



HAL
open science

Vérification automatique des montages d'usinage par vision : application à la sécurisation de l'usinage

Bilal Karabagli

► **To cite this version:**

Bilal Karabagli. Vérification automatique des montages d'usinage par vision : application à la sécurisation de l'usinage. Automatique / Robotique. Université Toulouse le Mirail - Toulouse II, 2013. Français. NNT : 2013TOU20107 . tel-01018079


HAL Id: tel-01018079

<https://theses.hal.science/tel-01018079>

Submitted on 3 Jul 2014

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

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



UNIVERSITÉ DE TOULOUSE - INSTITUT CLEMENT ADER-
IUT DE FIGEAC

ÉCOLE DOCTORALE SYSTEMES

T H È S E

pour obtenir le titre de

**Docteur en Systèmes automatiques et systèmes
industriels**

Présentée et soutenue par

Bilal KARABAGLI

**Vérification automatique des
montages d'usinage par vision**

Application à la sécurisation de l'usinage

préparée à l'IUT de Figeac et à l'ICA Albi,

soutenue le 06 Novembre 2013

Jury :

<i>Rapporteurs :</i>	Samia BOUCHAFA-BRUNEAU	- IBISC (Université d'Evry Val d'Essonne)
	Abdelmalik TALED-AHMED	- LAMIH (Université de Valenciennes)
<i>Directeurs :</i>	Jean-José ORTEU	- ICA (Mines Albi)
	Thierry SIMON	- ICA et IUT DE FIGEAC/LRPmip
<i>Examineurs :</i>	Claire LARTIGUE	- LURPA (ENS CACHAN)
	Michel DEVY	- LAAS (CNRS)
	Olivier AUBRETON	- L2i-CNRS (Université de Bourgogne)
	Hervé PINGAUD	- (CUFR J-F. Champollion)
<i>Invité :</i>	Djemel ZIOU	- MOIVRE (Université de Sherbrooke, Québec).



REMERCIEMENTS

Je tiens à adresser en premier lieu mes plus chaleureux remerciements à mes directeurs de thèse Jean-José ORTEU et Thierry SIMON. Ils n'ont pas simplement accepté de diriger ma thèse. Ils m'ont transmis la passion de la recherche et ont donné du sens aux termes de professionnalisme et d'organisation d'un projet de recherche, ce mélange de compétences théoriques, organisationnelles et de gestion du temps.

J'ai pu apprécier aussi leur dimension humaine et leur constant soutien durant ces trois années de thèse. Je veux leur exprimer ma plus profonde gratitude.

Je suis très honoré que Samia BOUCHAFA-BRUNEAU et Abdelmalik TALEB-AHMED aient accepté de rapporter sur mon travail de thèse.

Je remercie très sincèrement Claire LARTIGUE, Michel DEVY, Olivier AUBRETON et Hervé PINGAUD d'avoir accepté d'examiner ma thèse.

Je souhaite remercier très chaleureusement Marie-Flavie AUCLAIR-FORTIER, Djemel ZIOU et Daniel LEVESQUE de m'avoir accueilli dans le centre MOIVRE pendant la période du stage à l'université de Sherbooke, Québec.

Je tiens à remercier aussi les ingénieurs et techniciens de l'entreprise Noomeo[©], les collègues de l'IUT de Rodez et le CIRTES de Carmeaux d'avoir accepté de numériser les différents montages d'usinage par différents systèmes de numérisation tridimensionnelle. Je n'oublie pas les membres de l'équipe traitement et compréhension d'image de l'IRIT et en particulier Jean-Denis DUROU.

Un grand merci à tous les personnels de l'IUT de Figeac et en particulier les membres du LRPmip, avec qui j'ai collaboré pendant ces 3 ans. Je voudrais remercier en particulier, Patricia GONTIER de son soutien durant les différents stades de la thèse, Alain FERMY pour la conception des montages d'usinage, Safia CAILLAU pour sa gentillesse, Annie CAVARROC pour les démarches administratives. Je remercie aussi les collègues du l'ICA (Mines Albi) et, en particulier, Esther RAMIREZ et Catherine MAFFRE.

Durant ces trois années de thèse, j'ai eu la chance de découvrir un petit coin de FRANCE, la région du LOT, la ville de Figeac. J'ai beaucoup apprécié la beauté de cet endroit et la sympathie de ses citoyens. Je souhaite à cette occasion remercier, Véronique, Pascal, Didier et tous ceux avec qui j'ai sympathisé!

Enfin, je souhaite remercier ma famille, mon extraordinaire épouse Randa et mes amis, en particulier mon ami d'enfance Khaled, pour leur soutien constant. Cette thèse, c'est un aboutissement de longues années d'études, je la dois beaucoup à mes sœurs et frères, à toi très chère Randa, et en particulier à mes parents exceptionnels qui ont toujours combattu pour nous assurer, à moi et à mes frères et sœurs un climat serein afin que nous focalisions notre attention seulement sur les études. Il m'est impossible de trouver les mots pour dire à quel point je suis fier d'eux, et à quel point je les aime. Je ne peux pas conclure cette page de remerciements sans penser à ma grand-mère et aux merveilleux moments de mon enfance passés avec elle, autour d'une "bonne tasse" de lait frais de chèvre et de sa nourriture "bio". Elle était fière de moi et je suis très fier d'elle.

À Mohammed Ben Abdellah, à mes parents,
à mes frères et sœurs et toute la famille
à toi très chère Randa.
Enfin, une pensée particulière pour ma grand mère.

Table des matières

1	INTRODUCTION GÉNÉRALE	1
1.1	Introduction	2
1.2	Problématique industrielle	3
1.3	Contribution de la thèse	6
1.4	Structuration du rapport	7
	Bibliographie	10
2	SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE	11
2.1	Introduction	12
2.2	Solution actuelle avec contact	13
2.3	La vision pour la détection de collision de l'outil d'usinage	15
2.3.1	La détection de collisions de l'outil d'usinage	15
2.3.2	La reconstruction tridimensionnelle de l'outil	17
2.4	La reconstruction tridimensionnelle du montage d'usinage	19
2.5	Systèmes optiques testés	22
2.5.1	Critères conditionnant le choix de système de mesure 3D	23
2.5.2	Capteur Minolta	25
2.5.3	Capteur Viu-scan	26
2.5.4	Capteur Optinum	26
2.5.5	Capteur Kinect	26
2.5.6	Résultats de numérisation 3D du montage Norelem	28
2.6	Configuration du système de vision	29
2.7	Conclusion	31
	Bibliographie	32
3	SÉLECTION DES PRIMITIVES 2D/3D POUR LA CONSTRUCTION DES GRAPHES	35
3.1	Introduction et problématique	35
3.2	Détection des contours	37
3.2.1	Modélisation de l'histogramme du module des gradients	39
3.2.2	Méthode initiale	40
3.2.3	Amélioration proposée	43
3.2.4	Résultats et comparaison	44
3.3	Sélection des primitives 2D	50
3.3.1	Sélection des cercles	52
3.3.2	Sélection des droites	53
3.4	Sélection des primitives 3D	58
3.4.1	Création de l'image théorique	58

3.4.2	Extraction des centres des cercles 3D	62
3.4.3	Extraction des points 3D d'intersection des droites	62
3.5	Conclusion	63
	Bibliographie	64
4	ESTIMATION AUTOMATIQUE DE POSE PAR APPARIEMENT DE GRAPHERS	67
4.1	Introduction et problématique	68
4.2	Notions sur les graphes	71
4.2.1	Définition d'un graphe	71
4.2.2	Matrice d'adjacence	72
4.2.3	Isomorphisme	72
4.3	Construction des graphes	73
4.4	Appariement de graphes pour le calcul de pose : Notre proposition . . .	77
4.4.1	Calcul des sous graphes	78
4.4.2	Détermination de la relation géométrique Φ	79
4.5	Calcul de pose et extraction des régions d'intérêts	87
4.5.1	Modélisation de la caméra	87
4.5.2	Détermination des paramètres intrinsèques de la caméra	88
4.5.3	Calcul des paramètres extrinsèques de la caméra	89
4.6	Évaluation de la méthode	92
4.6.1	Évaluation de la méthode par rapport au bruit	96
4.6.2	Évaluation de la méthode par rapport à la topologie des graphes	98
4.7	Vérification de la pose et cas limites	101
4.7.1	L'impact de la matrice de rotation	101
4.7.2	L'impact du vecteur de translation	103
4.7.3	Cas du montage totalement différent du modèle numérique	104
4.8	Conclusion	105
	Bibliographie	106
5	ANALYSE LOCALE ET PRISE DE DÉCISION	109
5.1	Introduction et problématique	110
5.2	Méthodes de segmentation par régions testées	111
5.2.1	La segmentation en régions en utilisant les graphes (<i>NCut</i>)	113
5.2.2	Méthodes de segmentation en régions utilisant l'histogramme	115
5.2.3	Détermination de la forme de l'objet : stratégie proposée	118
5.2.4	Comparaison des méthodes	120
5.3	Le squelette comme outil déterminant la structure topologique	123
5.3.1	Méthodes d'amincissement	123
5.3.2	Méthodes à partir de la carte de distance	123
5.3.3	Méthodes à partir du diagramme de Voronoï	123
5.3.4	Méthodes de morphologie mathématique	124
5.3.5	Régularisation du squelette	124
5.4	Appariement de squelettes et calcul du score	125

5.4.1	Appariement de squelettes à partir de leur graphe de proximité : Méthode proposée	125
5.4.2	Calcul du score d'appariement	125
5.5	Calcul du score d'évaluation des contours et dimension	127
5.6	Les relations entre les paramètres de décision	128
5.6.1	Le squelette <i>VS</i> les contours et la dimension	128
5.6.2	Les contours <i>VS</i> le squelette et la dimension	129
5.6.3	La dimensions <i>VS</i> le squelette et les contours	130
5.7	Système expert flou pour la prise de décision	130
5.7.1	Définition des ensembles flous	131
5.7.2	Les principales opérations floues	133
5.7.3	Fuzzification	134
5.7.4	Défuzzification et décision	135
5.8	Avons-nous besoin de l'information tridimensionnelle ?	143
5.9	Conclusion	143
	Bibliographie	145
6	COMMENT OBTENIR L'INFORMATION 3D À PARTIR D'UNE SEULE CAMÉRA ?	149
6.1	Introduction et problématique	149
6.2	Formation de l'image	152
6.3	Relation entre le flou et l'éloignement	154
6.4	Méthodes de perception de la profondeur (<i>DFD</i>)	156
6.4.1	Méthode exploitant le flou sur les contours [Zhuo 2011]	157
6.4.2	Méthode adoptée [Deschenes 2003]	160
6.5	Expérimentation et résultats	163
6.6	La réflectance de l'image pour l'estimation de la carte de profondeur : <i>Shape From Shading</i>	167
6.7	Création de l'image RGBD	169
6.8	Conclusion	170
	Bibliographie	172
7	CONCLUSION ET PERSPECTIVES	175
7.1	Conclusion	175
7.2	Perspectives	177
A	RÉSULTATS	179
A.1	Les résultats sont illustrés sur des pages de format A3	179
	Bibliographie	191

Table des figures

1.1	<i>Quelques éléments modulaires constituant les montages d'usinage</i>	2
1.2	<i>Chaîne numérique de traitement de l'usinage: Problématiques de recherche</i>	3
1.3	<i>Chaîne numérique simplifiée</i>	4
1.4	<i>Quelques configurations du montage d'usinage, (a) modèle numérique donné en CAO, (b) configuration conforme au modèle CAO (c) configuration non conforme au modèle CAO, e.g. bride de fixation déplacée (d) configuration non conforme au modèle CAO, e.g. bride de fixation déplacée et élément de bridage en plus (e) configuration non conforme au modèle CAO, e.g. bride de plus et bride qui n'a pas la bonne dimension, (f) modèle numérique de la pièce finale.</i>	5
1.5	<i>Chaîne numérique simplifiée : solution proposée</i>	6
1.6	<i>Différents montages testés : (a) Montage sur plaque Norelem, (b) Modèle numérique correspondant, (c) Montage "Semelle", (d) Modèle numérique correspondant, (e) Montage "Minimanche" de réalisation de support commande de vol pour avion, (f) Modèle numérique correspondant.</i>	9
1.7	<i>Organigramme du rapport de la thèse et objectif de chaque partie.</i>	10
2.1	<i>Organigramme du chapitre 2 et objectif de chaque partie.</i>	12
2.2	<i>Bras MMT FARO GAGE pour la mesure tridimensionnelle</i>	13
2.3	<i>Principe de fonctionnement d'un palpeur</i>	14
2.4	<i>Palpeur à transmission IR monté sur la broche [Peyrucat 2013].</i>	14
2.5	<i>Algorithme proposé par [Ahmad 2010] pour la détection et correction de la trajectoire de l'outil</i>	16
2.6	<i>Principe de l'algorithme de détection des coins</i>	16
2.7	<i>Dimensionnement de l'outil d'usinage, (a): porte outil, (b): modèle 3D de porte outil, (c): cas de collision dû au mauvais dimensionnement de l'outil (zone rouge)</i>	17
2.8	<i>L'effet de la source lumineuse dans l'extraction des contours de l'outil [Zhang 2011]</i>	18
2.9	<i>Principe de FMBR X. ZHANG et al. [Zhang 2011]</i>	18
2.10	<i>Résultat de la reconstruction 3D de l'outil d'usinage [Zhang 2011], (a) : outils utilisés dans l'expérimentation, (b) : résultat de reconstruction 3D</i>	19
2.11	<i>Installation du système de vision [Tian 2010]</i>	19
2.12	<i>Algorithme proposé par [Tian 2010]</i>	20
2.13	<i>Reconstruction 3D théorique du montage d'usinage</i>	22
2.14	<i>Critères de choix de système de numérisation 3D (www.innoumetric.com)</i>	24
2.15	<i>Système Minolta[®], (A): capteur Minolta vi-9i, (B): Table de rotation</i>	25
2.16	<i>Système Viiu-scan[®] de Creaform</i>	26

2.17	<i>Système Optinum[®] Noomeo</i>	27
2.18	<i>(A): Capteur Kinect[®] de Microsoft, (B): Schéma de principe de fonctionnement comme est donné par Microsoft, les capteurs en bleu sont des composants facultatifs pour l'estimation de la profondeur de la scène</i>	27
2.19	<i>(a) : Modèle 3D du montage d'usinage, (b) : Montage d'usinage composé d'une : palette (rouge), pions de positionnement (vert), brut (bleu) et bride de fixation (jaune)</i>	28
2.20	<i>Résultats de numérisation 3D du montage Norelem (cas conforme)</i>	29
2.21	<i>Configuration du système de vision sur deux types de machines, (a) Zone de préparation du montage d'usinage, (b) le montage préparé est envoyé à la zone d'usinage, (c) machine d'usinage ne disposant pas d'une zone de préparation du montage, la préparation se fait directement dans la zone de l'usinage, (d) schéma des axes de la machine.</i>	30
3.1	<i>Organigramme du chapitre 3 et objectif de la partie.</i>	36
3.2	<i>Classification des méthodes de binarisation selon G. Papari et N. Petkov [Papari 2011], et le positionnement de la méthode proposée</i>	38
3.3	<i>Modélisation de l'histogramme du module des gradients par une exponentielle décroissante avec, $\alpha = 0.2319$</i>	42
3.4	<i>Résultat du seuillage par la méthode initiale sur des images de la base BSDS300.</i>	43
3.5	<i>Modélisation de l'histogramme par la méthode proposée. Image de la base BSDS300.</i>	44
3.6	<i>Seuil de binarisation. (a): point de croisement des tangentes D1 et D2. (b): Zoom avant sur le deuxième mode.</i>	44
3.7	<i>Résultat d'un extrait de la base BSDS300</i>	47
3.8	<i>Résultat d'un extrait de la base BSDS300</i>	48
3.9	<i>Courbes PR : en rouge continu la méthode Otsu [Ku 2005] avec $F - score = 0.50$, en vert continu, la méthode d'entropie [Kapur 1985] avec $F - score = 0.57$, en noir continu la méthode initiale avec $F - score = 0.48$ et en bleu continu, la méthode proposée avec $F - score = 0.56$</i>	49
3.10	<i>Courbes ROC : en rouge la méthode Otsu [Ku 2005], en vert la méthode d'entropie [Kapur 1985], en noir la méthode initiale et en bleu la méthode proposée.</i>	49
3.11	<i>Diagramme d'exactitude: en bleu la valeur moyenne d'exactitude, et en vert l'écart-type correspondant. Le calcul a été fait sur des centaines d'images tirées aléatoirement de la base BSDS300.</i>	50
3.12	<i>Évaluation du temps de calcul (CPU Time) de seuil de binarisation pour chaque méthode</i>	50
3.13	<i>Seuillage automatique: (a) : image du module des gradients calculée par l'opérateur Prewitt, (b) : résultat de binarisation par la méthode initiale, (c) : résultat de binarisation par la méthode Otsu, (d) : résultat de binarisation par la méthode entropie et (e) : résultat de binarisation par la méthode proposée.</i>	51

