & psychophysics* 41.2, p. 87–116 (cf. p. 141).
- (2000). “The complementary brain : Unifying brain dynamics and modularity”. *Trends in cognitive sciences* 4.6, p. 233–246 (cf. p. 141).
- Gunther, K. L. et Dobkins, K. R. (2002). “Individual differences in chromatic (red/green) contrast sensitivity are constrained by the relative number of L- versus M-cones in the eye”. *Vision Research* 42.11, p. 1367–1378 (cf. p. 23).
- Guth, S. L., Massof, R. W. et Benzscharwel, T. (1980). “Vector model for normal and dichromatic color vision”. *JOSA* 70.2, p. 197–212 (cf. p. 23).
- Hall, J. et Guyton, A. (2011). “Textbook of Medical Physiology”. *Textbook of Medical Physiology*. Sous la dir. d’Elsevier Publications. 11th ed. Chap. Chapitre 5, p. 619 (cf. p. 13).
- Hanazawa, A., Komatsu, H. et Murakami, I. (2000). “Neural selectivity for hue and saturation of colour in the primary visual cortex of the monkey”. *European Journal of Neuroscience* 12.5, p. 1753–1763 (cf. p. 24).
- Hansen, T. et Gegenfurtner, K. R. (2013). “Higher order color mechanisms : Evidence from noise-masking experiments in cone contrast space”. *Journal of Vision* 13.1, p. 26.1–21 (cf. p. 32).
- Hansen, T. et Gegenfurtner, K. R. (2009). “Independence of color and luminance edges in natural scenes”. *Visual neuroscience* 26.1, p. 35–49 (cf. p. 31, 42, 45, 48, 53, 55, 58, 71, 83, 134, 135).
- (2017). “Color contributes to object-contour perception in natural scenes”. 17, p. 1–19 (cf. p. 31, 115).
- Hansen, T., Olkkonen, M. et al. (2006). “Memory modulates color appearance”. *Nature neuroscience* 9.11, p. 1367 (cf. p. 141).
- Heidemann, G. (2006). “The principal components of natural images revisited”. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 5, p. 822–826 (cf. p. 71).
- Hering, E. (1878). “Outlines of a theory of the light sense.” (cf. p. 10).
- Hofer, H. (2005). “Organization of the Human Trichromatic Cone Mosaic”. *Journal of Neuroscience* 25.42, p. 9669–9679 (cf. p. 15).
- Horwitz, G. D., Chichilnisky, E. J. et Albright, T. D. (2005). “Blue-yellow signals are enhanced by spatiotemporal luminance contrast in macaque V1”. *Journal of Neurophysiology* 93.4, p. 2263–2278 (cf. p. 24).
- (2007). “Cone inputs to simple and complex cells in V1 of awake macaque”. *Journal of neurophysiology* 97.4, p. 3070–3081 (cf. p. 24).

- Hubel, D. H. et Wiesel, T. N. (1968a). "Receptive Fields and Functional Architecture of monkey striate cortex". *Journal of Physiology* 195, p. 215–243 (cf. p. 25).
- Hubel, D. H. et Wiesel, T. N. (1959). "Receptive fields of single neurones in the cat's striate cortex". *The Journal of physiology* 148.3, p. 574–591 (cf. p. 25).
- (1962). "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex". *The Journal of physiology* 160.1, p. 106–154 (cf. p. 25).
- (1968b). "Receptive fields and functional architecture of monkey striate cortex". *The Journal of physiology* 195.1, p. 215–243 (cf. p. 44).
- Hurvich, L. M. et Jameson, D. (1957). "An opponent-process theory of color vision." *Psychological review* 64.6p1, p. 384 (cf. p. 10).
- Ing, A. D. (2010). "Region grouping in natural foliage scenes : Image statistics and human performance". *Journal of Vision* 10.4, p. 1–19 (cf. p. 124).
- Ingling, C. R., Barley, J. P. et Ghani, N. (1996). "Chromatic content of spectral lights". *Vision Research* 36.16, p. 2537–2551 (cf. p. 81).
- Ingling, C. R. et Tsou, B. H. (1988). "Spectral sensitivity for flicker and acuity criteria." *Journal of the Optical Society of America A* 5.8, p. 1374–1378 (cf. p. 54).
- Ingling, C. R. et Martinez-Uriegas, E. (1983). "The relationship between spectral sensitivity and spatial sensitivity for the primate rg X-channel". *Vision research* 23.12, p. 1495–1500 (cf. p. 23).
- Itti, L., Koch, C. et Niebur, E. (1998). "A model of saliency-based visual attention for rapid scene analysis". *IEEE Transactions on pattern analysis and machine intelligence* 20.11, p. 1254–1259 (cf. p. 136).
- Jacobs, G. H. (1993). "The distribution and nature of colour vision among the mammals". *Biological Reviews* 68.3, p. 413–471 (cf. p. 10).
- Jennings (2015). "Detection of chromatic and luminance distortions in natural scenes". *Journal of the Optical Society of America* 32.9, p. 1613–1622 (cf. p. 31).
- Jennings, B. J. et Kingdom, F. A. (2017). "Chromatic blur perception in the presence of luminance contrast". *Vision Research* 135, p. 34–42 (cf. p. 31).
- Jindrová, H. (1998). *Vertebrate phototransduction : Activation, recovery, and adaptation* (cf. p. 13).
- Johnson, A. P., Kingdom, F. a. a. et Baker, C. L. (2005). "Spatiochromatic statistics of natural scenes : first- and second-order information and their correlational structure." *Journal of the Optical Society of America A* 22.10, p. 2050–2059 (cf. p. 44, 48, 71, 83).
- Johnson, E. N., Hawken, M. J. et Shapley, R. (2004). "Cone inputs in macaque primary visual cortex." *Journal of neurophysiology* 91.6, p. 2501–14 (cf. p. 24).
- (2008). "The orientation selectivity of color-responsive neurons in macaque V1". *Journal of Neuroscience* 28.32, p. 8096–8106 (cf. p. 25, 27).
- Jones, M. et Love, B. C. (2011). "Bayesian fundamentalism or enlightenment ? On the explanatory status and theoretical contributions of Bayesian models of cognition". *Behavioral and Brain Sciences* 34.4, p. 169–188 (cf. p. 37).
- Jost, T. et al. (2005). "Assessing the contribution of color in visual attention". *Computer Vision and Image Understanding* 100.1-2 SPEC. ISS. P. 107–123 (cf. p. 6).