3.14	<i>Différentes éléments de bridage (Catalogue CARRLANE)</i>	51
3.15	<i>Détection des cercles : (a) image des niveaux de gris, (b) : détection des cercles sur l'image des contours. Les cercles mentionnés avec un centre jaune reflètent les cercles pertinents</i>	53
3.16	<i>Calcul des régions portant les segments : (a) principe de rectangle approximatif de la région [Grompone von Gioi 2010], (b) test $k(r, x)$, (c) résultat de calcul de régions sur l'image du montage d'usinage</i>	54
3.17	<i>Détection des segments de droites : (a) les segments; l'ellipse jaune représente une région possédant des segments qui pourraient être représentés par une seule droite (b) : les droites portées par les segments.</i>	55
3.18	<i>Critère d'orientation (α)</i>	56
3.19	<i>Critère de distance (d)</i>	56
3.20	<i>Sélection des coins : (a) les droites des segments, (b) résultat de calcul obtenu après fusion des régions. Les intersections des droites sont mentionnées en jaune.</i>	57
3.21	<i>Reconnaissance des éléments par transformation géométrique</i>	58
3.22	<i>Pose virtuelle donnée par la matrice $[T_v]$</i>	59
3.23	<i>Image théorique du modèle CAO, (a) image du montage d'usinage, (b) image des contours</i>	61
3.24	<i>Extraction des centres des cercles 3D à partir de l'image théorique 2D et leurs correspondants dans le modèle CAO.</i>	62
3.25	<i>Extraction des points d'intersection 3D à partir de l'image théorique 2D et leurs correspondants dans le modèle CAO.</i>	63
4.1	<i>Organigramme du chapitre 4 et objectif de chaque partie.</i>	69
4.2	<i>Matrice d'adjacence et matrice de degré de connectivité des nœuds.</i>	72
4.3	<i>Graphes isomorphismes : (a) Graphe non-orienté, (b) Graphe isomorphisme au graphe de la figure (a).</i>	73
4.4	<i>Graphes de proximités : (a) principe de calcul du graphe SIG, (b) principe de calcul du graphe de Delaunay</i>	74
4.5	<i>Principe de calcul du graphe Gabriel : (a) graphe de Gabriel, (b) graphe de Delaunay</i>	75
4.6	<i>Construction automatique des graphes : cas du montage conforme.</i>	75
4.7	<i>Construction automatique des graphes pour différentes configurations non-conformes du montage d'usinage</i>	76
4.8	<i>L'impact de "l'infidélité" du modèle CAO à la réalité dans la conception du montage d'usinage, (a) image de la bride réelle, (b) modèle numérique de la bride correspondante à (a), (c) contours sur la zone de la bride sur l'image réelle du montage d'usinage, (d) contours sur la zone de la bride dans l'image théorique.</i>	77
4.9	<i>Calcul de la coupe par analyse de la nature d'arêtes: Critère de partitionnement.</i>	79
4.10	<i>Calcul des sous-graphes pour différentes configurations du montage d'usinage</i>	80

4.11	<i>Rainures de fixation sur la table d'usinage¹.</i>	82
4.12	<i>Principe de changement du repère et la recherche de la configuration optimale ϕ : Le nœud P_3 du graphe G et le nœud P_{ref} du graphe H présentent les nœuds de changement du repère. les colonnes (b), (c) et (d) représentent les différentes configurations Φ_i, la configuration optimale est donnée par la colonne (d) ϕ_i^{Opt}.</i>	83
4.13	<i>Détermination du voisinage pour la configuration Φ_i^{Opt} de la figure 4.12. Les deux graphes G en bleu et H en vert sont projetés dans le même espace de variation dimensionnelle.</i>	84
4.14	<i>Résultat de la mise en correspondance des graphes réel (G) et théorique (H) : (a) Configuration retenue Φ_2, (b) : Tableau des appariements, la ligne "grisée" indique les nœuds de changement de repère. (c) : Diagramme des barres pour différentes configuration Φ_q, dans la config.1 correspond à Φ_0 pour laquelle, aucune rotation de plan n'est appliquée. La config.2 correspond à une rotation de plan $\Phi_1 = \pi/2$. Cette configuration donne de meilleurs appariements, e.g. un maximum de nœuds appariés avec une distance minimale. La config.3 correspond à une rotation de plan $\Phi_2 = \pi$. Enfin, la config.4 correspond à une rotation du plan $\Phi_2 = 3\pi/2$</i>	85
4.15	<i>Mise en correspondance automatique des graphes pour différentes configurations non conforme du montage d'usinage</i>	86
4.16	<i>Modèle de caméra "pinhole" et relation entre ses paramètres.</i>	87
4.17	<i>Définition des paramètres de la matrice de rotation.</i>	90
4.18	<i>Estimation des paramètres intrinsèques et de la pose de la caméra par rapport au montage d'usinage, dans le cas conforme : (a) La pose de la caméra exprimée par rapport au repère du montage d'usinage, donné par le modèle CAO, (b) la pose de la caméra exprimée par rapport au repère de la caméra, (c) la matrice des paramètres intrinsèques déterminés lors du calibrage hors-ligne et (d) la matrice des paramètres extrinsèques ou de la pose de la caméra par rapport au montage d'usinage.</i>	91
4.19	<i>Localisation des éléments de bridage: (a) graphe des attributs (ARG) du modèle CAO, (b) projection du modèle CAO sur l'image réelle du montage d'usinage, (c) Calcul du graphe des attributs après estimation de la pose. (d) localisation des éléments de la scène à partir du ARG projeté.</i>	92
4.20	<i>Localisation automatique des régions d'intérêts pour différentes configurations conformes et non conformes du montage Norelem[©].</i>	93
4.21	<i>Localisation automatique des régions d'intérêts pour différentes configurations conformes et non conformes du montage semelle.</i>	94
4.22	<i>Localisation automatique des régions d'intérêts pour différentes configurations conformes et non conformes du montage Minimanche.</i>	95

4.23	<i>Schéma simplifié de l'étude de la robustesse des paramètres de chaque section du chapitre par rapport à la précision de la pose. La ligne "verte" présente le test de l'évaluation du bruit impulsif, la ligne "rouge" présente le test de l'évaluation du bruit gaussien et la ligne "bleu" illustre le test de l'évaluation de la topologie des graphes.</i>	96
4.24	<i>Évaluation de la pose par rapport au bruit : (a) Erreur moyenne en pixels par rapport à l'axe x et y, (a) Écart-type moyen en pixels par rapport à l'axe x et y, (c) Images théoriques utilisées pour le test.</i>	97
4.25	<i>Évaluation de la pose par rapport aux coordonnées des points d'intérêts.</i>	98
4.26	<i>Évaluation de la pose par rapport au nombre des nœuds appariés pour le cas conforme.</i>	98
4.27	<i>Évaluation de la pose par rapport à la structuration du graphe, i.e. distribution spatiale des nœuds : (a) Erreur moyenne en pixels par rapport à l'axe x et y, (b) Écart-type moyen en pixels par rapport à l'axe x et y, (c) Configurations des graphes correspondants. La configuration 29 représente la configuration dont la valeur moyenne de l'erreur de pose est minimale, avec une moyenne en $Mean(x, y) = (1.2063, 2.3697)$ et un écart-type $std(x, y) = (1.3732, 1.6096)$</i>	100
4.28	<i>Détection de l'anomalie dans le calcul de pose à partir du graphe ARG : (a) Graphe ARG_{th} selon sa définition en CAO, (b) : projection du modèle numérique CAO sur l'image du montage réalisé, (c) : Graphe ARG_r calculé après estimation de la pose, (d) : Segmentation d'image et extraction des régions d'intérêts, (e) Matrice adjacence du graphe ARG_{th} selon la définition CAO, (f) Matrice adjacence du graphe ARG_r. Les cases rouges dans la matrice (f) représentent l'anomalie, décrit par la définition d'un nouveau voisinage qui n'est pas défini en ARG_{th}.</i>	101
4.29	<i>Influence de point de vue dans détermination des régions d'intérêts. (a) la composante Z du vecteur translation veut $Z = 700mm$, (b) la composante Z du vecteur translation veut $Z = 1200mm$ et (c) cas d'une pose correcte avec une bride qui n'a pas la bonne dimension</i>	104
4.30	<i>Robustesse de notre méthode pour des cas extrêmes. (a) Erreur moyenne de la pose, (b) écart-type moyen calculé sur les points appariés, (c) modèle numérique du montage Norelem[©] projeté sur l'image du montage Semelle, (d) Graphe des attributs du montage Norelem[©] projeté sur l'image du montage Semelle, (e) localisation des régions d'intérêts</i>	105
5.1	<i>Organigramme du chapitre 5 et objectif de chaque partie.</i>	111
5.2	<i>Décomposition du montage d'usinage en régions d'intérêt des éléments de bridage pour la reconnaissance de formes. (a) Image du montage d'usinage, (b),(c),(d),(e),(f),(g),(h) et (i) imagerie des régions d : brut, Bride gauche, bride droite, pion 1, pion 2, pion 3, pion 4 et pion 5, successivement.</i>	112

5.3	<i>Images binaires de référence des éléments de bridage du montage Semelle. (a) image binaire de la bride1, (b) image binaire du brut (c) image binaire de la bride2, (d) image binaire du pion1, (e) image binaire du pion2, (f) image binaire du pion3, (g) image binaire du pion4 et (h) image binaire du pion5.</i>	112
5.4	<i>Principe de l'algorithme Ncut [Shi 2000] : (A) un simple graphe et la coupe de poids minimale ; (B) tableau des associations et des coupes pour ce graphe. assoc et cut sont calculés à partir de la matrice de poids W.</i>	114
5.5	<i>Exemple d'application de l'algorithme NCut sur l'image du pion du montage Minimanche. (a) le degré de partitionnement calculé sur les contours, (b) les contours de partitionnement et (c) les régions de l'image.</i>	115
5.6	<i>Modélisation de l'histogramme de l'image du pion du montage Semelle par l'algorithme GMM-EM. Application à la segmentation par régions</i>	117
5.7	<i>éléments structurants utilisés dans la morphologie d'image. (a) de voisinage 4, (b) de voisinage 8.</i>	118
5.8	<i>Opération de dilatation en utilisant un filtre de voisinage 8×8.</i>	119
5.9	<i>Opération de dilatation en utilisant un filtre de voisinage 8×8.</i>	119
5.10	<i>Opérations Opening et Closing en utilisant un élément structurant de voisinage 4×4.</i>	120
5.11	<i>Résultat de différentes étapes de traitement proposées pour le calcul de la forme de l'objet, (a) image de luminance de la région de brut, (b) binarisation par la méthode d'Otsu, (c) application de l'opération de morphologie mathématique, ouverture et (c) restitution de la forme de l'objet par le calcul de la plus grande région.</i>	120
5.12	<i>Résultats de comparaison des méthodes de segmentation par régions testées.</i>	122
5.13	<i>Mise en correspondance de squelettes du brut. (a) image du brut réel calculé, (b) image de référence du brut théorique, (c) squelette correspondant à (a), (d) squelette correspondant à (b), (e) score d'appariement pour chaque nœud du graphe réel, (f) appariement des graphes</i>	126
5.14	<i>Calcul du score d'évaluation de contours et le score de dimension. (a) image binaire du brut réel calculé, (b) image des contours du brut réel, (c) image binaire du brut de référence, (d) image des contours du brut de référence, (e) score d'évaluation de contours, (f) score de dimension donné par la surface</i>	128
5.15	<i>Deux formes ayant le même squelette, mais elles n'ont pas la même dimension ni les mêmes contours.</i>	129
5.16	<i>Deux formes ayant les mêmes contours, mais elles n'ont pas la même dimension ni le même squelette.</i>	129
5.17	<i>Deux formes ayant la même dimension, mais elles n'ont pas le même squelette, ni les mêmes contours.</i>	130
5.18	<i>Définition du degré d'appartenance.</i>	133
5.19	<i>Exemple d'opérations sur les ensembles flous : (a) intersection floue, (b) complément flou et (c) union floue</i>	134

5.20	<i>Définition des règles floues, (a) architecture général du système expert flou, (b) règle de score d'appariement de squelettes, (c) règle de score CE, (d) règle de score de dimension et (e) règle de la décision</i>	136
5.21	<i>Relation 2D entre les règles floues, (a) relation 2D entre l'ensemble flou CE et l'ensemble flou Dimension, (b) relation 2D entre l'ensemble flou CE et l'ensemble flou Score d'appariement de squelettes, (a) relation 2D entre l'ensemble flou Dimension et l'ensemble flou Score d'appariement de squelettes</i>	136
5.22	<i>Relation 3D entre les variables d'entrées et la variable de sortie floues</i>	137
5.23	<i>Relation entre les méthodes de fuzzification et la forme des règles floues. Exemple de www.mathworks.com</i>	137
5.24	<i>Principe de l'inférence Mamdani et agrégation des règles de sortie, (a) calcul de taux d'appartenance, (c) inférence et calcul de la valeur de la sortie.</i>	138
5.25	<i>Inférence Sugeno et agrégation des règles de sortie, (a) calcul de taux d'appartenance, (c) inférence et calcul de la valeur de la sortie.</i>	138
5.26	<i>Nécessité de l'information 3D, (a) Image du montage Norelem sans les éléments de bridage analysés, (b) localisation des régions d'intérêt, (c) Segmentation par la méthode Ncut, (d) Segmentation par la méthode proposée, (e) segmentation par la méthode GMM et (f) carte de profondeur de la scène.</i>	144
6.1	<i>Organigramme du chapitre 6.</i>	151
6.2	<i>Formation d'image : (a) image nette formée par les rayons Paraxiaux, (b) image nette focalisée, (c) réponse impulsionnelle correspondant au système donné en (b), (d) image formée par un ajustement de l'ouverture (L) et (e) réponse impulsionnelle du système donné en (d).</i>	153
6.3	<i>Relation la taille du pixel et la profondeur du champs par rapport au flou optique.</i>	154
6.4	<i>Réponse impulsionnelle du système optique donnée par [Stokseth 1969], (a) le modèle Stokseth, (b) la tâche floue créée sur l'image.</i>	154
6.5	<i>Relation flou-éloignement, (a) cas de flou proche au plan image, e.g. $\xi < 0$, (b) cas de flou loin au plan image, e.g. $\xi \geq 0$</i>	155
6.6	<i>Relation flou-éloignement par rapport à la caméra.</i>	156
6.7	<i>Perception du flou optique sur les images, (a) contours flous sur le texte, (b) texture floue.</i>	156
6.8	<i>Estimation de l'indice de profondeur sur les contours [Zhuo 2011].</i>	158
6.9	<i>Résultat de l'estimation de la carte de profondeur sur des images réelles donnée par la méthode [Zhuo 2011], (a) images couleurs, (b), matrice éparse, (c) carte de profondeur</i>	159
6.10	<i>Résultat de l'estimation de la carte de profondeur sur des images du montage d'usinage par la méthode [Zhuo 2011], (a) images du montage d'usinage, (b), matrice sparse, (c) carte de profondeur estimée.</i>	160
6.11	<i>Relation flou-éloignement par rapport à la caméra.</i>	163

6.12	<i>Calibration de la caméra pour la détermination de la relation entre le flou et l'éloignement.</i>	164
6.13	<i>Détermination des valeurs de σ_i, avec $i \in R, G, B$ (a) σ_R, (b) σ_G, (c) σ_B</i>	165
6.14	<i>Détermination des valeurs du rapport des α_{RG}, α_{RB} et α_{GB}</i>	165
6.15	<i>Relation entre le flou et l'éloignement pour les 3 canaux couleurs.</i>	166
6.16	<i>Résultat de l'estimation de la carte de profondeur. (a) image de 3 plans, (b) carte de profondeur estimée, (c) image du montage d'usinage sur la plaque Norelem[©], (d) carte de profondeur estimée.</i>	167
6.17	<i>Résultat de l'estimation de la carte de profondeur à partir de méthode shape from shading. (a) image du montage d'usinage sur la plaque Norelem[©] après application de ressuage et création de texture, (b) carte de profondeur estimée.</i>	169
6.18	<i>Création de l'image RGBD : (a) nuage de points 3D, (b) Zoom avant sur le nuage du point, (c) image de profondeur (d) image de profondeur après application du filtre médian.</i>	170
6.19	<i>Segmentation de l'image RGBD : (a) image de luminance de la région du brut, (b) image de profondeur de la région du brut, (c) image RGBD de la région du brut (d) image de luminance de la région de la bride de fixation, (e) image de profondeur de la région de la bride, (f) image RGBD de la région de la bride</i>	171
6.20	<i>Représentation de l'intérêt de l'information 3D pour la segmentation : (a) image de luminance du montage d'usinage, dont les régions d'intérêt sont remises au fond, "0", (b) image de profondeur donnant l'apparition de l'anomalie, (c) image segmentée par la méthode d'Otsu.</i>	171

Liste des tableaux

2.1	<i>Quelques critères d'évaluation des systèmes de numérisation 3D. Par exemple en utilisant une notation entre ★ et ★★★★★ (moins important et très important)</i>	23
3.1	<i>Corrélation des experts</i>	42
3.2	<i>Table de confusion</i>	45

NOTATIONS

CAO	Conception Assistée par Ordinateur
FAO	Fabrication Assistée par Ordinateur
CFAO	Conception et Fabrication Assistée par Ordinateur
CL	Cutter Location file
UGV	Usinage à Grande Vitesse
CN	Commande Numérique
MOCN	Machine Outil à Commande Numérique
ROI	Région(s) d'intérêt
<i>VP</i> (vrai positif)	Correctement détecté comme positif par le test
<i>FP</i> (faux positif)	Déclaré positif par le test alors qu'il est négatif en réalité
<i>VN</i> (vrai négatif)	Correctement déclaré comme négatif
<i>FN</i> (faux négatif)	Déclaré négatif par le test alors qu'il est positif en réalité
<i>GSG</i>	Les plus grand sous-graphes
CE	Évaluation du contours
DFD	Depth From Defocus
DFF	Depth From Focus
SFS	Shape From Shading

INTRODUCTION GÉNÉRALE

Sommaire

1.1	Introduction	2
1.2	Problématique industrielle	3
1.3	Contribution de la thèse	6
1.4	Structuration du rapport	7
	Bibliographie	10

Mots clés : Montage d'usinage, Chaîne numérique du traitement de l'usinage, FAO, CN.

1.1 Introduction

"L'usinage à porte fermée", est un terme fréquemment employé par les PME de l'usinage des pièces mécaniques destinées à l'industrie de l'aéronautique et de l'automobile. Il désigne l'automatisation sécurisée du processus d'usinage. Pour répondre à ce besoin, ce travail vient contribuer à la mise en œuvre d'un système de vision permettant la validation de conformité des montages d'usinage en vue d'assurer la sécurité de l'usinage opéré sur des machines outils de type fraiseuse, notamment grande vitesse, pour lesquelles les réparations sont **très coûteuses**.

Dans le domaine de l'industrie, la réalisation d'une pièce mécanique nécessite une étude technique de fabrication menée par le bureau des méthodes. Ce dernier, doit générer le programme de pilotage des axes de la machine permettant la réalisation de la pièce finale par enlèvement de la matière. Le code de pilotage des axes de la machine cible est calculé en tenant compte du modèle géométrique du montage d'usinage donné en CAO. Ce dernier a été conçu selon la forme de la pièce à réaliser et le type de machine cible. Le montage d'usinage est un dispositif qui permet de fixer le brut à usiner contre les efforts de coupe pendant la phase de l'usinage. Il est composé essentiellement par des éléments de fixation, éléments d'appui et éléments de positionnement. La figure 1.1 montre quelques éléments modulaires constituant les montages d'usinage.

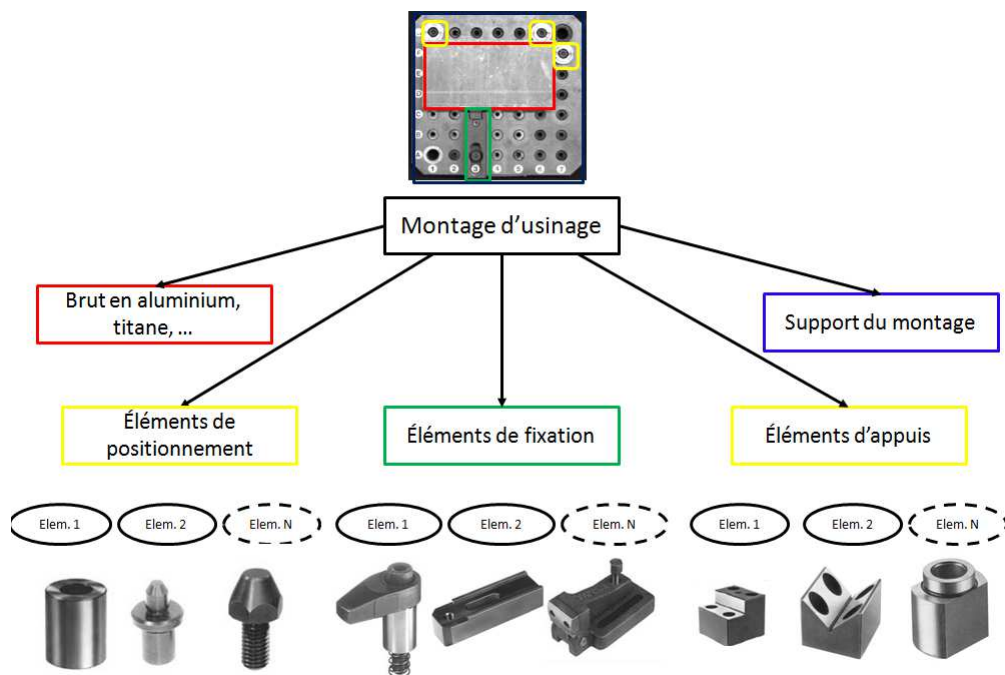


Figure 1.1 – Quelques éléments modulaires constituant les montages d'usinage

1.2 Problématique industrielle

Les difficultés de l'usinage sur des machines outils à commande numérique (MOCN) peuvent être placées essentiellement aux différentes étapes de la chaîne numérique de traitement de l'usinage. La figure 1.2 montre, par les étoiles, les difficultés qui sont à l'origine des thématiques de recherche sur la sécurité des machines outils et le contrôle de l'usinage.

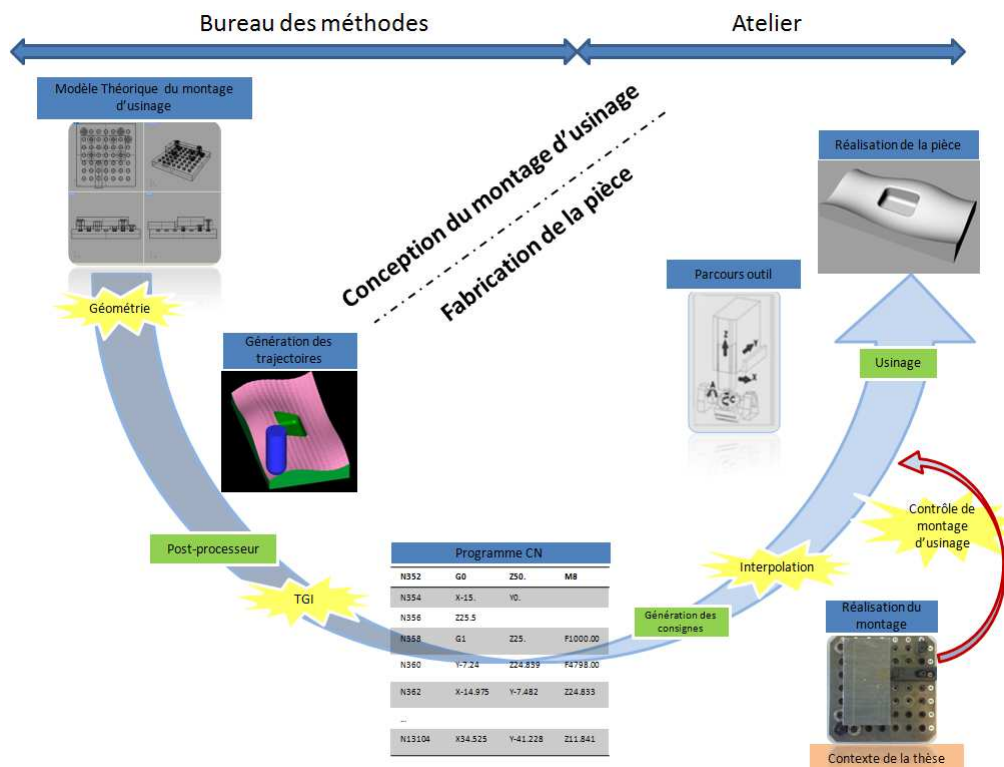


Figure 1.2 – Chaîne numérique de traitement de l'usinage : Problématiques de recherche

Dans la phase de conception du montage d'usinage, la création des formes des pièces mécaniques selon leurs fonctionnalités imposées par le cahier des charges fait appel à un processus intellectuel interprétant les contraintes pour **l'organisation des formes géométriques** [Caux 2012] des pièces à l'aide des boîtes à outils du logiciel de conception. La réalisation de la pièce par enlèvement de matière nécessite de configurer et de dimensionner les éléments de bridage pour tenir compte des efforts de coupe et des phénomènes vibratoires de l'outil. Ces différents aspects conditionnent la trajectoire de l'outil dans une génération automatique par FAO et permettent l'analyse de l'accessibilité de l'outil d'usinage via un post-processeur, dédié à la machine cible, autorisant une simulation d'usinage.

Les post-processeurs gèrent les transformations géométriques inverses (TGI) dans le lien entre l'outil et le brut maintenu. La **TGI** et la gestion des points singuliers, pour assurer la fluidité et la continuité de la trajectoire d'usinage, restent une difficulté majeure. Or le parcours de l'outil doit être calculé pour respecter les spécifications

géométriques de la pièce finie, en évitant les discontinuités géométriques et les changements brusques en orientation et translation de l'outil dans l'espace d'usinage. Le lecteur intéressé par les détails trouvera des explications de ce phénomène pour une machine 5 axes usinant par flanc d'outil dans la thèse de P-Y. Pechard [Pechard 2009].

La commande numérique (CN) des axes de déplacement interprète un code standard représentant les positions, vitesses de coupe et de déplacement, en s'appuyant sur des interpolations différentes (linéaire, circulaire et polynomiale, etc.). Certains experts proposent de conditionner les stratégies de l'usinage par une vérification d'interpolations opérées par la CN dans une phase d'étalonnage. Cet étalonnage permet de modifier les programmes via le post-processeur afin d'intégrer les spécifications géométriques et les états de surfaces liés au comportement vibratoire de la machine. Pour plus de détails sur l'analyse dynamique, cinématique et la modélisation géométrique des machines d'usinage, le lecteur est invité à consulter la monographie de C. Tournier *et al.*[Tournier 2010].

La génération de trajectoire consiste à calculer le parcours de l'outil qui passe par les points pilotés, points essentiels pour respecter la forme estimée à partir du modèle CAO et selon le type de la machine (3 ou 5 axes, etc.). Le programme FAO fourni par le bureau des méthodes est classiquement constitué d'une séquence de points et de divers paramètres d'usinage comme, par exemple, le changement d'outils, la vitesse de déplacement, la vitesse d'usinage, constituant un programme appelé «code G» **norme ISO 6983**, ou plus récemment le code «STEP-NC» **norme ISO 10303**. Cependant, les logiciels de CFAO ne produisent que des solutions génériques d'usinage dont le format n'est pas reconnu par la CN. Une fois le programme transféré à l'atelier, l'opérateur doit l'intégrer dans la machine à l'aide de l'interprétation opérée par un post-processeur dédié. La figure 1.3 montre la chaîne numérique simplifiée pour la réalisation d'une pièce. Le post-processeur est donc un élément indispensable dans la chaîne numérique de traitement, car il permet d'adapter le fichier CL : *Cutter Location file* aux spécificités de la machine.

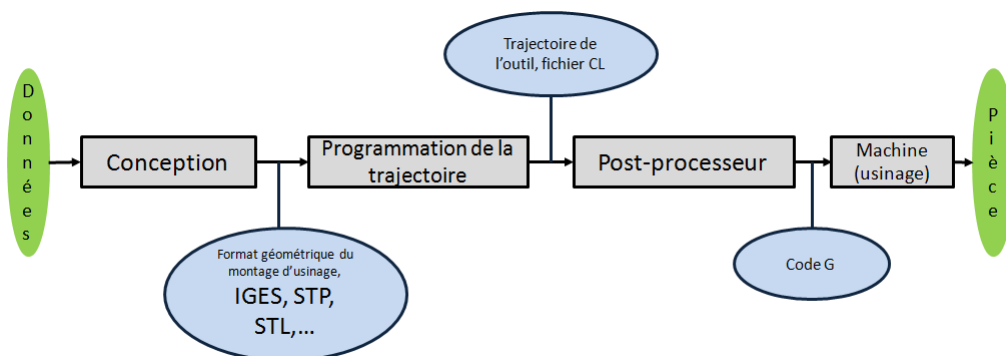


Figure 1.3 – Chaîne numérique simplifiée

Dans ce travail, nous nous intéressons aux problèmes de vérification automatique des montages d'usinage avant de lancer l'opération d'usinage proprement dite. L'objectif consiste à s'assurer que l'ensemble des éléments de bridage constituant le montage

d'usinage sont **présents**, au **bon endroit** et ont la **bonne forme et dimension**. Nous utilisons pour cela, le modèle numérique de référence donné par le fichier CAO, car le parcours de l'outil a été calculé par rapport à ce dernier.

De plus, il n'est pas rare que dans les phases d'optimisation des trajectoires de coupe, phases importantes de mise au point assurant un gain de productivité, l'opérateur soit amené à déplacer les éléments de bridage du brut à usiner. Ce déplacement n'est pas toujours transféré dans le système de FAO pour le calcul de la nouvelle trajectoire, à cause du coût prohibitif de ce transfert en terme de temps. Ce transfert est remplacé par un essai de validation dans le vide pas toujours maîtrisé. Cet essai devient impossible pour les processus d'usinage grande vitesse auxquels nous nous intéressons. En dehors de cette phase de mise au point, l'augmentation permanente des carnets de commande dans le milieu de production des pièces aéronautiques ou automobiles impose qu'une machine soit utilisée pour l'usinage de différentes pièces, parfois dans une même journée, demandant à l'opérateur d'assurer des montages multiples. Une éventuelle erreur du montage peut ainsi provoquer des incidents graves. La figure 1.4 montre quelques configurations "conforme" et "non conforme" du montage d'usinage, réalisées par l'opérateur, et le modèle numérique de la pièce finale.

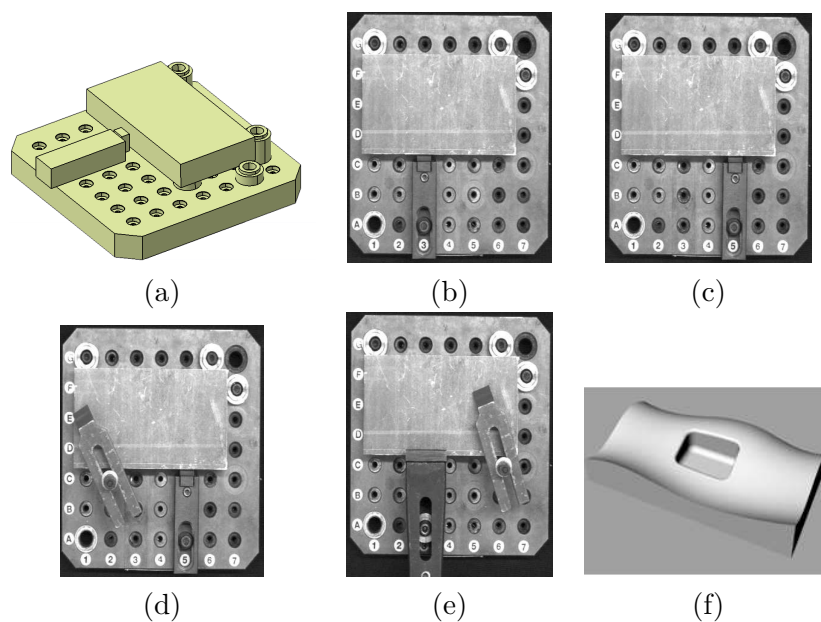


Figure 1.4 – Quelques configurations du montage d'usinage, (a) modèle numérique donné en CAO, (b) configuration conforme au modèle CAO (c) configuration non conforme au modèle CAO, e.g. bride de fixation déplacée (d) configuration non conforme au modèle CAO, e.g. bride de fixation déplacée et élément de bridage en plus (e) configuration non conforme au modèle CAO, e.g. bride de plus et bride qui n'a pas la bonne dimension, (f) modèle numérique de la pièce finale.

1.3 Contribution de la thèse

Si on reprend la figure 1.3, on peut remarquer que la chaîne numérique du traitement est unidirectionnelle et constitue une boucle ouverte. Les changements menés par l'opérateur au pied de la machine, à la fois, au niveau du programme dans la CN et de la configuration réalisée du montage d'usinage ne sont pas pris en compte dans la génération de la nouvelle trajectoire. Il manque un retour d'information "feedback" pour le contrôle et la correction de la trajectoire de l'outil. De plus le passage du modèle CAO au fichier CL, entraîne des pertes d'informations géométriques entre le modèle CAO et le code CN.