- Judd, D. B. (1951). "Report of U.S. Secretariat Committee on Colorimetry and Artificial Daylight". *Proceedings of the Twelfth Session of the CIE, Stockholm (vol. 1, pp. 11)*. Paris : Bureau Central de la CIE. (Cf. p. 16, 53, 170).
- Kersten, D. et al. (2004). "Object perception as Bayesian inference". *Annual review of psychology*. 55, p. 1–34 (cf. p. 37).
- Kingdom, F. A. A., Beauce, C. et Hunter, L. (2004). "Colour vision brings clarity to shadows". *Perception* 33.8, p. 907–914 (cf. p. 5, 114, 115).
- Kingdom, F. A. A. (2008). "Perceiving light versus material". *Vision Research* 48.20, p. 2090–2105 (cf. p. 114).
- Kingdom, F. A. (2003). "Color brings relief to human vision". *Nature Neuroscience* 6.6, p. 641–644 (cf. p. 115).
- Kingdom, F. et Prins, N. (2016). *Psychophysics : A practical introduction second edition*. Academ Press (cf. p. 93).
- Kiper, D. C., Fenstemaker, S. B. et Gegenfurtner, K. R. (1997). "Chromatic properties of neurons in macaque area V2". *Visual neuroscience* 14.6, p. 1061–1072 (cf. p. 24).
- Knight, R. et al. (1998). "Rods affect S-cone discrimination on the Farnsworth–Munsell 100-hue test". *Vision research* 38.21, p. 3477–3481 (cf. p. 13).
- Knill, D. C. et Richards, W. (1996). *Perception as Bayesian inference*. Cambridge University Press (cf. p. 37).
- Knoblauch, K. (1995). "Dual bases in dichromatic color space". *Colour Vision Deficiencies XII*. Springer, p. 165–176 (cf. p. 169).
- Knoblauch, K. et Shevell, S. K. (2001). "Relating cone signals to color appearance : Failure of monotonicity in yellow/blue". *Visual Neuroscience* 18.06, p. 901–906 (cf. p. 81).
- Koffka, K. (1935). *Principles of Gestalt Psychology, International Library of Psychology, Philosophy and Scientific Method* (cf. p. 5).
- König, A. et Dieterich, C. (1892). *Die Grundempfindungen in normalen und anomalen Farbensystemen und ihre Intensitätsverteilung im Spektrum*. L. Voss (cf. p. 20).
- Kraskov, A., Stögbauer, H. et Grassberger, P. (2004). "Estimating mutual information". *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 69.6 2, p. 1–16 (cf. p. 51).
- Krauskopf, J., Gegenfurtner, K. R. et Sharpe, L. T. (1999). *Higher order color mechanisms*. Cambridge U. Press (cf. p. 22).
- Krauskopf, J., Williams, D. R. et Heeley, D. W. (1982). "Cardinal directions of color space". *Vision Research* 22.9, p. 1123–1131 (cf. p. 21–23).
- Kremers, J., Lee, B. B. et Kaiser, P. K. (1992). "Sensitivity of macaque retinal ganglion cells and human observers to combined luminance and chromatic temporal modulation". *JOSA A* 9.9, p. 1477–1485 (cf. p. 22).
- Kremers, J., Silveira, L. C. L. et al. (2016). "The Retinal Processing of Photoreceptor Signals". *Human Color Vision*. Springer, p. 33–70 (cf. p. 21).
- Kries, J. von (1902). "Chromatic Adaptation [Translation : D. MacAdam, "Sources of Color Science", MIT press, 1970.]" (cf. p. 54).
- Land, E. H. et McCann, J. J. (1971). "Lightness and Retinex Theory". *Journal of the Optical Society of America* 61.1, p. 1 (cf. p. 81).
- Larimer, J., Krantz, D. H. et Cicerone, C. M. (1975). "Opponent process additivity-II. Yellow/blue equilibria and nonlinear models". *Vision Research* 15.6, p. 723–731 (cf. p. 81).

- Lee, B. B., Martin, P. R. et Valberg, A. (1988). “The physiological basis of heterochromatic flicker photometry demonstrated in the ganglion cells of the macaque retina.” *The Journal of physiology* 404.1, p. 323–347 (cf. p. 16, 21, 22).
- Lee, B. B., Valberg, A. et al. (1987). “An account of responses of spectrally opponent neurons in macaque lateral geniculate nucleus to successive contrast”. *Proc. R. Soc. Lond. B* 230.1260, p. 293–314 (cf. p. 21).
- Lee, B. B. (2004). “Paths to colour in the retina”. *Clinical and Experimental Optometry* 87.4-5, p. 239–248 (cf. p. 21).
- Lee, B. B. (2011). “Visual pathways and psychophysical channels in the primate.” *The Journal of physiology* 589.Pt 1, p. 41–7 (cf. p. 22).
- (2014). “Color coding in the primate visual pathway : a historical view”. *Journal of the Optical Society of America A* 31.4, A103–A112 (cf. p. 21).
- Lee, B. B., Martin, P. R., Valberg, A. et Kremers, J. (1993). “Physiological mechanisms underlying psychophysical sensitivity to combined luminance and chromatic modulation”. *JOSA A* 10.6, p. 1403–1412 (cf. p. 22).
- Lee, T.-w., Wachtler, T. et Sejnowski, T. J. (2002). “Color opponency is an efficient representation of spectral properties in natural scenes”. 42, p. 2095–2103 (cf. p. 43).
- Lennie, P., Pokorny, J. et Smith, V. C. (1993). “Luminance”. *J Opt Soc Am A* 10.6, p. 1283–1293 (cf. p. 16).
- Lennie, P. et D’Zmura, M. (1988). “Mechanisms of color vision.” *Critical reviews in neurobiology* 3.4, p. 333–400 (cf. p. 32).
- Lennie, P., Krauskopf, J. et Sclar, G. (1990). “Chromatic mechanisms in striate cortex of macaque”. *Journal of Neuroscience* 10.2, p. 649–669 (cf. p. 24).
- Lennie, P. et Movshon, J. A. (2005). “Coding of color and form in the geniculostriate visual pathway (invited review).” *Journal of the Optical Society of America. A, Optics, image science, and vision* 22.10, p. 2013–33 (cf. p. 7).
- Livingstone, M. S. et Hubel, D. H. (1987). “Psychophysical evidence for separate channels for the perception of form, color, movement, and depth”. *Journal of Neuroscience* 7.11, p. 3416–3468 (cf. p. 30).
- Lovell, P. G. et al. (2006). “Evaluation of a multiscale color model for visual difference prediction”. *ACM Transactions on Applied Perception* 3.3, p. 155–178 (cf. p. 33).
- Lucas, P. W. et al. (1998). “Colour cues for leaf food selection by long-tailed macaques (*Macaca fascicularis*) with a new suggestion for the evolution of trichromatic colour vision”. *Folia Primatologica* 69.3, p. 139–154 (cf. p. 5).
- Mamassian, Y., Landy, M. et Maloney, L. T. (2002). “Bayesian Modelling of Visual Perception”. *Probabilistic Models of the Brain : Perception and Neural Function*, p. 13–36 (cf. p. 37).
- Marr, D. (1982). “Vision : A computational investigation into the human representation and processing of visual information. MIT Press”. *Cambridge, Massachusetts* (cf. p. 32, 45).
- Martin, P. R. et al. (2001). “Chromatic sensitivity of ganglion cells in the peripheral primate retina”. *Nature* 410.6831, p. 933 (cf. p. 22).
- Maxwell, J. C. (1855). *On the Theory of Colours in Relation to Colour-blindness...* (Cf. p. 10).
- Mollon, J. D. (1989). “‘Tho’ she kneel’d in that place where they grew...’ — the uses and origins of primate colour vision.” *J. exp. Biol* 146.2, p. 21–38 (cf. p. 5).