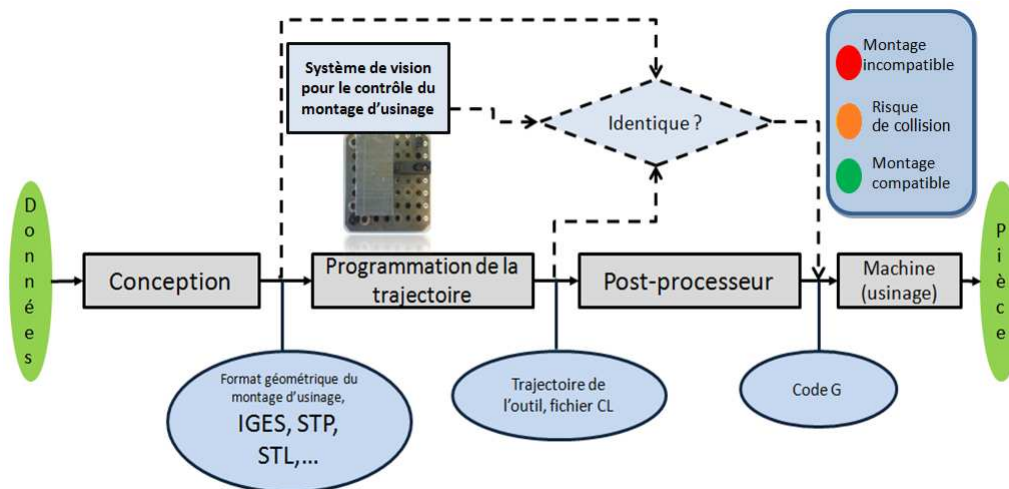


Figure 1.5 – Chaîne numérique simplifiée : solution proposée

L'automatisation sécurisée de l'usinage nécessite une rétroaction passive *feedback* sur les étapes de la chaîne du traitement numérique pour contrôler et avertir l'opérateur de toute modification provoquant un problème de collision de l'outil d'usinage avec les éléments de bridage. Pour remédier ce problème, nous proposons un contrôle permanent par un système de vision constitué d'une seule caméra. La figure 1.5 montre le positionnement de notre solution dans la chaîne du traitement simplifiée. Dans cette nouvelle chaîne proposée, trois paramètres sont nécessaires pour assurer la sécurité et le contrôle du montage d'usinage, à savoir le modèle de référence donné en CAO, l'image du montage réalisé, acquise par la caméra, et enfin un paramètre supplémentaire donné par la trajectoire de l'outil. Ce paramètre pourrait permettre de réaliser une simulation de la trajectoire de l'outil dans le cas **orange**, e.g. le montage d'usinage n'est pas conforme au modèle CAO, mais la trajectoire de l'outil ne provoque pas de collision. Le voyant **rouge**, indique à l'opérateur que le montage n'est pas conforme et peut provoquer des collisions de l'outil avec les éléments de bridage. Le voyant **vert**, indique la bonne réalisation du montage d'usinage. A ce stade, l'opérateur peut lancer l'opération de l'usinage en toute sécurité!

D'autres contraintes imposées par le cahier des charges doivent être prises en compte dans la stratégie menée pour la proposition de la solution. En effet, la so-

lution proposée doit être **rapide** pour respecter le gain de production dans le cycle de fabrication. En plus elle doit être à un **coût accessible** pour les PME de l'usinage. Raison pour laquelle nous avons privilégié une solution monoculaire, ou de même point de vue, pour pallier le problème d'encombrement au sein de la machine.

Le cahier des charges impose de prendre en compte trois paramètres principaux, à savoir **le temps de calcul, le coût du système et la simplicité**, et enfin, **l'encombrement** pour pouvoir installer le système de contrôle au sein de la machine.

1.4 Structuration du rapport

Rares sont les travaux de recherche menés sur cette problématique. Une étude bibliographique a permis d'identifier quelques travaux récents en liaison directe ou indirecte avec notre solution.

Les systèmes optiques sont déjà utilisés dans le domaine de l'industrie et leur progression montre l'intérêt de l'utilisation d'un tel système. L'apparition des capteurs moins chers donnant une carte de profondeur en temps réel de la scène, e.g. La KINECT[®], a fait orienter nos recherches au début de la première année de ce travail vers le test de quelques systèmes commerciaux de numérisation tridimensionnelles sans contact, dans l'objectif de faire une comparaison directe 3D/3D du montage d'usinage réalisé, e.g. numérisé et le modèle de référence 3D donné en CAO. Nous avons constaté que la reconstruction tridimensionnelle du montage d'usinage nécessite un temps important et qui ne correspond pas aux exigences du cahier des charges. Ce qui a orienté nos travaux sur une solution monoculaire. Dans ce cas et afin d'assurer une meilleure visibilité de la scène, nous avons proposé une configuration de la caméra par rapport au montage d'usinage. Le chapitre 2 présente le travail mené dans ce sens.

Notre objectif, maintenant, repose sur l'exploitation des contours, porteurs de l'information de la géométrie et de la forme. Ces contours trouvent leurs correspondants dans la représentation CAO par les arêtes des surfaces. Nous avons proposé une nouvelle méthode de calcul des contours à partir de la modélisation explicite de l'histogramme des modules de gradient. L'image des contours va permettre d'extraire les primitives pertinentes 2D permettant d'établir une mise en correspondance entre l'image et le modèle de référence CAO. Pour cela, nous avons créé une image théorique à partir du modèle CAO. Pour simplifier la mise en correspondance et optimiser le temps de calcul, nous représentons les primitives pertinentes par les graphes. Le détail de cette partie est présenté dans le chapitre 3.

Dans l'objectif de déterminer les régions d'intérêt représentant les emplacements *prévus* des éléments de bridage, nous projetons automatiquement le modèle CAO sur l'image du montage réalisé. Le problème revient, à trouver les correspondances entre les deux graphes : réel noté G donné par l'image de la caméra et théorique, noté H donné par l'image théorique. Nous avons proposé une nouvelle méthode s'appuyant sur l'information géométrique pour déterminer la mise en correspondance des graphes. Nous avons évalué le calcul de pose par rapport aux paramètres de la méthode proposée

et ainsi les paramètres photométriques de la scène. Le détail est présenté dans le chapitre 4.

Une fois la pose calculée et les régions d'intérêts localisées, nous nous intéressons à l'analyse locale de chaque région pour la prise de décision sur la conformité de cette région à l'élément de bridage attendu et par conséquent la conformité du montage d'usinage au modèle CAO. Les éléments de bridage ont souvent une forme polyédrique ; nous avons utilisé les squelettes pour décrire la forme de la région. La mise en correspondance entre les squelettes théorique et réel permet de quantifier l'appariement de la forme. Pour renforcer la décision nous exploitons les contours de l'objet en question ainsi que la dimension donnée par la superficie. Le chapitre 5 décrit le processus de l'analyse locale.

Dans certains cas, l'information 3D est indispensable, notamment pour répondre au cas "orange", nécessitant un test de la trajectoire par rapport à la représentation 3D de la scène. Pour réaliser une carte de profondeur dense de la scène en utilisant une solution monoculaire, un stage a été effectué à l'université de Sherbrooke, Québec, afin de travailler sur le sujet. Notre objectif consiste à estimer la carte de profondeur à partir d'une seule acquisition d'image couleur (RGB). Les premiers résultats sont encourageants et ils montrent la possibilité d'exploiter cette piste.

Enfin, nous concluons avec des perspectives et nous montrerons la limite de notre système de vision et les divers points nécessitant un travail approfondi.

Pour répondre aux besoins du cahier des charges et valider notre solution, nous avons testé la méthode sur différents montages d'usinage, présentant différentes formes possibles des éléments de bridage y compris le brut ou la pièce finale.

- Le montage sur la **plaque Norelem**[©] est un montage réalisé pour les besoins de cette thèse. La forme de la pièce à usiner présente des creux à différents endroits. Ces creux parfois sont en contact avec les éléments de bridage, i.e. la bride de fixation, ce qui présente un danger de collision dans le cas de déplacement de la bride. Ce montage est notre support principal dans ce travail.
- Le montage **Semelle** est constitué essentiellement par un brut "moulé" ayant déjà la forme de la pièce finale. Ce dernier est maintenu par deux brides et fixé par 5 pions. Ce montage a été réalisé pour percer un trou sur le brut afin de lui donner sa forme finale.
- Le montage **Minimanche** est un autre type de montage d'usinage dont le support de fixation est différent du support Norelem[©]. En effet, la diversité des formes de pièces demande au concepteur de concevoir des supports adaptables, assurant la fixation du brut. C'est le cas du montage Minimanche qui permet de réaliser un support de *joystick* pour la commande de vol sur des avions.

D'autre part, il est à noter que certaines pièces complexes doivent être réalisées en deux ou plusieurs phases d'usinage, à chaque étape d'usinage la pièce intermédiaire est considérée comme le brut pour la phase d'usinage en cours. La figure 1.6 illustre les trois montages.

Enfin, l'organigramme de la figure 1.7 illustre les chapitres du manuscrit et l'objectif de chaque partie.

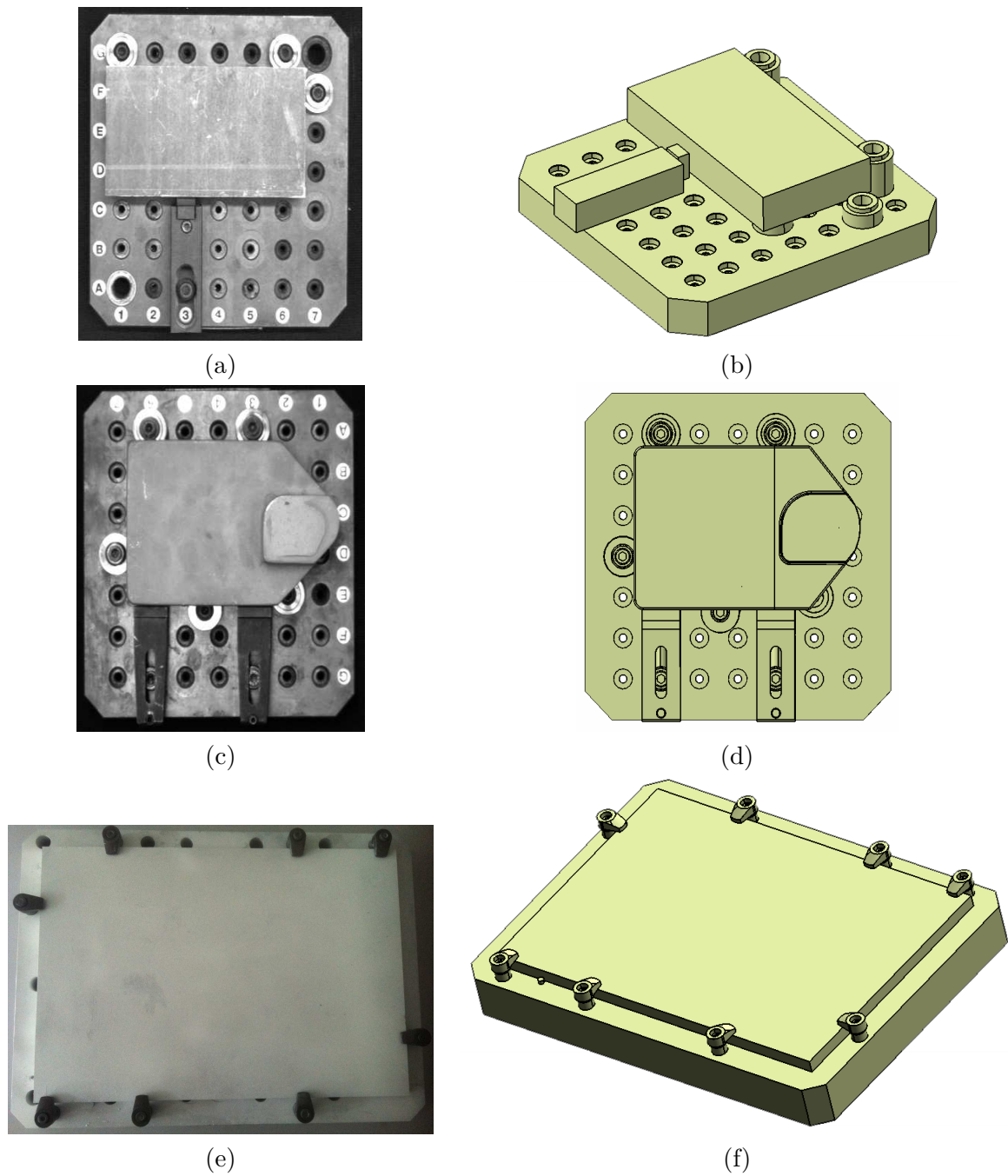


Figure 1.6 – Différents montages testés : (a) Montage sur plaque Norelem, (b) Modèle numérique correspondant, (c) Montage "Semelle", (d) Modèle numérique correspondant, (e) Montage "Minimanche" de réalisation de support commande de vol pour avion, (f) Modèle numérique correspondant.

Objectif :

Localisation des éléments de la scène

Prise de décision sur la conformité du montage réalisé

... pour répondre au cas "Orange"

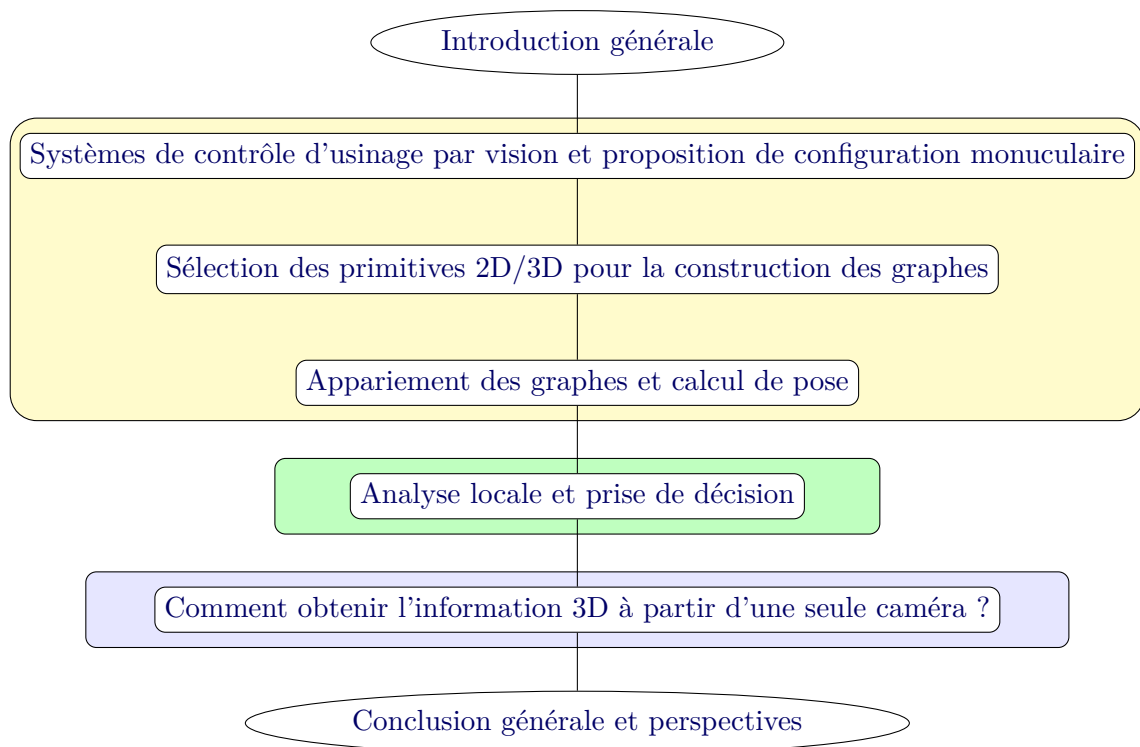


Figure 1.7 – Organigramme du rapport de la thèse et objectif de chaque partie.

Bibliographie

- [Caux 2012] M. Caux. Méthodologie de maîtrise des variations géométriques des produits en conception, fabrication et contrôle dans le contexte de l'usine numérique. PhD thesis, ENS CACHAN, 29 Juin 2012. (Cité en page 3.)
- [Pechard 2009] P.-Y. Pechard. Génération de trajectoires d'usinage grande vitesse 5 axes par flanc d'outil : intégration d'un critère de fluidité. PhD thesis, ENS CACHAN, 07 Juillet 2009. (Cité en page 4.)
- [Tournier 2010] C. Tournier et et coll. Usinage à Grande Vitesse. page 400, 2010. (Cité en page 4.)

SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE

"L'homme et sa sécurité doivent constituer la première préoccupation de toute aventure technologique". Albert Einstein (1879-1955)

Sommaire

2.1	Introduction	12
2.2	Solution actuelle avec contact	13
2.3	La vision pour la détection de collision de l'outil d'usinage	15
2.3.1	La détection de collisions de l'outil d'usinage	15
2.3.2	La reconstruction tridimensionnelle de l'outil	17
2.4	La reconstruction tridimensionnelle du montage d'usinage	19
2.5	Systèmes optiques testés	22
2.5.1	Critères conditionnant le choix de système de mesure 3D	23
2.5.2	Capteur Minolta	25
2.5.3	Capteur Viu-scan	26
2.5.4	Capteur Optinum	26
2.5.5	Capteur Kinect	26
2.5.6	Résultats de numérisation 3D du montage Norelem	28
2.6	Configuration du système de vision	29
2.7	Conclusion	31
	Bibliographie	32

Mots clés : Contrôle d'usinage, Systèmes avec contact, Systèmes sans contact, CFAO, Critères de sélection d'un système de numérisation 3D, Configuration proposée.

Chapitre 2. SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION 1 ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE

2.1 Introduction

Pour assurer la sécurité de la machine pendant la phase d'usinage, une assistance par ordinateur est nécessaire. En effet, chaque étape du traitement numérique de l'usinage : conception, génération de la trajectoire, calcul des commandes d'axes (CN), etc. est validée par un outil informatique de simulation numérique, CATIA[®], MasterCam[®], etc. , pour les plus utilisés en aéronautique dans notre environnement industriel, avant de transmettre les données à l'étape suivante : génération de la trajectoire et programmation de la CN, etc. Nous nous intéressons dans ce chapitre à quelques solutions actuelles de mesure avec contact, section 2.2, leurs avantages et inconvénients, ainsi que des solutions nouvelles sans contact utilisant la vision pour le contrôle et la mesure en usinage, section 2.4. Nous avons envisagé au début une solution 3D/3D pour la vérification du montage d'usinage. Raison pour laquelle nous avons numérisé le montage d'usinage avec quelques systèmes commerciaux de mesure tridimensionnelle sans contact. Nous présentons, dans la section 2.5 quelques critères de choix d'un tel système de numérisation 3D par rapport à l'application envisagée.

Enfin, pour des raisons d'incompatibilité des systèmes testés par rapport aux exigences du cahier des charges, nous avons proposé une solution monoculaire à partir d'une seule image. Nous présentons dans la section 2.6, la configuration du système de vision proposé au sein de la machine, pour une meilleure acquisition de la scène.

Le chapitre est structuré comme le montre l'organigramme de la figure 2.1.

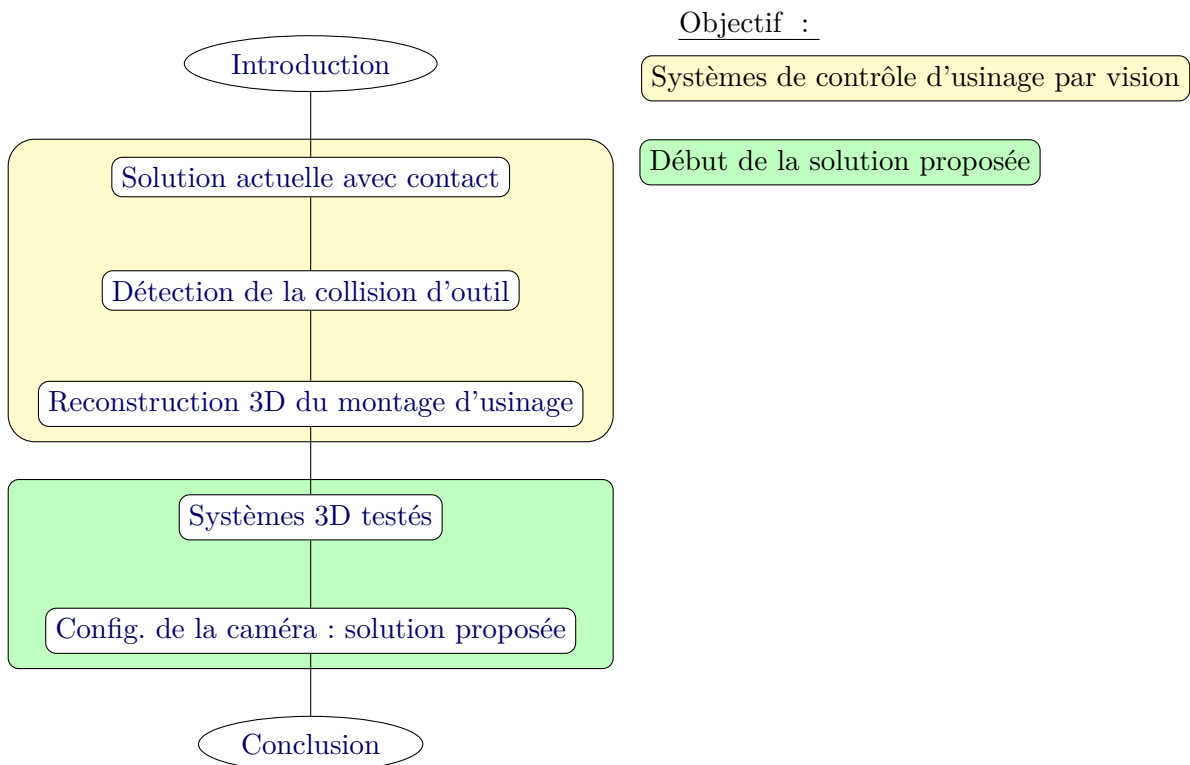


Figure 2.1 – Organigramme du chapitre 2 et objectif de chaque partie.

2.2 Solution actuelle avec contact

Le temps nécessaire pour la réalisation d'une pièce mécanique "finie" prend une place importante dans le cycle de production. Les mesures manuelles qui emploient principalement des comparateurs à cadran ne présentent pas une solution optimale, car elles dépendent des compétences de l'opérateur ou d'un transfert de la pièce sur une MMT¹, ce qui n'est pas toujours pratique dans la fabrication en série. A ce stade de fabrication, dans l'atelier, le contrôle du montage d'usinage se fait aujourd'hui par un palpeur [Peyrucat 2013], un exemple de bras robotisé de mesure 3D est présenté sur la figure 2.2.

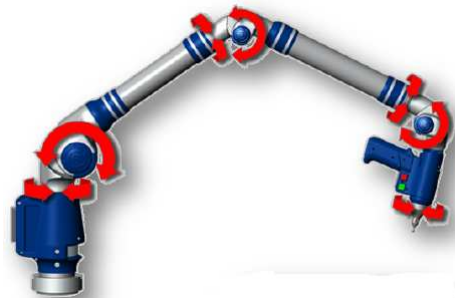


Figure 2.2 – Bras MMT FARO GAGE pour la mesure tridimensionnelle

La machine à mesure tridimensionnelle est caractérisée par une répétabilité allant jusqu'à 0,018 mm et des précisions de l'ordre de quelques *micromètres* suivant la norme **ISO 10360-2**, utilisée pour évaluer les MMT. Une MMT est composée essentiellement d'un palpeur à déclenchement. Ce dernier est constitué d'une tête de mesure-palpeur-stylet motorisé ou manuelle et d'une bille. Cet ensemble est régulièrement complété par un changeur de palpeur et/ou changeur de bille permettant de réaliser différentes configurations de palpation : choix de longueur de stylet (L), diamètre de bille ($2R$), etc. Comme le montre la figure 2.3. Le principe de fonctionnement d'un palpeur repose sur la génération d'un signal de mémorisation lors du contact entre la bille du palpeur et le point local de l'objet. Il existe plusieurs technologies, qui se distinguent par la technique mise en œuvre pour déclencher la mesure. Il peut s'agir d'une mesure de choc avec un capteur piézoélectrique ; d'une mesure de force en utilisant un palpeur à contact sec avec ressort calibré ; d'une mesure de déplacement avec un détecteur inductif, etc. Cette opération enregistre la position du palpeur par rapport au repère du système, bras manipulateur, MMT, etc. permettant des mesures répondant, généralement, aux besoins en précision. Cette précision de mesure dépend ainsi du degré de liberté et des performances du système de positionnement utilisé.

D'autres solutions récentes reposent sur l'utilisation d'un palpeur infrarouge (*IR*) monté sur la broche. Cette solution peut intervenir à la fois dans les mesures en cours de cycle permettant un usinage *adaptatif* et lors des premiers contrôles après usinage dans l'objectif de diminuer le temps d'arrêt de la machine pour le transfert de la pièce et les réglages manuels qui lui sont associés. Le principe de fonctionnement d'un palpeur

1. Machine à Mesure Tridimensionnelle

Chapitre 2. SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION 1 ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE

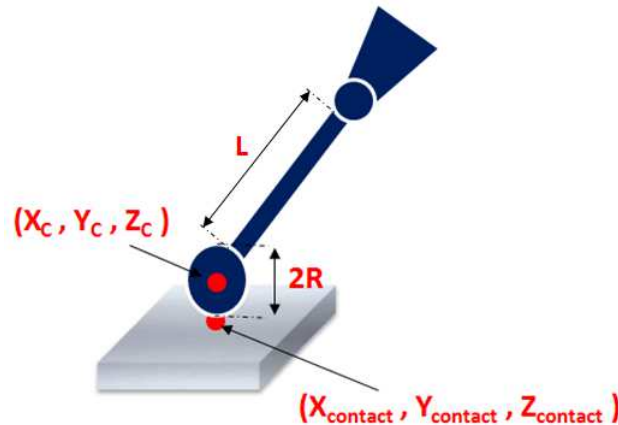


Figure 2.3 – Principe de fonctionnement d'un palpeur

infrarouge s'appuie sur la transmission d'un signal infrarouge au récepteur installé dans l'espace de travail de la machine, lors du contact du palpeur avec la surface de la pièce. La stabilité de transmission du signal est assurée à la fois par la fréquence de transmission utilisée, $75Hz$ pour éviter toute perturbation avec les signaux de 50 et $60Hz$ fréquemment utilisés dans les ateliers, et le débit du signal employé, ainsi que les angles assez larges de réception. Un exemple d'un palpeur infrarouge est illustré sur la figure 2.4.



Figure 2.4 – Palpeur à transmission IR monté sur la broche [Peyrucat 2013].

Cependant, ces solutions avec contact permettent de relever les mesures tridimensionnelles de certaines zones pertinentes du montage d'usinage sans donner une information complète sur sa conformité en regard du parcours de l'outil. Le contrôle est réalisé par quelques dizaines de points de mesure. Il est facilement compréhensible que ce procédé peut être "leurré" par des objets non prévus et donc non testés même si les points relevés sont des points jugés pertinents par rapport à la trajectoire de l'outil. De plus, ils sont d'un coût relativement élevé pour les PME de l'usinage.

Des travaux plus récents portant sur des solutions sans contact, par vision, ont été identifiés dans la littérature. Ils montrent l'intérêt de l'utilisation d'un tel système de vision pour le contrôle de l'usinage.

Afin de situer nos travaux, nous proposons de présenter quelques systèmes utilisant la vision pour le suivi et l'optimisation de la trajectoire, section 2.3.1, la mesure tridimensionnelle de l'outil d'usinage, section 2.3.2 et la reconstruction tridimensionnelle du montage d'usinage, section 2.4.

2.3 La vision pour la détection de collision de l'outil d'usinage

Nous nous intéressons dans cette partie aux travaux qui portent sur le suivi et l'optimisation de la trajectoire de l'outil, ainsi qu'à la modélisation tridimensionnelle de ce dernier pour la détection de collision.

2.3.1 La détection de collisions de l'outil d'usinage

Dans les travaux de R. AHMAD *et al.* [Ahmad 2010], les auteurs ont proposé une solution de détection et de localisation de la collision dans l'espace 2D de la caméra pour une trajectoire rectiligne entre le point de départ de l'outil et la cible (le brut à usiner). Ils proposent un algorithme de SET²/RESET³ pour le calcul et l'optimisation de la nouvelle trajectoire. L'algorithme proposé est structuré en 3 étapes, comme le montre la figure 2.5.

- **Détection d'objet :**

Dans cette étape, la détection de l'objet est effectuée par l'intermédiaire d'une opération de binarisation en utilisant le *package* du logiciel Maxtor Inspector[®] (Maxtor Electronic Systems Ltd).

- **Identification de l'objet :**

Avec une connaissance *a priori* des éléments dans la scène théorique, les objets dans la scène réelle sont identifiés dans la même représentation spatiale que la scène théorique par comparaison directe. Le processus d'interprétation de la trajectoire (code G) peut être utilisé comme donnée pour le système de vision. Cela nécessite une étape de calibration de la caméra dans l'enceinte de la machine. La position de l'outil dans l'espace 2D de l'image peut être utilisée comme donnée pour l'algorithme RESET, pour la génération de la nouvelle trajectoire en présence d'obstacles.

- **Optimisation de la trajectoire de l'outil, algorithme SET/RESET :**

Le processus commence par une binarisation de l'image réelle et théorique. Les images binaires résultantes représentent l'espace d'interprétation de la trajectoire, tel que le fond soit représenté par '0' (noir) et l'objet '1' (blanc). Ensuite, l'algorithme procède initialement, à la détermination du chemin le plus court (cas d'un trajet en droite) entre le point de départ (position actuelle de l'outil) et le point d'arrivée (point de production, brut). Si cette trajectoire est marquée par la présence d'un obstacle à l'endroit (x_o, y_o) dans l'image, le processus

2. SET : Safe and Efficient Trajectory

3. RESET : Rectangular Enveloped object Safe and Efficient Trajectory

Chapitre 2. SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE

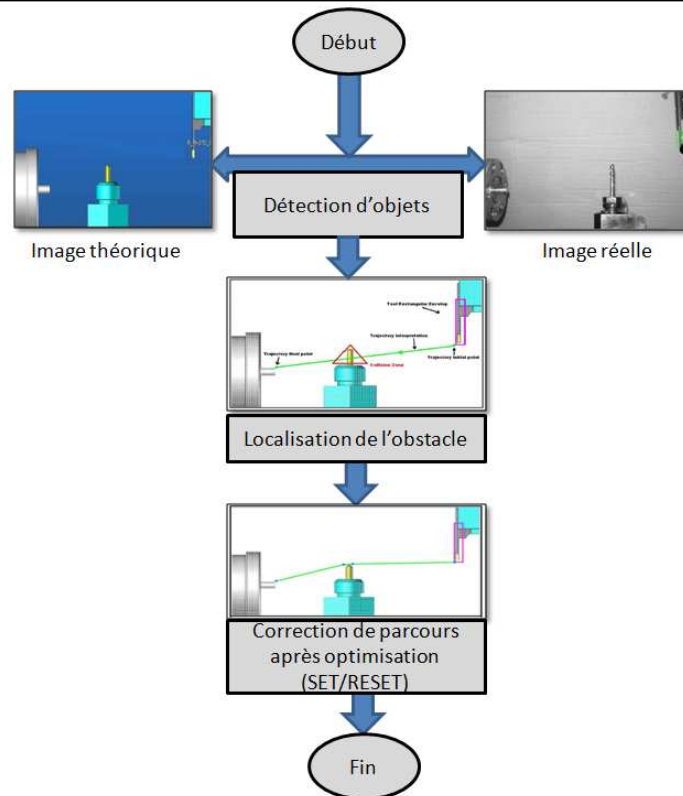


Figure 2.5 – *Algorithme proposé par [Ahmad 2010] pour la détection et correction de la trajectoire de l'outil*

fait appel à un algorithme de détection des coins autour de l'obstacle pour la recherche des coins appartenant au fond dans ce voisinage du point critique. Le principe de l'algorithme de détection des coins, est présenté sur la figure 2.6. A partir de ces coins, l'algorithme RESET est utilisé pour calculer la nouvelle trajectoire.

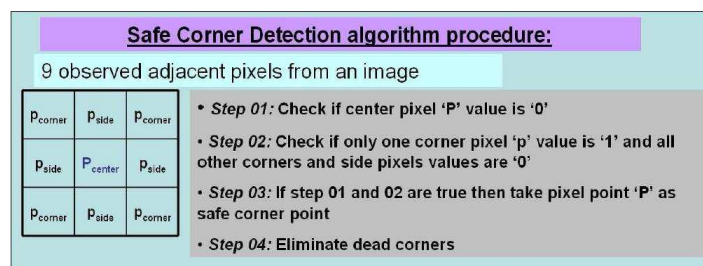


Figure 2.6 – *Principe de l'algorithme de détection des coins*

L'algorithme RESET, est une extension de l'algorithme SET qui tient compte de l'espace de définition de l'outil d'usinage, il propose une correction de la

trajectoire en tenant compte des coordonnées des coins de l'enveloppe convexe de l'outil utilisé.

Cette partie de modélisation tridimensionnelle de l'outil d'usinage a fait l'objet des travaux de X. ZHANG *et al.* [Zhang 2011] sur la reconstruction 3D de l'outil.

2.3.2 La reconstruction tridimensionnelle de l'outil

La modélisation tridimensionnelle pour le calcul de la trajectoire de l'outil est une donnée importante pour la génération sans collision des commandes numériques d'axes. Il est préférable de simuler la trajectoire proposée par le programme FAO, pour détecter les éventuelles collisions, avec l'ensemble outil et porte outil dans l'environnement théorique avant de transférer le programme à l'atelier. La figure 2.7 montre le modèle 3D de l'outil et de son support, ainsi qu'un exemple de collision dû à un mauvais outil, de dimensions ou de forme inadaptées.

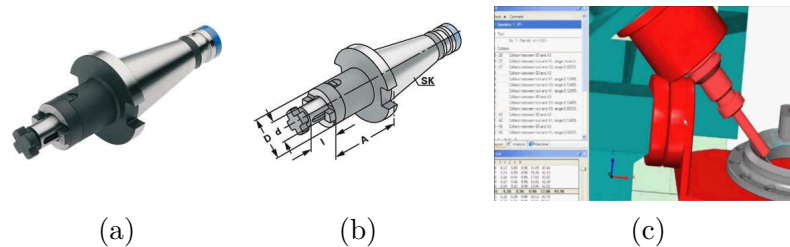


Figure 2.7 – Dimensionnement de l'outil d'usinage, (a) : porte outil, (b) : modèle 3D de porte outil, (c) : cas de collision dû au mauvais dimensionnement de l'outil (zone rouge)

X. ZHANG *et al.* [Zhang 2007b] ont proposé une solution sans contact reposant sur l'utilisation d'une caméra pour la modélisation 3D de l'ensemble de l'outil et du support. Dans un autre travail complémentaire [Zhang 2008], les auteurs proposent une estimation de pose (rotation et translation) de l'outil avec une reconstruction 3D de ce dernier dans l'objectif de suivre les mouvements de l'outil et de prédire les collisions. Le système est composé d'une caméra placée en face de l'outil et devant une source de lumière. La source de lumière a pour fonction de remédier aux problèmes de flou, permettant une fermeture de l'iris donnant une augmentation de la profondeur de champ, pour l'extraction des contours. La figure 2.8 illustre l'expérimentation réalisée.

Le programme de la reconstruction tridimensionnelle procède en 2 étapes :

- Premièrement, les paramètres intrinsèques et extrinsèques dans l'espace de la machine sont déterminés à partir d'une étape de calibration hors ligne. La position relative de l'outil par rapport à la caméra est identifiée.
- Deuxièmement, les contours de l'outil sont calculés à partir de l'image réelle, et chaque contour est divisé en plusieurs segments basés sur l'analyse de la courbure, en utilisant l'algorithme *Digital Straight Segment (DSS)* [Reinhard 2004] avec le concept de *FMBR*⁴.

4. Forward Minimal Bounding Rectangle

Chapitre 2. SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE

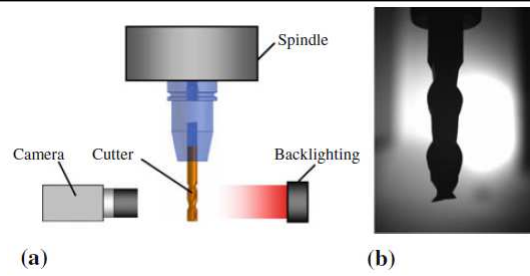


Figure 2.8 – L'effet de la source lumineuse dans l'extraction des contours de l'outil [Zhang 2011]

Le principe *FMBR* permet de déterminer le rectangle englobant le contour dans sa forme rectiligne. Un rectangle est défini par sa largeur, fixée à 2 pixels, sa longueur définie par la taille du contours et son orientation θ fixée comme étant l'angle entre la direction du contour et le centre de gravité. La figure 2.9 montre les paramètres constituant le principe *FMBR* pour deux point P_i et P_j .

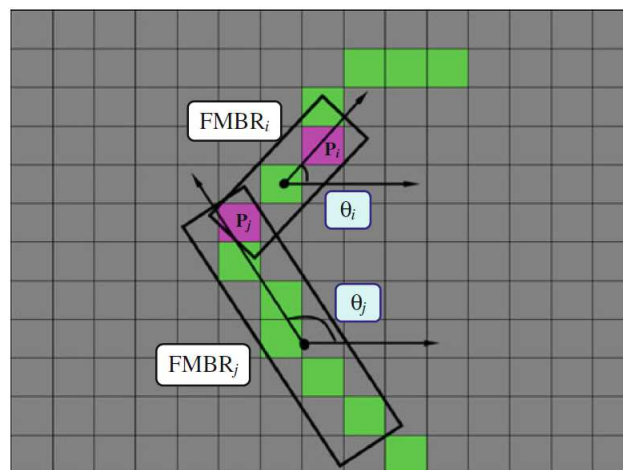


Figure 2.9 – Principe de *FMBR* X. ZHANG et al. [Zhang 2011]

Chaque segment est classifié en contour apparent ou contour invisible (occulté) par la caméra. Enfin, par symétrie, le modèle 3D de l'outil est reconstruit par remplissage de la section occultée par une fonction de génération qui est un chaînage entre les points terminaux en utilisant la surface de révolution (*SOR*)⁵[Colombo 2004]. Le résultat de cette reconstruction 3D est illustré sur la figure 2.10.

Les deux solutions évoquées précédemment dans la section 2.3.1 et 2.3.2, peuvent être classées dans la phase de génération de trajectoire comme une étape de vérification de la trajectoire de l'outil par simulation en ligne.

D'autres travaux ont porté leur intérêt sur la reconstruction tridimensionnelle de

5. La surface de révolution est une surface dans l'espace euclidien, créée par une rotation d'une courbe plane de la forme de l'objet, autour d'un axe dans le plan de cette dernière.

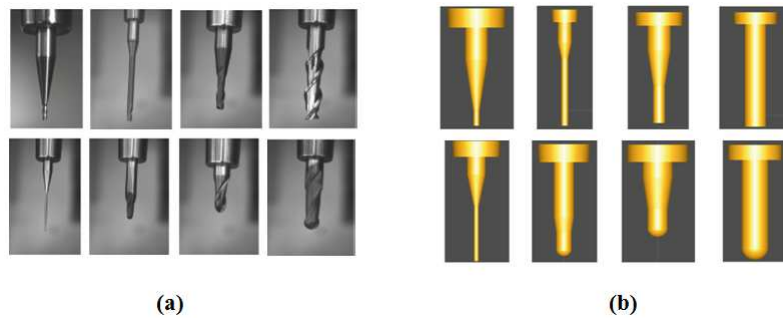


Figure 2.10 – Résultat de la reconstruction 3D de l'outil d'usinage [Zhang 2011], (a) : outils utilisés dans l'expérimentation, (b) : résultat de reconstruction 3D

l'ensemble du montage d'usinage, travaux présentés par X. TIAN *et al.* [Tian 2010]. La partie suivante présente le seul travail identifié dans la littérature qui soit en liaison directe avec notre proposition de solution.

2.4 La reconstruction tridimensionnelle du montage d'usinage

Le système proposé par les auteurs [Tian 2010] repose sur l'utilisation de la vision binoculaire pour la reconstruction du montage d'usinage. La figure 2.11 montre l'installation prototype du système au sein de la machine d'usinage.

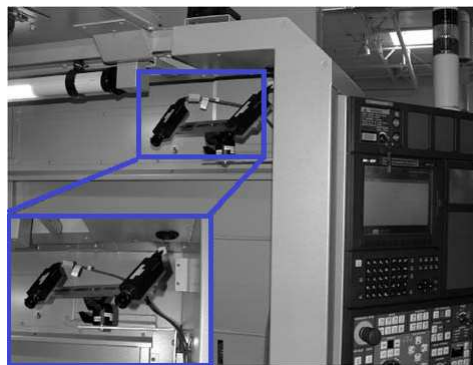


Figure 2.11 – Installation du système de vision [Tian 2010]

L'objectif est de reconstruire le montage d'usinage réel (réalisé par l'opérateur) à partir du modèle théorique et des éléments solides 3D (*Module Library*) de bridage définis dans une bibliothèque d'éléments solides 3D, comme le montre la figure 2.12 résumant le principe de l'algorithme proposé.

Cette reconstruction 3D s'appuie sur l'appariement des droites issues à la fois de l'image théorique et de l'image réelle.

L'algorithme est structuré en 3 parties.

- **Création de l'image théorique de la scène :**

Chapitre 2. SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE

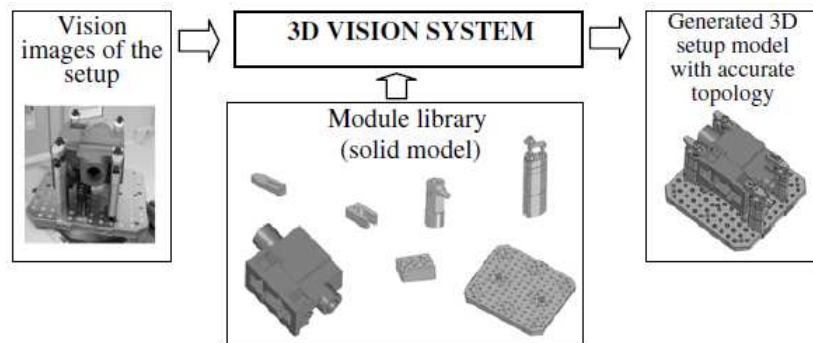


Figure 2.12 – Algorithme proposé par [Tian 2010]

La scène théorique est créée en projetant le modèle CAO 3D du montage d'usinage sur le plan (uv) de la caméra virtuelle [Kutulakos 1998]. Cette projection est donnée par les paramètres extrinsèques du système de vision, ce qui permet de garantir une **similarité géométrique** de la scène entre l'image réelle et théorique. Il est difficile de donner une luminance et une texture ressemblant à la réalité compte tenu des variations d'éclairage et de la nature des matières. Donc, seuls les éléments géométriques peuvent servir à la comparaison. L'image théorique est utilisée comme image de fond (*background*), et l'image réelle pour l'objet (*foreground*) dans l'étape d'extraction des contours.

- **Extraction des contours :**

L'extraction des contours du montage d'usinage qui s'opère sur des régions d'intérêt (ROI)⁶ se résume, dans un premier temps, à la recherche des coins en utilisant le détecteur de Harris. Toutes les coordonnées 2D des coins détectés sont enregistrées dans une matrice C de taille $(2 \times N)$ avec $C_i = (x_i, y_i)$ les coordonnées pour chaque point i . Deuxièmement, une opération de filtrage est appliquée pour éliminer les coins issus de bruit (éclairage, reflet, etc.) ou du fond. Pour cela, l'image théorique est utilisée comme support de comparaison directe. Si les coordonnées 2D du coin (x_i, y_i) extrait de l'image réelle appartiennent au contour dans l'image théorique, ce coin est classifié au fond, il est éliminé, sinon il appartient à l'objet et il est enregistré dans la nouvelle matrice $C_j = (x_j, y_j)$, $j = 1 \dots M$, avec $M \leq N$.

Ensuite, les coins **isolés** sont considérés comme des coins venant du bruit, ils sont supprimés. Le processus de validation du contour E_{ij} est défini comme suit :

- 1- La direction θ et la taille du contour L_{ij} supposé sont calculés.
- 2- Pour chaque pixel de E_{ij} , on détermine les directions des gradients des 8 pixels voisins. Pour cette région le gradient maximum dans la direction θ est calculé.
- 3- Si ce maximum local du gradient de ce pixel est supérieur à un seuil (T), fixé comme étant la valeur moyenne du gradient, ce pixel appartient au contour et le compteur C_p est incrémenté, ainsi si un autre pixel dans ce voisinage a une

6. ROI : Region Of Interest

valeur de gradient supérieure à (T) le compteur est aussi incrémenté. Tout ces pixels sont considérés comme des pixels de contour.

4- Si le ratio du compteur Cp par rapport à la taille du contour L_{ij} est supérieur à 80% l'hypothèse de contour est validée. Sinon, le contour n'est pas pris en compte.

5- Dans le cas où plusieurs contours passent par le même coin $C(x_i, y_i)$ et selon la même direction θ , ces contours droits sont remplacés par un seul contour en résolvant l'ensemble des équations des droites au sens des moindres carrées.

- **Établissement de correspondance entre les contours dans les deux images droite et gauche :**

Afin de reconstruire le modèle 3D de l'objet ou d'une partie de l'objet à partir des modèles solides 3D des éléments de bridage, l'algorithme procède par l'établissement d'une correspondance entre les segments extraits de l'image droite et ceux de l'image gauche. C'est le principe d'appariement de la vision binoculaire. Pour cela, les auteurs utilisent l'algorithme de *block matching* [Faugeras 1993], [Tian 2002] qui estime une carte de disparité entre les images droite et gauche, en commençant par les plus longs segments, ceux ayant le plus de possibilité de correspondance. Une étape d'optimisation de la distance entre le segment de la scène et son correspondant 3D dans la bibliothèque des éléments solides est établie en minimisant l'expression donnée par l'équation (2.1) par la méthode du quaternion.

$$Minimise(Dist) = \min\left(\sum_{i=1}^k (Dist(StartP_{E_s}, E_m) + Dist(EndP_{E_s}, E_m))\right) \quad (2.1)$$

Avec k le nombre total de segments appariés, $StartP_{E_s}$ est le point de départ du contour E_s et $EndP_{E_s}$ est son point d'arrivée.

Cette méthode a permis de reconstruire le modèle réel à partir des éléments 3D définis dans le modèle CAO avec une précision allant jusqu'à 1mm pour une image de 1024×768 . La figure 2.13 montre le résultat de reconstruction.

L'adoption de l'hypothèse reposant sur l'extraction des segments limite l'application de cette méthode aux objets polyédriques. De plus, la solution proposée par les auteurs demeure un outil de reconstruction du montage d'usinage théorique à partir du réel, mais ne tient pas compte des mauvaises configurations qui peuvent avoir été réalisées. Notre proposition de solution englobe plus de formes géométriques et a comme objectif de valider le montage d'usinage vis-à-vis le modèle CAO. Pour des raisons initiales, évoquées en introduction générale, sur la nécessité d'intégration de l'information tridimensionnelle pour l'analyse du montage, nous avons testé quelques systèmes optiques de numérisation. La section 2.5 décrit les systèmes testés, leurs inconvénients et avantages par rapport à notre application industrielle.

Chapitre 2. SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION 2ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE

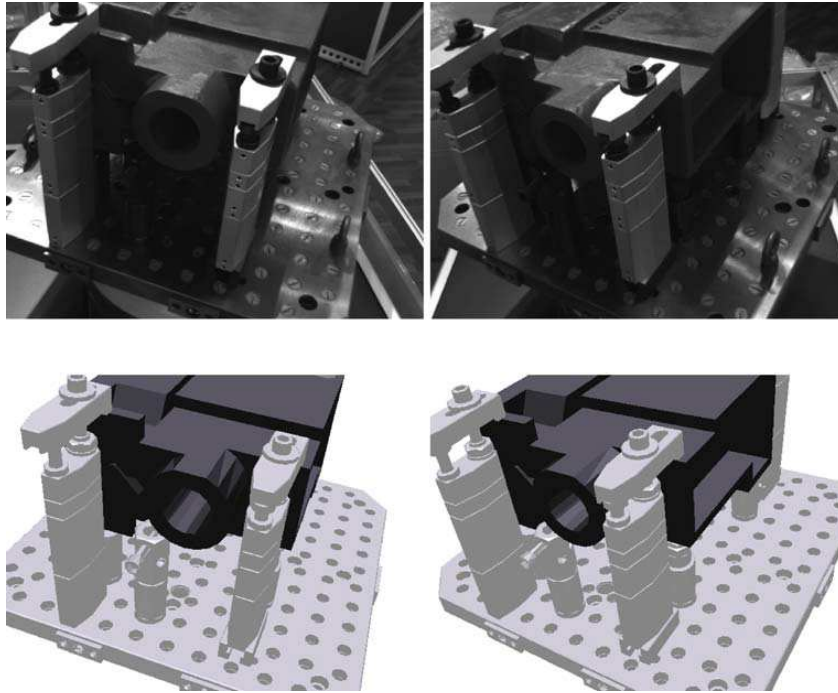


Figure 2.13 – Reconstruction 3D théorique du montage d'usinage

2.5 Systèmes optiques testés

Posséder une représentation 3D du montage d'usinage pourrait permettre d'enrichir davantage l'analyse de forme et géométrique des éléments de bridage constituant le montage d'usinage. Cette représentation peut prendre la forme brute d'un nuage de point 3D ou les formes symboliques, plan, cylindre, polyèdre, etc. pour autoriser une comparaison numérique ou symbolique du montage d'usinage par rapport au montage théorique. Au cours de ces dernières années, les systèmes de mesure sans contact utilisant une caméra ont fortement progressé, notamment dans le secteur industriel. Ils s'implantent de plus en plus dans la surveillance, la mesure et la rétro-conception, etc. Parmi cette famille de systèmes d'acquisition, on trouve les systèmes optiques de mesure tridimensionnelle. Cette section présente les systèmes commerciaux testés, e.g. Minolta[®], Optinum[®], Vii-scan[®] et la Kinect[®], pour la mesure tridimensionnelle et leur classification selon la technologie utilisée par rapport aux critères de coût, de temps de calcul et de précision de mesure, etc.