- Mollon, J. D. et Bowmaker, J. K. (1992). "The spatial arrangement of cones in the primate fovea". *Nature* 360.6405, p. 677 (cf. p. 15).
- Moutoussis, K. (2015). "The physiology and psychophysics of the color-form relationship : A review". *Frontiers in Psychology* 6.NOV, p. 1–17 (cf. p. 31).
- Mullen, K. T. et Kingdom, F. A. A. (2002). "Differential distributions of red-green and blue-yellow cone opponency across the visual field". *Visual Neuroscience* 19.1, p. 109–118 (cf. p. 56).
- Mullen, K. T. et Losada, M. a. (1994). "Evidence for separate pathways for color and luminance detection mechanisms." *Journal of the Optical Society of America. A, Optics, image science, and vision* 11.12, p. 3136–3151 (cf. p. 22, 81).
- Mullen, K. T. (1985). "The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings." *Journal of Physiology* 359.1, p. 381–400 (cf. p. 26, 44).
- Mullen, K. T. et Sankeralli, M. J. (1999). "Evidence for the stochastic independence of the blue-yellow, red-green and luminance detection mechanisms revealed by subthreshold summation". *Vision Research* 39.4, p. 733–745 (cf. p. 22).
- Nathans, J. (1999). "The evolution and physiology of human color vision : insights from molecular genetic studies of visual pigments". *Neuron* 24.2, p. 299–312 (cf. p. 5).
- Neumann, H., Pessoa, L. et Hansen, T. (2001). "Visual filling-in for computing perceptual surface properties". *Biological Cybernetics* 85.5, p. 355–369 (cf. p. 141).
- Normann, R. A. et Werblin, F. S. (1974). "Control of retinal sensitivity : I. Light and dark adaptation of vertebrate rods and cones". *The Journal of general physiology* 63.1, p. 37–61 (cf. p. 13).
- Oliva, A. et Schyns, P. G. (2000). "Diagnostic Colors Mediate Scene Recognition". *Cognitive Psychology* 41, p. 176–210 (cf. p. 5).
- Olkkonen, M., Hansen, T. et Gegenfurtner, K. R. (2008). "Color appearance of familiar objects : Effects of object shape, texture, and illumination changes". *Journal of vision* 8.5, p. 13–13 (cf. p. 141).
- Olmos, A. et Kingdom, F. (2004a). *McGill Calibrated Colour Image Database* (cf. p. 46, 81, 86, 116, 125).
- Olmos, A. et Kingdom, F. A. A. (2004b). "A biologically inspired algorithm for the recovery of shading and reflectance images". *Perception* 33.12, p. 1463–1473 (cf. p. 115).
- Olshausen, B. a. et Field, D. J. (2005). "How close are we to understanding v1?" *Neural computation* 17.8, p. 1665–99 (cf. p. 81).
- Párraga, C. a., Brelstaff, G. et al. (1998). *Color and luminance information in natural scenes*. (Cf. p. 33, 44, 48, 54, 82, 83, 94, 125).
- Párraga, C., Troscianko, T. et Tolhurst, D. (2002). "Spatiochromatic Properties of Natural Images and Human Vision". *Current Biology* 12.6, p. 483–487 (cf. p. 44, 48, 83).
- Peyvandi, S., Nieves, J. L. et Gilchrist, A. (2013). "On the Information Content along Edges in Trichromatic Images". 3, p. 236–239 (cf. p. 45, 71).
- Ho-Phuoc, T. et al. (2012). "When viewing natural scenes, do abnormal colors impact on spatial or temporal parameters of eye movements?" *Journal of Vision* 12.2, p. 4 (cf. p. 6, 136).
- Pinna, B. (2008). "Watercolor illusion". *Scholarpedia* 3.1. revision #127995, p. 5352 (cf. p. 141, 142).

- Provenzi, E. et al. (2016). “On the second order spatiochromatic structure of natural images”. *Vision Research* 120, p. 22–38 (cf. p. 31).
- Regan, B. C. et al. (2001). “Fruits, foliage and the evolution of primate colour vision”. *Philosophical Transactions of the Royal Society of London B : Biological Sciences* 356.1407, p. 229–283 (cf. p. 5).
- Reid, R. C. et Shapley, R. M. (1992). “Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus”. *Nature* 356.6371, p. 716 (cf. p. 22).
- (2002). “Space and time maps of cone photoreceptor signals in macaque lateral geniculate nucleus”. *Journal of Neuroscience* 22.14, p. 6158–6175 (cf. p. 22).
- Rentzeperis, I. et al. (2014). “Distributed processing of color and form in the visual cortex”. *Frontiers in Psychology* 5.AUG, p. 1–14 (cf. p. 24).
- Rieke, F. et Rudd, M. E. (2009). “The Challenges Natural Images Pose for Visual Adaptation”. *Neuron* 64.5, p. 605–616 (cf. p. 82).
- Rizzolatti, G. et al. (1987). “Reorienting attention across the horizontal and vertical meridians : evidence in favor of a premotor theory of attention”. *Neuropsychologia* 25.1, p. 31–40 (cf. p. 136).
- Roorda, A. et Williams, D. R. (1999). “The arrangement of the three cone classes in the living human eye”. *Nature* 397.6719, p. 520 (cf. p. 15).
- Rowe, M. H. (2002). “Trichromatic Color Vision in Primates”. *Physiology* 17.3, p. 93–98 (cf. p. 32).
- Rubin, J. M. et Richards, W. A. (1982). “Color Vision and Image Intensities : When are Changes Material?” 226, p. 215–226 (cf. p. 114).
- Ruderman, D. L., Cronin, T. W. et Chiao, C.-C. (1998). “Statistics of cone responses to natural images : implications for visual coding”. *Journal of the Optical Society of America A* 15.8, p. 2036 (cf. p. 42, 43, 54, 55, 71, 81).
- Rushton, W. A. H. (1972). “Visual pigments in man”. *Photochemistry of vision*. Springer, p. 364–394 (cf. p. 13).
- Russell, R. et Sinha, P. (2007). “Real-world face recognition : The importance of surface reflectance properties”. *Perception* 36.9, p. 1368–1374 (cf. p. 5).
- Rust, N. C. et Movshon, J. A. (2005). “In praise of artifice.” *Nature neuroscience* 8.12, p. 1647–1650 (cf. p. 81).
- Sankeralli, M. J. et Mullen, K. T. (1997). “Postreceptoral chromatic detection mechanisms revealed by noise masking in three-dimensional cone contrast space”. *Journal of the Optical Society of America A* 14.10, p. 2633–2646 (cf. p. 22).
- Sankeralli, M. J. et Mullen, K. T. (1996). “Estimation of the L-, M-, and S-cone weights of the postreceptoral detection mechanisms”. *Journal of the Optical Society of America A* 13.5, p. 906 (cf. p. 99).
- Schluppeck, D. et Engel, S. a. (2002). “Color opponent neurons in V1 : a review and model reconciling results from imaging and single-unit recording.” *Journal of vision* 2.6, p. 480–92 (cf. p. 24).
- Schnapf, J. L., Kraft, T. W. et Baylor, D. A. (1987). “Spectral sensitivity of human cone photoreceptors”. *Nature* 325.6103, p. 439 (cf. p. 15, 16).
- Schrater, P. et Kersten, D. (2002). “Vision, Psychophysics and Bayes”. *Probabilistic Models of the Brain : Perception and Neural Function*, p. 37–60 (cf. p. 37).

- Schwartz, O. et Simoncelli, E. P. (2001). “Natural signal statistics and sensory gain control”. *Nature Neuroscience* 4.8, p. 819–825 (cf. p. 82).
- Shannon, C. E. et Weaver, W. (1949). “The mathematical theory of communication”. *Urbana, IL : University of Illinois Press* (cf. p. 34).
- Shapley, R. et Hawken, M. J. (2011). “Color in the Cortex : single- and double-opponent cells”. *Vision Research* 51.7, p. 701–717 (cf. p. 24, 31).
- Sharpe, L. T. et al. (2005). “A luminous efficiency function, $V^*(\lambda)$, for daylight adaptation”. *Journal of Vision* 5.11, p. 3 (cf. p. 16, 99).
- (2011). “A luminous efficiency function, $VD65^*(\lambda)$, for daylight adaptation : A correction”. *Color Research and Application* 36.1, p. 42–46 (cf. p. 53).
- Shevell, S. K. et Kingdom, F. A. A. (2008). “Color in Complex Scenes”. *Annu. Rev. Psychol* 59, p. 143–66 (cf. p. 31).
- Silveira, L. C. L. et al. (2005). “Comparative anatomy and physiology of the primate retina”. *The primate visual system : A comparative approach*, p. 127–160 (cf. p. 21).
- Simoncelli, E. P. et Olshausen, B. A. (2001). “Natural Image Statistics and Neural Representation”. *Annual Review of Neuroscience* 24.1, p. 1193–1216 (cf. p. 34, 42).
- Simpson, E. H. (1951). “The interpretation of interaction in contingency tables”. *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 238–241 (cf. p. 58).
- Smith, V. C. et Pokorny, J. (1975). “Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm”. *Vision Research* 15.2, p. 161–171 (cf. p. 20, 48, 53, 65, 86, 116, 170).
- Solomon, S. G. (2005). “Chromatic Gain Controls in Visual Cortical Neurons”. *Journal of Neuroscience* 25.19, p. 4779–4792 (cf. p. 24, 82).
- Solomon, S. G. et Lennie, P. (2007). “The machinery of colour vision.” *Nature reviews. Neuroscience* 8.4, p. 276–286 (cf. p. 7, 11).
- Solomon, S. G., Peirce, J. W. et Lennie, P. (2004). “The impact of suppressive surrounds on chromatic properties of cortical neurons”. *Journal of Neuroscience* 24.1, p. 148–160 (cf. p. 24).
- Spence, I. et al. (2006). “How color enhances visual memory for natural scenes?”. *Psychological Science* 17.1, p. 1–6 (cf. p. 6).
- Stabell, B. et Stabell, U. (2002). “Effects of rod activity on color perception with light adaptation”. *JOSA A* 19.7, p. 1249–1258 (cf. p. 13).
- Stabell, U., Definitions, A. et Apparatus, B. (1998). “Chromatic rod – cone interaction during”. *America* 15.11, p. 2809–2815 (cf. p. 13).
- Standring, S. et Gray, H. (2008). “Gray’s anatomy : the anatomical basis of clinical practice”. Churchill Livingstone. Chap. Chapitre 4, p. 690 (cf. p. 13).
- Stiles, W. S. et Burch, J. M. (1955). “Interim report to the Commission Internationale de l’Eclairage, Zurich, 1955, on the National Physical Laboratory’s investigation of colour-matching (1955)”. *Optica Acta : International Journal of Optics* 2.4, p. 168–181 (cf. p. 17).
- Stockman, A. et Sharpe, L. T. (1999). “Cone spectral sensitivities and color matching”. *Color vision : From Genes to Perception*. 1855. Chap. Chapitre 2, p. 53–87 (cf. p. 20).
- Stockman, A. et Brainard, D. H. (2010). “Color vision mechanisms”. *The Optical Society of America Handbook of Optics*, p. 11.1–11.104 (cf. p. 32).

- Stockman, A. et Sharpe, L. T. (2000). “The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype”. *Vision Research* 40.13, p. 1711–1737 (cf. p. 20, 53, 54, 65).
- Stromeyer, C. F., Cole, G. R. et Kronauer, R. E. (1985). “Second-site adaptation in the red-green chromatic pathways”. *Vision Research* 25.2, p. 219–237 (cf. p. 23, 99).
- Stromeyer, C. F., Kronauer, R. E. et Cole, G. R. (1983). “Adaptive mechanisms controlling sensitivity to red-green chromatic flashes”. *Colour vision : Physiology and psychophysics*, p. 313–330 (cf. p. 23).
- Sumner, P. et Mollon, J. D. (2000a). “Chromaticity as a signal of ripeness in fruits taken by primates”. *Journal of Experimental Biology* 203.13, p. 1987–2000 (cf. p. 5).
- Sumner, P. et Mollon, J. D. (2000b). “Catarrhine photopigments are optimized for detecting targets against a foliage background”. *Journal of Experimental Biology* 203.13, p. 1963–1986 (cf. p. 5).
- Sun, H. et al. (2006). “Specificity of cone inputs to macaque retinal ganglion cells”. *Journal of neurophysiology* 95.2, p. 837–849 (cf. p. 22).
- Switkes, E., Bradley, a. et De Valois, K. K. (1988). “Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings.” *Journal of the Optical Society of America. A, Optics and image science* 5.7, p. 1149–62 (cf. p. 81).
- Tadmor, Y. et Tolhurst, D. J. (2000). “Calculating the contrasts that retinal ganglion cells and LGN neurones encounter in natural scenes”. *Vision research* 40.22, p. 3145–3157 (cf. p. 31).
- Tailby, C., Solomon, S. G. et Lennie, P. (2008). “Functional Asymmetries in Visual Pathways Carrying S-Cone Signals in Macaque”. *Journal of Neuroscience* 28.15, p. 4078–4087 (cf. p. 22).
- Tanaka, J. W. et Presnell, L. M. (1999). “Color diagnosticity in object recognition”. *Perception & Psychophysics* 61.6, p. 1140–1153 (cf. p. 5).
- Tanaka, J., Weiskopf, D. et Williams, P. (2001). “The role of color in high-level vision”. *Trends in cognitive sciences* 5.5, p. 211–215 (cf. p. 6).
- Tansley, B. W. et Boynton, R. M. (1976). “A line, not a space, represents visual distinctness of borders formed by different colors”. *Science* 191.4230, p. 954–957 (cf. p. 23).
- Tappen, M., Freeman, W. et Adelson, E. (2003). “Recovering intrinsic images from a single image”. *Advances in Neural Information Processing Systems* 15 (cf. p. 115).
- To, M. P. S. et al. (2010). “Perception of suprathreshold naturalistic changes in colored natural images.” *Journal of Vision* 10.4, 12 :1–22 (cf. p. 33).
- Tolhurst, D. J., Tadmor, Y. et Chao, T. (1992). “Amplitude spectra of natural images”. *Ophthalmic and Physiological Optics* 12.2, p. 229–232 (cf. p. 45).
- Valberg, A. (2005). *Light Vision Color*, p. 181–202 (cf. p. 15).
- Van Den Berg, R. et al. (2012). “Optimal inference of sameness”. *Proceedings of the National Academy of Sciences* 109.8, p. 3178–3183 (cf. p. 37).
- Viénot, F. (2016). “Cone fundamentals : A model for the future of colorimetry”. *Lighting Research and Technology* 48, p. 5–13 (cf. p. 20).
- Vilankar, K. et Golden, J. (2014). “Local edge statistics provide information regarding occlusion and nonocclusion edges in natural scenes”. *Journal of vision* 14.2014, p. 1–21 (cf. p. 124).