Définir les besoins en numérisation, en terme de rapidité d'acquisition, précision, facilité d'utilisation, etc. , par rapport à l'application envisagée, le gain de production et les conditions photométriques des milieux de mesure, etc. induit un choix de la technologie d'acquisition tridimensionnelle à adopter.

2.5.1 Critères conditionnant le choix de système de mesure 3D

Nous avons testé la famille des systèmes optiques pour leur rapidité d'acquisition. La taille moyenne des montages d'usinage à numériser varie entre 200 et 1000 mm, ce qui conforte notre sélection. De plus la précision, ici relativement moins importante ($\leq 1mm$) du fait que notre but est d'intégrer l'information 3D pour la validation du montage d'usinage. La figure 2.14 montre un classement possible de ces différentes techniques de numérisation tridimensionnelles sans contact selon les critères donnés par le tableau 2.1.

Tableau 2.1 – Quelques critères d'évaluation des systèmes de numérisation 3D. Par exemple en utilisant une notation entre ★ et ★★★★★ (moins important et très important)

Critère d'évaluation	Minolta vi-9i	ViuScan	Optinum	Kinect
Réflectivité	★	★	★	★★
Flexibilité	★	★★★	★★★	★★
Portabilité	★	★★★★	★★★★	★★★
Volume (taille de l'objet)	Petit à Grand	Petit à Moyen	Petit à Moyen	Moyen à Grand
Résolution	400 μm	0,100 mm	0.5mm	9.3 μm
Précision	50 μm	50 μm	100 μm	$f(Z^2)$ ⁷
Temps de calcul	★	★	★★	★★★★
Coût	★	★	★★	★★★★★
Autres

- **La réflectivité des objets** est définie comme la possibilité de mesurer un spectre large de réflectivité, de surfaces foncées jusqu'aux surfaces pâles, de surfaces mates jusqu'aux surfaces polies. Pour plusieurs technologie optiques, la transparence/ reflet du matériel demeure un problème majeur. Cependant, il est possible de remédier à ce problème de reflet spéculaire en appliquant un produit de ressuage sur la surface, e.g. cire, poudre blanche, etc. Par contre, ce traitement ajoute une épaisseur non uniforme à l'objet, affectant l'incertitude de mesure qui peut être importante pour certaine applications en médecine, technologies de précision, etc. , par exemple.
- **La flexibilité** est la capacité du capteur à accéder à des endroits profonds de l'objet à numériser. Selon l'application envisagée, la flexibilité peut être évaluée selon les critères suivants :
 - *La possibilité de mesurer des points dans le fond de trous*
 - *La facilité à mesurer des données selon différentes prises de vues*
 - *La profondeur de champs, etc.*

7. Proportionnelle à la distance au carrée de la surface a numérisée [Kourosh 2012], avec $Z \in]50cm, 500cm[$

Chapitre 2. SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE

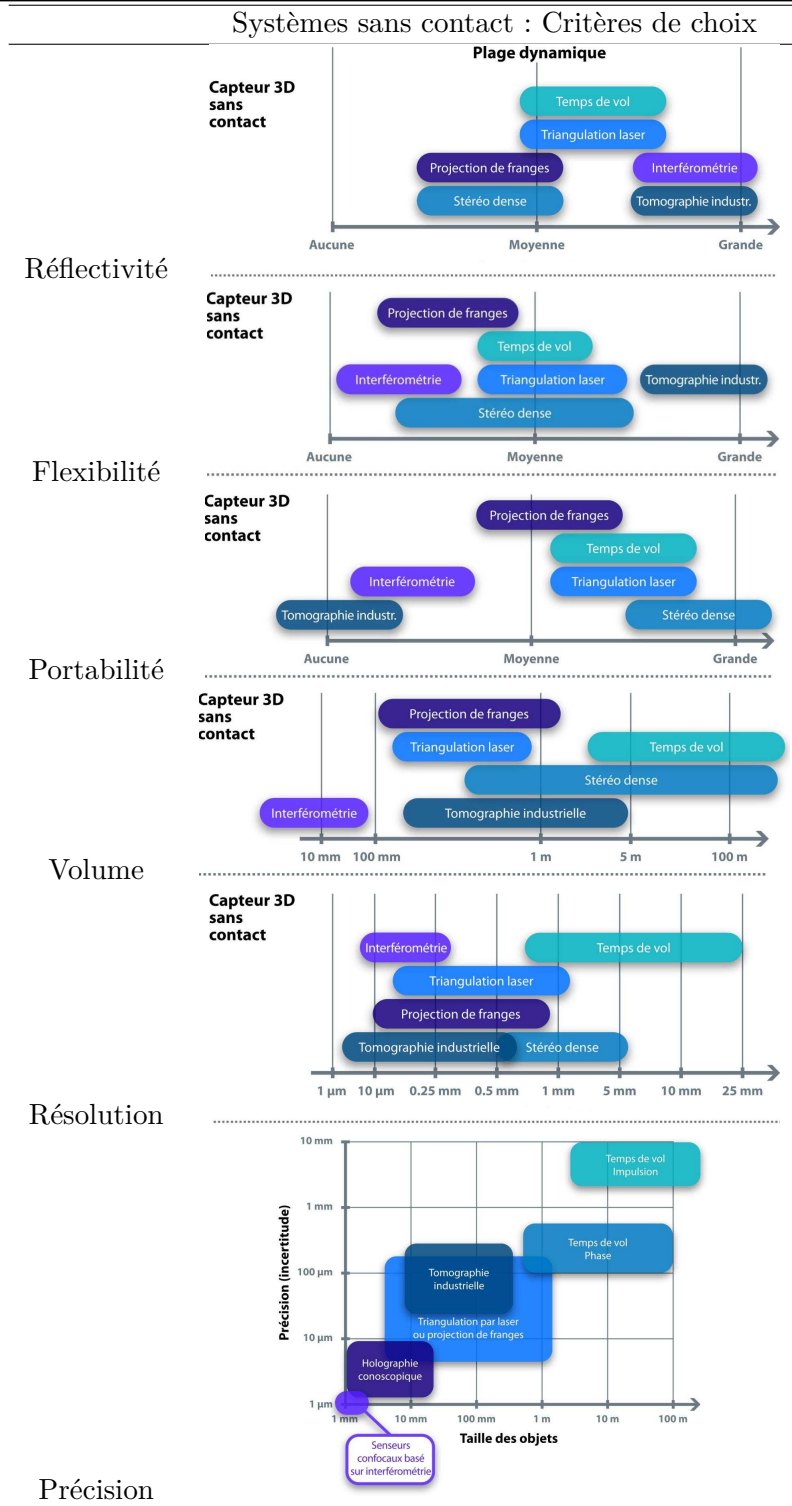


Figure 2.14 – Critères de choix de système de numérisation 3D (www.innovmetric.com)

- **La portabilité** concerne la plus au moins grande facilité à déplacer le capteur de mesure au lieu de déplacer l'objet à mesurer. Ce critère doit être considéré pour évaluer le temps d'acquisition.
- **Le volume** réfère à la taille de l'objet pouvant être mesuré. Le volume mesurable dépend plus souvent de la méthode de référence que de la technologie utilisée du capteur lui-même, e.g. triangulation, corrélation, etc.
- **La résolution** peut être groupée en deux parties : la résolution latérale, qui représente l'espace latéral entre les points. La résolution en profondeur qui représente la plus petite différence de profondeur mesurable. Les deux résolutions sont critiques pour mesurer les surfaces, les trous et les bordures de la pièce avec "fidélité".
- **Le précision** est définie comme la capacité d'un capteur à mesurer la pièce le plus fidèlement possible au sens de l'incertitude de mesure. Le mot incertitude est souvent utilisé pour décrire la précision.

2.5.2 Capteur Minolta

Le système Minolta présenté dans la figure 2.15.a, est un système de mesure sans contact qui s'appuie sur l'utilisation de la technologie de triangulation par méthode des sections lumineuses. Il permet de numériser la forme du montage d'usinage par *profilométrie* : projection d'une ligne laser à balayage. Il est souvent utilisé avec une table de rotation - figure 2.15.b - dont l'ensemble est commandé par un logiciel de numérisation dédié, afin d'aligner facilement et automatiquement les numérisations réalisées dans différentes prises de vues. En plus, le volume de ce capteur, ne permet pas de l'installer au sein de la machine.

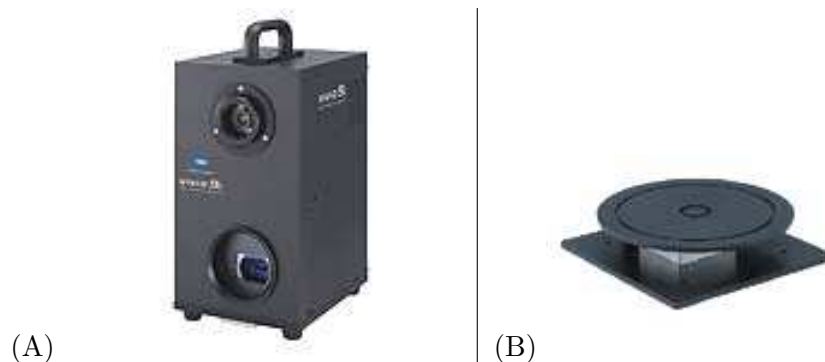


Figure 2.15 – Système Minolta[®], (A) : capteur Minolta vi-9i, (B) : Table de rotation

La précision de mesure peut aller jusqu'au $4\mu\text{m}$ pour une distance de mesure de 450mm. Cependant, la flexibilité reste faible sur certaines zones présentant des creux, dans le montage d'usinage. Ce qui fait que l'acquisition du montage nécessite un temps *mort* pour le changement de point de vue et l'orientation du montage par rapport au capteur.

Chapitre 2. SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION 2.5.3 Capteur Vii-scan

Le système Vii-scan[®] illustré sur la figure 2.16 est un système de mesure sans contact portable basé sur la triangulation entre un faisceau laser croisé projeté sur l'objet et l'image vue par les caméras (gauche, droite). Ce système nécessite une étape de calibration hors-ligne avant toute opération de numérisation. L'acquisition se fait à l'aide des marqueurs collés sur l'objet dans différentes surfaces et à différents endroits assurant au moins un repérage de 4 pastilles à chaque séquence de numérisation. La précision de ce système peut aller jusqu'à $50\mu\text{m}$ pour une profondeur de champs de 300mm . Cependant, le temps nécessaire pour la numérisation, y compris le pré et post-traitement, ne respecte pas les exigences du cahier des charges en terme du gain de production.



Figure 2.16 – Système Vii-scan[®] de Creaform

2.5.4 Capteur Optinum

Le capteur Optinum[®] montré sur la figure 2.17 est un système de mesure sans contact portable reposant sur la corrélation d'images issues des deux caméras (droite et gauche) avec éclairage par un mouchetis pour la reconstruction tridimensionnelle de l'objet. La précision du système annoncée peut aller jusqu'à $100\mu\text{m}$ pour une distance de prise de vue de 400mm . Le système ne nécessite pas une étape de calibration hors-ligne et le recalage des prises de vues se fait sans l'utilisation des marqueurs, pastilles autocollantes. Cependant, l'inconvénient majeur par rapport à notre application est le temps nécessaire pour l'opération de numérisation.

2.5.5 Capteur Kinect

La Kinect[®] de Microsoft est un capteur destiné à la console de jeux Xbox360, permettant de contrôler les jeux vidéos avec les gestes du corps humain, sans l'utilisation de la manette : par l'utilisation du modèle squelette humain. Le capteur est constitué d'une caméra infrarouge (IR) pour la mesure de déformation entre le mouchetis projeté en IR et sa forme caractéristique "initiale". Il est composé aussi d'une caméra couleur non utilisée pour l'estimation de la profondeur, mais qui peut servir pour la création de la texture dans l'image de profondeur. Le capteur et son principe de fonctionnement



Figure 2.17 – Système Optinum[®] Noomeo

sont montrés sur la figure 2.18. La précision de la Kinect[®] dépend de la mesure de disparité. Cette dernière est calculée en fonction de la distance au carré du capteur par rapport à l'objet [Kourosh 2012]. La Kinect[®] répond aux critères de choix pour notre application, à savoir le coût, la rapidité d'acquisition, la précision et, en particulier, la possibilité de représenter la mesure tridimensionnelle, image RGBD, dans la même résolution que l'image couleur, image RGB. Cependant, le champ de vision est assez large par rapport à la taille de l'objet, ce qui réduit la résolution du montage d'usinage dans l'image et par conséquent, peut contribuer à cacher quelques détails qui pourraient être importants dans la prise de décision sur la conformité du montage d'usinage.

Nous avons testé ces systèmes de numérisation tridimensionnelles pour le montage sur la plaque Norelem[®]. Les résultats de numérisation sont présentés dans la section 2.5.6 qui suit.

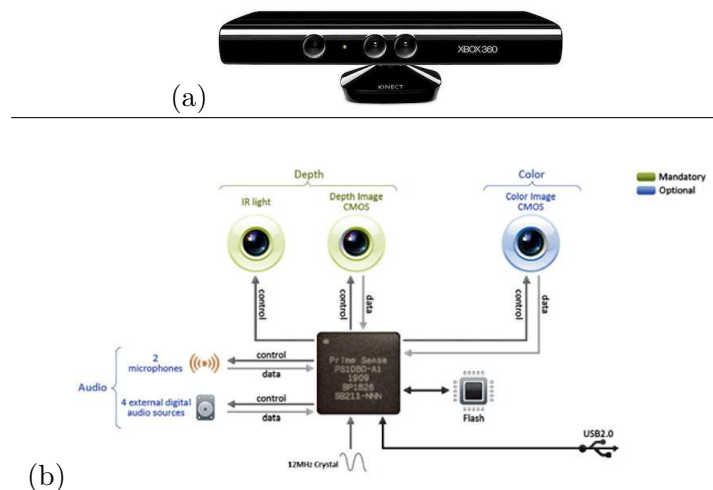


Figure 2.18 – (A) : Capteur Kinect[®] de Microsoft, (B) : Schéma de principe de fonctionnement comme est donné par Microsoft, les capteurs en bleu sont des composants facultatifs pour l'estimation de la profondeur de la scène

Chapitre 2. SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE

2.5.6 Résultats de numérisation 3D du montage Norelem

Nous utilisons dans cette partie le montage d'usinage de la plaque Norelem® présenté dans la figure 2.19. La qualité de numérisation est relativement acceptable, par contre, cette opération nécessite un pré et un post-traitement (pour la recalage, filtrage de nuage des points 3D, etc.) consommant un temps important pour obtenir un tel résultat.

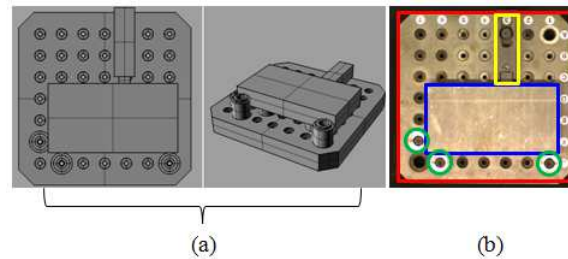


Figure 2.19 – (a) : Modèle 3D du montage d'usinage, (b) : Montage d'usinage composé d'une : palette (rouge), pions de positionnement (vert), brut (bleu) et bride de fixation (jaune)

Ainsi la préparation de la scène par ressuyage et contrôle d'éclairage représente un coût prohibitif et ils ne sont pas toujours maîtrisés dans les conditions photométriques de l'atelier. Seul le capteur Kinect® fournit une image de profondeur de la scène à la cadence vidéo et dans la même résolution que l'image de chrominance (RGB) permettant une corrélation 2D/3D dans l'espace réel pour la reconnaissance de forme par exemple sans mettre en œuvre des techniques de fusion des données.

Ces dispositifs restent encombrants, l'habitude en production de pièces mécaniques est de garantir un espace de manipulation au sein de la machine pour l'assemblage. Les systèmes possédant un faible encombrement sont privilégiés. C'est pour cette raison que nous nous sommes intéressés à des **techniques monoculaires**, notamment par exploitation du flou optique (Depth From Defocus) [Deschenes 2003], [Paolo 2007], [Zhuo 2011], [Karabagli 2013] que nous abordons ultérieurement.