- Vincent, B. T. (2015). "A tutorial on Bayesian models of perception". *Journal of Mathematical Psychology* 66, p. 103–114 (cf. p. 37).
- Von Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. T. 9. Voss (cf. p. 10, 20, 35).
- Von Wartburg, R. et al. (2005). "The influence of colour on oculomotor behaviour during image perception". *NeuroReport* 16.14, p. 1557–1560 (cf. p. 6).
- Vorobyev, M. et Osorio, D. (1998). "Receptor noise as a determinant of colour thresholds." *Proceedings. Biological sciences / The Royal Society* 265.1394, p. 351–8 (cf. p. 111).
- Vos, J. J. (1978). "Colorimetric and photometric properties of a 2-deg fundamental observer." *Color Research and Application*, 3, 125–128 (cf. p. 16, 53, 170).
- Wachtler, T., Lee, T. W. et Sejnowski, T. J. (2001). "Chromatic structure of natural scenes." *Journal of the Optical Society of America. A, Optics, image science, and vision* 18.1, p. 65–77 (cf. p. 43, 71).
- Wachtler, T., Sejnowski, T. J. et Albright, T. D. (2003). "Representation of color stimuli in awake macaque primary visual cortex". *Neuron* 37.4, p. 681–691 (cf. p. 24).
- Walters-Williams, J. et Li, Y. (2009). "Estimation of mutual information : A survey". *International Conference on Rough Sets and Knowledge Technology*. Springer, p. 389–396 (cf. p. 51).
- Wandell, B. A. (1995). *Foundations of vision*. T. 8. Sinauer Associates Sunderland, MA (cf. p. 13, 15, 17, 26).
- Wandell, B. A. et Winawer, J. (2011). "Imaging retinotopic maps in the human brain". *Vision research* 51.7, p. 718–737 (cf. p. 24).
- Webster, M. a., Miyahara, E. et al. (2000). "Variations in normal color vision. I- Cone-opponent axes". *Journal of the Optical Society of America. A, Optics, image science, and vision* 17.9, p. 1545–1555 (cf. p. 99).
- Webster, M. A. (2015). "Visual Adaptation". *Annual Review of Vision Science* 1.1, p. 547–567 (cf. p. 26).
- Webster, M. A. et Mollon, J. D. (1997). "Adaptation and the Color Statistics of Natural Images". *Vision Res* 37.23, p. 3283–3298 (cf. p. 42).
- Wei, X.-X. et Stocker, A. A. (2015). "A Bayesian observer model constrained by efficient coding can explain 'anti-Bayesian' percepts". *Nature Neuroscience* 18.10, p. 1509–1517 (cf. p. 34).
- Werner, J. S. et Wooten, B. R. (1979). "Opponent chromatic mechanisms : Relation to photopigments and hue naming". *Journal of the Optical Society of America* 69.3, p. 422–434 (cf. p. 81).
- Wichmann, F. A., Sharpe, L. T. et Gegenfurtner, K. R. (2002). "The contributions of color to recognition memory for natural scenes." *Journal of Experimental Psychology : Learning, Memory, and Cognition* 28.3, p. 509 (cf. p. 6).
- Wiesel, T. N. et Hubel, D. H. (1966). "Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey." *Journal of neurophysiology* 29.6, p. 1115–1156 (cf. p. 21).
- Williams, D. R., MacLeod, D. I. A. et Hayhoe, M. M. (1981). "Foveal tritanopia". *Vision research* 21.9, p. 1341–1356 (cf. p. 15).
- Witzel, C. et Gegenfurtner, K. R. (2018). "Color Perception : Objects, Constancy, and Categories". *Annual review of vision science* 4, p. 475–499 (cf. p. 31).

- Wurm, L. H. et al. (1993). "Color improves object recognition in normal and low vision." *Journal of Experimental Psychology : Human perception and performance* 19.4, p. 899 (cf. p. 5).
- Wyszecki, G. et Stiles, W. S. (1982). *Color science*. T. 8. Wiley New York (cf. p. 16).
- Yarbus, A. L. (1967). "Eye movements during perception of complex objects". *Eye movements and vision*. Springer, p. 171–211 (cf. p. 136).
- Yoonessi, A., Kingdom, F. a. a. et Alqawlaq, S. (2008). "Is color patchy?" *Journal of the Optical Society of America. A, Optics, image science, and vision* 25.6, p. 1330–1338 (cf. p. 31).
- Yoshioka, T., Dow, B. M. et Vautin, R. G. (1996). "Neuronal mechanisms of color categorization in areas V1, V2 and V4 of macaque monkey visual cortex". *Behavioural brain research* 76.1-2, p. 51–70 (cf. p. 24).
- Young, T. (1807). *A course of lectures on natural philosophy and the mechanical arts : in two volumes*. T. 2. Johnson (cf. p. 20).
- Zaidi, Q. et Halevy, D. (1993). "Visual mechanisms that signal the direction of color changes". *Vision research* 33.8, p. 1037–1051 (cf. p. 169).
- Zeki, S. M. (1978). "Functional specialisation in the visual cortex of the rhesus monkey". *Nature* 274.5670, p. 423 (cf. p. 31).
- Zhou, C. et Mel, B. W. (2008). "Cue combination and color edge detection in natural scenes." *Journal of vision* 8 VN - re.4, p. 4–25 (cf. p. 45, 71).



Conversion de l'espace des couleurs RGB vers l'espace DKL

Afin de définir un espace de couleur DKL, nous devons spécifier les trois axes de celui-ci. Deux approches permettent de définir ces axes, et sont formellement équivalentes (Knoblauch, 1995) :

- L'une basée sur les directions cardinales, c'est-à-dire en termes de stimuli qui isolent chacun des mécanismes post-récepteurs (Zaidi et Halevy, 1993).
- L'autre basée sur les mécanismes post-récepteurs, en termes de propriétés des réponses des mécanismes (Brainard, 1996).

Par définition, dans l'espace DKL, ces axes sont identiques : les directions cardinales sont celles qui correspondent aux réponses des trois canaux post-récepteurs et qui permettent également d'isoler celles-ci.

Pour spécifier les directions cardinales de notre espace DKL, nous utilisons une procédure relative à l'écran de présentation des stimuli et basée sur la caractéristique d'isolation de réponse. La procédure se décompose en trois grandes étapes :

- Dans un premier temps, nous mesurons les primaires RGB de l'écran pour une gamme de 18 niveaux d'intensité. C'est-à-dire que pour 18 valeurs de $k \in [0, 255]$, nous mesurons la luminance associée à l'affichage de pixels (R, G, B) de valeurs $(k, 0, 0)$, $(0, k, 0)$ et $(0, 0, k)$. Ces mesures sont réalisées grâce à un spectroradiomètre (CS 2000, Konica-Minolta, Inc), dans des conditions comparables à celles de présentation des stimuli dans le cadre de nos expériences : en affichant des carrés colorés sur un fond gris neutre de coordonnées RGB=(127,127,127). Les courbes de luminance mesurées ne sont pas linéaires et peuvent être modélisées par des fonctions $f(x) = x^\gamma$, avec des valeurs de

γ différentes pour chacune des primaires R, G et B. Nous estimons ainsi les valeurs $(\gamma_R, \gamma_G, \gamma_B)$ permettant de convertir les valeurs RGB de notre écran dans l'espace sRGB, qui est l'espace RGB standard linéaire.

- Nous calculons ensuite les réponses respectives de chacun des types de cônes L, M et S. Une fois les coordonnées des pixels converties dans l'espace sRGB, les valeurs LMS peuvent être obtenues par une transformation linéaire se décomposant en deux étapes. Les valeurs sRGB sont, dans un premier temps, projetées dans l'espace tristimulus XYZ, puis dans l'espace LMS, basé sur les fonctions fondamentales de Smith et Pokorny (1975). Ces deux transformations correspondent formellement à deux matrices de passage, dont nous noterons le produit \mathbf{T} . Nous ajoutons aux résultats obtenus suite à ces transformations les valeurs (L_b, M_b, S_b) qui sont les réponses résiduelles de l'écran, mesurées lorsque tous ses pixels sont éteints (RGB=(0,0,0)). Finalement, les coordonnées LMS sont obtenues par l'opération :

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \mathbf{T} \begin{bmatrix} sR \\ sG \\ sB \end{bmatrix} + \begin{bmatrix} L_b \\ M_b \\ S_b \end{bmatrix} \quad (\text{A.1})$$

où \mathbf{T} est la matrice de passage de l'espace sRGB vers l'espace LMS, (sR, sG, sB) les coordonnées sRGB des pixels dans l'espace RGB standard linéaire et (L_b, M_b, S_b) les réponses résiduelles de l'écran.

- Nous imposons finalement deux contraintes pour déterminer la pondération relative de chacun des types de cônes dans les mécanismes post-récepteurs :

- L'axe de luminance correspond à la courbe $V_M(\lambda)$, fonction d'efficacité lumineuse modifiée par Judd-Vos (Judd, 1951 ; Vos, 1978). Par définition, la somme des fonctions de sensibilité des cônes L et M de Smith et Pokorny (1975) est égale à $V_M(\lambda)$. L'axe de luminance de notre espace DKL est donc : $Lum = L + M$

- Les deux axes de chrominance s'interceptent au niveau des points gris. En d'autres termes, les points gris dans l'espace RGB (i.e. $RGB = (\gamma, \gamma, \gamma)$ avec $\gamma \in [0, 255]$) sont achromatiques dans l'espace DKL. En supposant donc que l'axe Rouge-Vert (RG) peut être décrit comme une différence pondérée des réponses des cônes L et M et l'axe Bleu-Jaune (BY) comme une différence pondérée de la réponses des cônes S et de la somme de celles des cônes L et M, alors nous avons le système d'équations suivant :

$$\text{suivant : } \begin{cases} RG &= L - \alpha M \\ BY &= 2S - \beta(L + M) \end{cases}$$

où α et β sont des paramètres dépendants de l'écran et qui doivent permettre de vérifier $RG = BY = 0$ pour tout les pixels gris. Ceci est, en particulier, vrai pour le point blanc de notre écran (RGB=(255,255,255)) de coordonnées (L_0, M_0, S_0) dans l'espace d'excitation des cônes. En résolvant le système précédent nous obtenons

$$\text{donc : } \begin{cases} \alpha &= \frac{L_0}{M_0} \\ \beta &= \frac{2S_0}{L_0 + M_0} \end{cases}$$

Finalement, les coordonnées DKL d'un pixel sont calculées à partir de ses coordonnées

LMS grâce à la matrice de passage suivante :

$$\begin{bmatrix} Lum \\ RG \\ BY \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -\alpha & 0 \\ -\beta & -\beta & 2 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (\text{A.2})$$

B

Propriétés des écrans d'expérimentation

Nous avons utilisé pour ces travaux deux écrans différents : Iiyama Vision Master Pro 513 et Philips 201B4. Les propriétés de chacun d'entre eux ont été mesurées à l'aide d'un spectroradiomètre (CS 2000, Konica-Minolta, Inc).

B.1 Iiyama Vision Master Pro 513

- Coordonnées xyY des trois primaires à leur maximum d'intensité :
 - xyY(R) = (0.6222, 0.3376, 9.0472)
 - xyY(G) = (0.2825, 0.6074, 23.9300)
 - xyY(B) = (0.1492, 0.0678, 2.4934)
- Spectre de radiance
- Non-linéarité des primaires
- Luminance résiduelle
(L_b, M_b, S_b) = (0.3028, 0.2241, 0.0345)

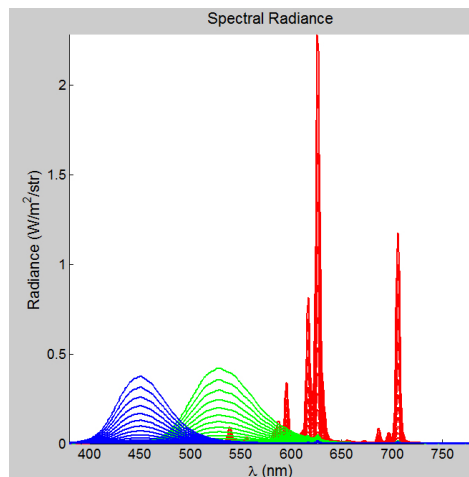


FIGURE B.1 – Spectres de radiance des trois primaires de l'écran Iiyama Vision Master Pro 513 mesurés à différents niveaux d'intensité.

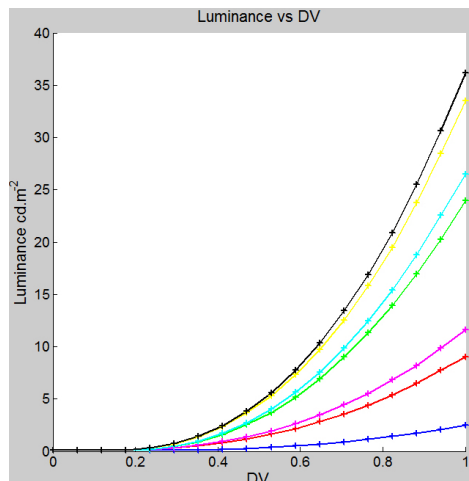


FIGURE B.2 – Luminance des trois primaires (R, G et B) de l'écran Iiyama Vision Master Pro 513 en fonction des valeurs digitales (DV, c'est-à-dire les valeurs RGB normalisées entre 0 et 1). Les courbes correspondant aux mélanges de paires de primaires (magenta, cyan et jaune) ainsi qu'à l'ensemble des trois primaires (gris) sont également présentées.

B.2 Philips 201B4

- Coordonnées xyY des trois primaires à leur maximum d'intensité :
 - $xyY(R) = (0.6221, 0.3296, 15.0485)$
 - $xyY(G) = (0.2956, 0.6016, 57.0515)$
 - $xyY(B) = (0.1519, 0.0677, 6.9209)$
- Spectre de radiance

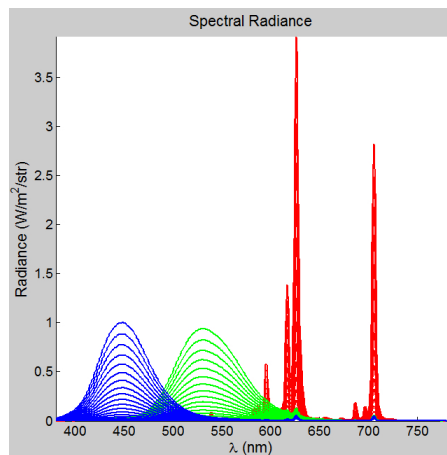


FIGURE B.3 – Spectres de radiance des trois primaires de l'écran Philips 201B4 mesurés à différents niveaux d'intensité.

- Non-linéarité des primaires

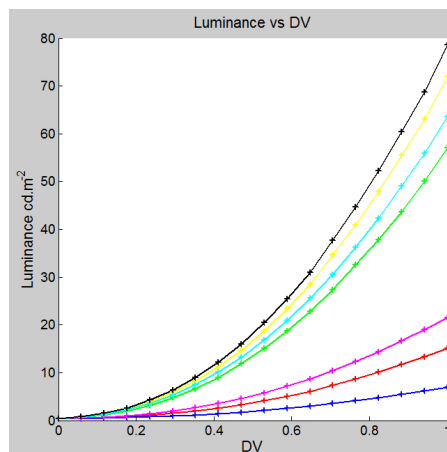


FIGURE B.4 – Luminance des trois primaires (R, G et B) de l'écran Philips 201B4 en fonction des DV. Les courbes correspondant aux mélanges de paires de primaires (magenta, cyan et jaune) ainsi qu'à l'ensemble des trois primaires (gris) sont également présentées.

- Luminance résiduelle
 $(L_b, M_b, S_b) = (1.4004, 1.0419, 0.1798)$

C

Valeurs d'information mutuelle estimées avec le
modèle de référence

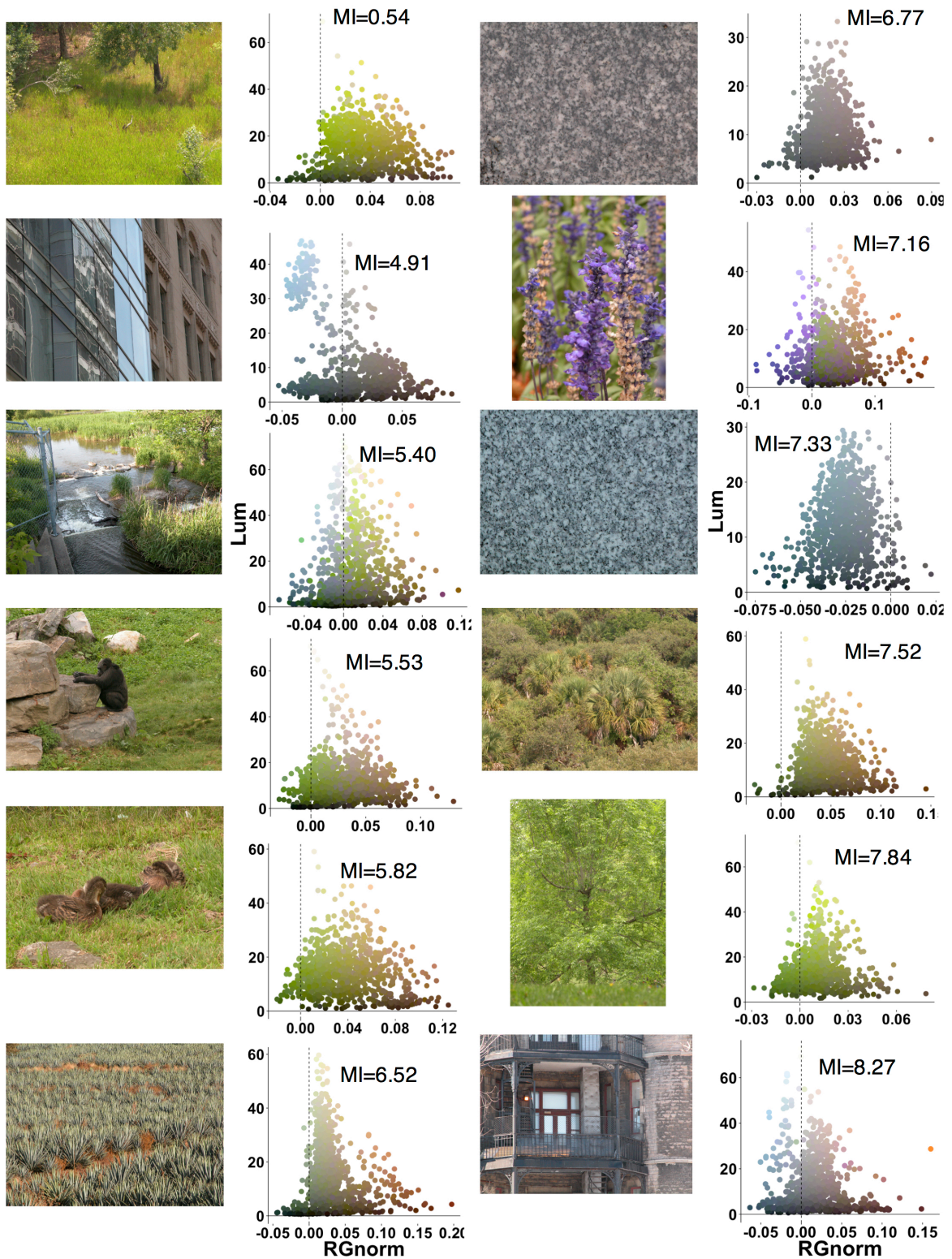


FIGURE C.1 – 20 images de plus faible information mutuelle entre contours de luminance et de chrominance Rouge-Vert obtenues avec le modèle de référence. Chaque image est présentée avec la valeur de l'information mutuelle entre contours Ed_RG et Ed_Lum ($\cdot 10^3$) et l'allure de la distribution de ses pixels dans un repère (RG,Lum).

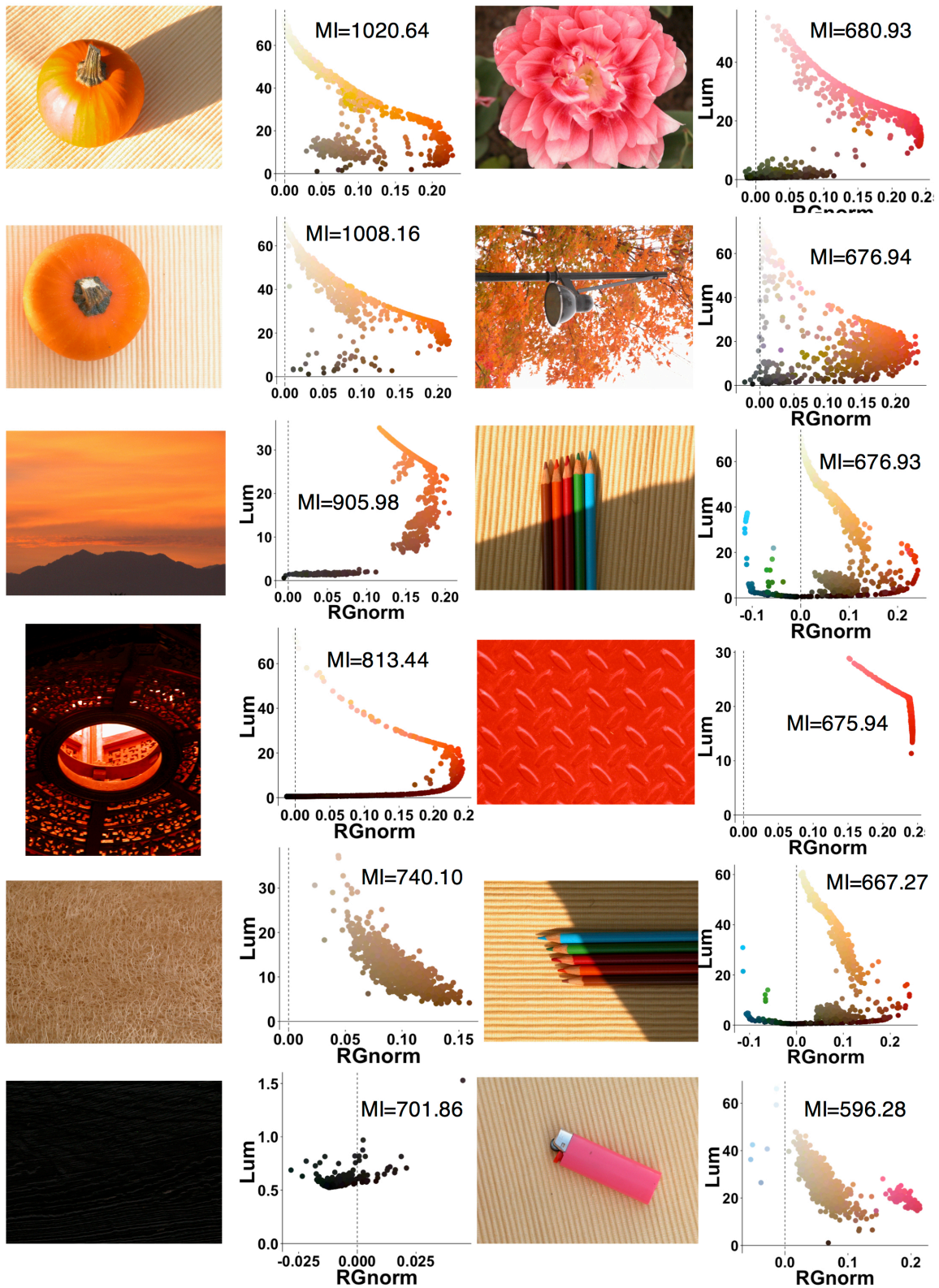


FIGURE C.2 – 20 images de plus forte information mutuelle entre contours de luminance et de chrominance Rouge-Vert obtenues avec le modèle de référence. Chaque image est présentée avec la valeur de l'information mutuelle entre contours Ed_RG et Ed_Lum ($\cdot 10^3$) et l'allure de la distribution de ses pixels dans un repère (RG,Lum).

Résumé — La couleur fait partie intégrante de notre expérience visuelle quotidienne et foisonne dans la représentation que nous avons de notre environnement. Pourtant, les mécanismes qui nous rendent cette information accessible et l'utilité de celle-ci pour la perception visuelle demeurent largement incompris et plusieurs hypothèses ont été émises à ces sujets. La vision des couleurs a été abondamment étudiée à partir de stimuli simples et artificiels, mais ces résultats ne peuvent pas trivialement être généralisés au cas de scènes plus complexes, qui nécessitent des investigations spécifiques. C'est cette problématique qui a principalement guidé notre démarche pour ces travaux de thèse. Nous nous sommes attachés à étudier les mécanismes sous-tendant le traitement de l'information de couleur (ou chrominance) et à explorer le rôle de celle-ci dans le cas de la perception de stimuli complexes et écologiquement pertinents, tels que les scènes naturelles. Nous proposons pour cela une analyse théorique visant à quantifier la redondance statistique des informations de luminance et de chrominance dans les scènes naturelles. Nous décrivons ensuite une série d'expériences psychophysiques que nous avons menées et développons un modèle d'encodage de l'information et de prise de décision permettant de tester différentes hypothèses mécanistiques concernant le traitement de l'information de chrominance par le système visuel. Enfin, nous mettons en évidence le rôle de la couleur dans une tâche spécifique : la classification de contours. L'ensemble de ces travaux permet d'éclairer en partie la compréhension des processus complexes entrant en jeu dans le traitement de la couleur par le système visuel et propose des outils méthodologiques, computationnels et expérimentaux pour l'étude de ces phénomènes perceptifs dans le cas des scènes naturelles.

Mots clés : perception visuelle, vision des couleurs, scènes naturelles, description mécanistique, observateur bayésien, modèles computationnels

Abstract — Color is an essential part of our daily visual experience and abounds in our representation of our environment. However, the mechanisms that make this information available and its usefulness for visual perception remain largely misunderstood and several hypotheses have been put forward on these topics. Color vision has been extensively studied using simple and artificial stimuli, but these results cannot be trivially generalized to more complex scenes, which require specific investigations. It is the main issue that has guided our approach for this thesis work. We focused on studying the mechanisms underlying the processing of color (or chrominance) information and exploring its role in the perception of complex and ecologically relevant stimuli, such as natural scenes. To this end, we propose a theoretical analysis to quantify the statistical redundancy of luminance and chrominance information in natural scenes. We then describe a series of psychophysical experiments we have conducted and provide an information encoding and decision-making model to test different mechanistic hypotheses regarding the processing of chrominance information by the visual system. Finally, we highlight the role of color in a specific task : the classification of edges. All of these studies partly shed light on the complex processes involved in the processing of color by the visual system and propose methodological, computational and experimental tools for the study of these perceptual phenomena in natural scenes.

Keywords : visual perception, color vision mechanisms, natural scenes, bayesian observer, computational perception