Les différents systèmes testés ne donnent pas satisfaction selon les critères de coût, rapidité, précision et représentation 3D de la scène. Cette information tridimensionnelle est utile pour dimensionner les objets et faciliter la segmentation, mais n'est pas toujours nécessaire pour la prise de décision sur la conformité du montage d'usinage, comme nous le montrons dans le chapitre 5. L'exploitation de la luminance pour extraire la forme de l'objet dans l'image 2D de ce dernier peut être utile et suffisante.

Pour répondre aux exigences du cahier des charges en terme de rapidité d'exécution, installation du système au sein de la machine et problème d'encombrement et la précision de mesure. La solution proposée privilégie un **système monoculaire**. Le choix de la configuration de la caméra par rapport au montage d'usinage permettant d'assurer une meilleure acquisition de la scène à partir d'une **seule acquisition d'image, one shoot**, est une étape importante. Cette partie est présentée dans la section 2.6.

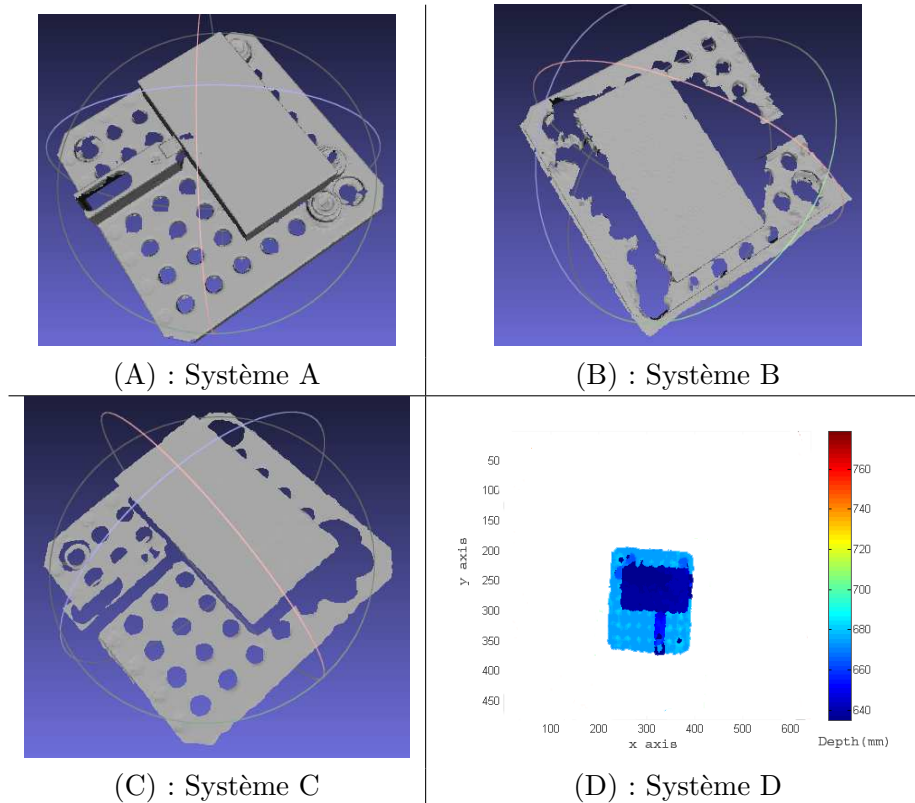


Figure 2.20 – Résultats de numérisation 3D du montage Norelem (cas conforme)

2.6 Configuration du système de vision

La finalité de ce travail vise la réalisation d'un capteur de vision intégrant acquisition, traitement et décision capable de traiter à la cadence de production la vérification de la conformité des montages d'usinage avant de lancer l'opération d'usinage proprement dite. Ceci nécessite des algorithmes simples, rapides et robustes utilisant des éléments significatifs de mise en correspondance, entre l'image réelle et le modèle 3D défini en CAO. Pour répondre à cette finalité, nous proposons un placement au-dessus, de la caméra, par rapport au montage d'usinage permettant de garantir une meilleure acquisition de la scène à partir d'une seule image, ce qui peut réduire aussi les traitements de la mise en correspondance. Nous avons envisagé dans un premier temps que l'ensemble des traitements serait supportable en capacité de calcul et mémorisation par des systèmes embarqués composés de circuit FPGA et DSP.

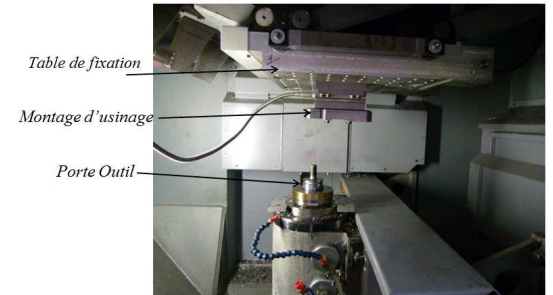
Selon le type de la machine, la transformation donnée entre le plateau, porteur du montage d'usinage, et le porte outil permet de calculer la matrice de projection du modèle CAO sur le plan (uv) .

En réalité, il existe différentes type du machines outils (MOCN) qu'on peut regrouper en machines disposant d'un espace réservé à la préparation du montage, pour une fabrication en série, et machines "ordinaires" dans lesquelles l'espace de préparation du

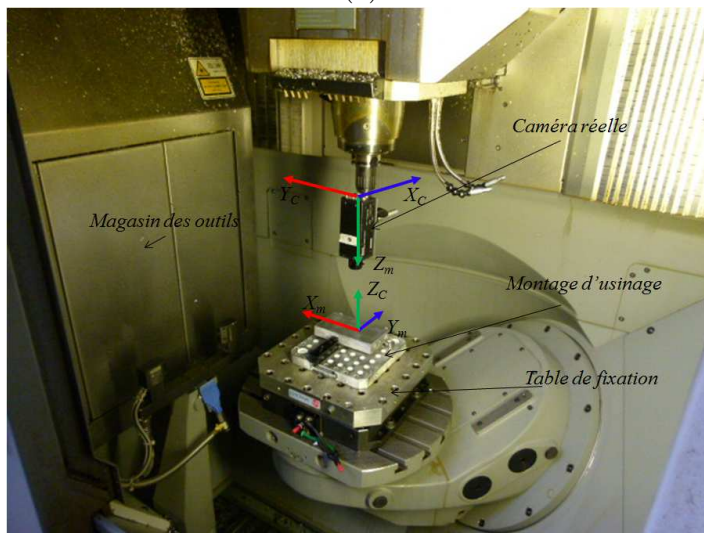
**Chapitre 2. SYSTÈMES DE CONTRÔLE D'USINAGE PAR VISION
30 ET NOTRE PROPOSITION DE CONFIGURATION MONOCULAIRE**



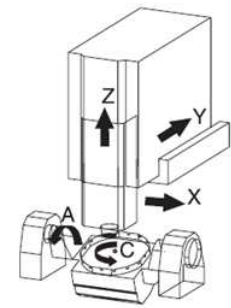
(a)



(b)



(c)



(d)

Figure 2.21 – Configuration du système de vision sur deux types de machines, (a) Zone de préparation du montage d'usinage, (b) le montage préparé est envoyé à la zone d'usinage, (c) machine d'usinage ne disposant pas d'une zone de préparation du montage, la préparation se fait directement dans la zone de l'usinage, (d) schéma des axes de la machine.

montage est celui de l'usinage. Ce qui conduit à adapter la solution de configuration proposée à ces situations.

Deux propositions pourraient être envisageables pour adapter notre solution aux différents types des machines :

- Pour des machines disposant d'un espace de préparation du montage d'usinage,

la caméra est installée dans cet endroit au dessus de la table de préparation du montage d'usinage. La figure 2.21.a en montre un exemple. Les axes Z_m et Z_c sont supposés confondus. Dans ce type de machines⁸, l'usinage se fait à l'envers, i.e. l'outil de coupe est placé au-dessous du montage d'usinage comme le montre la figure 2.21.b.

- Dans le cas où la machine ne dispose pas d'espace de préparation du montage d'usinage, nous proposons de placer la caméra dans le magasin des outils dans un endroit *prédéfini* par la machine avec les outils d'usinage. Pour vérifier le montage d'usinage réalisé, l'opérateur demande à la machine de chercher la caméra dans le magasin des outils (ligne de code, similaire au changement d'outil, à ajouter au début du programme G de CFAO). Les axes Z_m et Z_c sont supposés aussi confondus. Pour l'exemple illustré sur la figure 2.21.c, cette transformation est donnée par les angles des axes A et C. La figure 2.21.d en montre un exemple

2.7 Conclusion

Nous avons vu dans ce chapitre des travaux qui sont en relation directe ou indirecte avec notre problématique. Nous avons montré que ces travaux récents s'intéressent à la sécurité de la machine d'usinage en proposant diverses solutions sans contact, mono et binoculaire. Pour nous une solution bas coût, rapide et assez précise est privilégiée. Nous avons testé quelques systèmes commerciaux de numérisation tridimensionnelle. Malgré leur précision, ces systèmes ont un coût relativement élevé et nécessitent un temps de calcul assez important par rapport au gain de production, ainsi leur installation permanente au sein de la machine n'est pas toujours adaptée. Nous avons réfléchi sur une solution monoculaire permettant d'apporter une décision automatique sur la conformité du montage d'usinage, à partir d'une seule image. L'exploitation de l'information 3D peut enrichir davantage l'analyse de la scène et la prise de décision. Nous avons proposé un élément de solution permettant d'exploiter le flou optique (Depth From Defocus) à partir d'une seule image couleur, en utilisant les trois canaux (RGB). Cette solution sera abordée ultérieurement.

Les méthodes traitant la sécurisation de l'usinage **demeurent rares**. Seules celles évoquées dans cette partie ont été identifiées dans la littérature. Ce qui montre ainsi l'originalité d'une solution sans contact pour la sécurité de la machine. Néanmoins, ces méthodes restent un **outil d'aide** pour l'opérateur, dont la problématique générale est la sécurité des MOCN et la productivité.

La solution proposée privilégie un système sans contact monoculaire. Le choix de la configuration de la caméra par rapport au montage d'usinage permettant une meilleure acquisition de la scène et réduisant les traitements de mise en correspondance est une étape importante.

L'objectif de notre travail est de proposer un système permanent répondant aux contraintes imposées par le cahier des charges et évoquées précédemment. Le système

8. L'usinage "à l'envers" caractérisant cette machine, permet de gagner un gain de productivité pour certains types de pièces en aluminium.

doit être **capable de valider** le montage d'usinage en regard de la trajectoire de l'outil pour éviter les collisions, en laissant toujours la main à l'opérateur pour prendre la décision finale.

Nous présentons aussi, dans les chapitres suivants, les résultats obtenus sur différents types de montages d'usinage et différentes positions de la caméra par rapport au montage d'usinage, respectant l'hypothèse de configuration proposée. Les expérimentations ont été réalisées pour différentes caméras avec différents focales. Cette diversité montre aussi la robustesse de la méthode proposée pour le contrôle du montage d'usinage par vision.

Dans le chapitre suivant, nous nous intéressons à la sélection des primitives 2D/3D pour la construction des graphes. Ces derniers sont un outil pertinent pour représenter la scène et faciliter la mise en correspondance.

Bibliographie

- [Ahmad 2010] R. Ahmad, S. Tichadou et J.-Y. Hascoët. Integration of Vision Based Image Processing for Multi-Axis CNC Machine Tool Safe and Efficient Trajectory Generation. International Journal of Machine Engineering, vol. 10, no. 4, 2010. (Cité en pages vii, 15 et 16.)
- [Colombo 2004] Carlo . Colombo, Dario Comanducci, Alberto Del Bimbo et Alberto Del Federico Pernici. Accurate Automatic Localization of Surfaces of revolution for self-calibration and metric reconstruction. In In Proc. IEEE Workshop on Perceptual Organization in Computer Vision, page 67, 2004. (Cité en page 18.)
- [Deschenes 2003] F. Deschenes, D. Ziou et P. Fuchs. Improved Estimation of defocus blur and spatial shifts in spatial domain : A homotopy-based approach. Pattern Recognition, vol. 36, pages 2105–2125, 2003. (Cité en pages v, 28, 149, 150, 151, 160, 163, 166 et 177.)
- [Faugeras 1993] O. Faugeras, B. Hotz, M. Herve, T. Viéville, Z. Zhang, P. Fua, E. Théron, L. Moll, G. Berry, J. Vuillemin, P. Bertin et C. Proy. Real time correlation-based stereo : algorithm, implementations and applications. Rapport technique 1-49, INRIA Rhones-alpes, 1993. (Cité en page 21.)
- [Karabagli 2013] B. Karabagli, M.-F. Auclair-Fortier, T. Simon et J.-J. Orteu. Depth Map Estimation from a Single RGB image : Principale and Calibration. Stage à l'université de Sherbrooke, Québec, présenté à EDSys, Tarbes, 2013. (Cité en page 28.)
- [Kourosh 2012] Khoshelham Kourosh et Oude Elberink Sander. Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications. sensors, vol. 12, pages 1437–1454, 2012. (Cité en pages 23 et 27.)
- [Kutulakos 1998] KN Kutulakos et JR Vallino. Calibration-free augmented reality. IEEE Transactions on Visualization and Computer Graphics, no. 4(1), pages 1–20, 1998. (Cité en page 20.)

- [Paolo 2007] Favaro Paolo et Soatto Stefano. 3-D Shape Estimation and Image Restoration - Exploiting Defocus and Motion Blur. pages 1–249, 2007. (Cité en pages 28 et 152.)
- [Peyrucat 2013] J-F. Peyrucat, D. Torre, S. Carras, C. Corleto, L. Gonzalez et S. Durand. La mesure dimensionnelle par palpée. Magazine Contrôles, Essais, Mesures, no. 43, pages 33–41, 2013. (Cité en pages vii, 13 et 14.)
- [Reinhard 2004] Klett Reinhard et Rosenfeld Azriel. Digital geometry : Geometric methods for digital picture analysis. The Morgan Kaufmann series in computer graphics and geometric modeling, page 656, 2004. (Cité en page 17.)
- [Tian 2002] X. Tian, W. Hui, Z. Xionghui et R. Xueyu. Object modeling of multiple views using dual quaternion in reverse engineering. The International Journal of Advanced Manufacturing Technology, vol. 20, no. 7, pages 495–502, sep 2002. (Cité en page 21.)
- [Tian 2010] X. Tian, X. Zhang, K. Yamazaki et A. Hansel. A study on three-dimensional vision system for machining setup verification. Robotics and Computer-Integrated Manufacturing, vol. 26, no. 1, pages 46–55, feb 2010. (Cité en pages vii, 19 et 20.)
- [Zhang 2007] X. Zhang, X. Tian, Y. Kazuo et F. Makoto. 3D Reconstruction and Pose Determination of the Cutting Tool from a Single View. In ISVC (2)'07, pages 377–386, 2007. (Cité en page 17.)
- [Zhang 2008] X. Zhang, W. Tsang, X. Tian, Y. Kazuo et M. Masahiko. Automatic Segmentation of the Apparent Contour for 3D Modeling of Cutting Tools from Single View. In ISVC (2)'08, pages 772–781, 2008. (Cité en page 17.)
- [Zhang 2011] X. Zhang, W-M. Tsang, Y. Kazuo et M. Masahiko. A study on automatic on-machine inspection system for 3D modeling and measurement of cutting tools. Journal of Intelligent Manufacturing, may 2011. (Cité en pages vii, 17, 18 et 19.)
- [Zhuo 2011] S. Zhuo et T. Sim. Defocus Map Estimation from a Single Defocused Image. Pattern Recognition, vol. 44, no. 9, pages 1852–1858, 2011. (Cité en pages v, xiii, 28, 149, 150, 152, 157, 158, 159 et 160.)

SÉLECTION DES PRIMITIVES 2D/3D POUR LA CONSTRUCTION DES GRAPHES

" In science there is only physics all the rest is stamp collecting."
William Thomson (1824-1907)

Sommaire

3.1	Introduction et problématique	35
3.2	Détection des contours	37
3.2.1	Modélisation de l'histogramme du module des gradients	39
3.2.2	Méthode initiale	40
3.2.3	Amélioration proposée	43
3.2.4	Résultats et comparaison	44
3.3	Sélection des primitives 2D	50
3.3.1	Sélection des cercles	52
3.3.2	Sélection des droites	53
3.4	Sélection des primitives 3D	58
3.4.1	Création de l'image théorique	58
3.4.2	Extraction des centres des cercles 3D	62
3.4.3	Extraction des points 3D d'intersection des droites	62
3.5	Conclusion	63
	Bibliographie	64

Mots clés : Détection des contours, Sélection des primitives 2D, Sélection des primitives 3D, Image réelle, Image théorique.

3.1 Introduction et problématique

Dans ce chapitre nous apportons les premiers éléments de solution pour répondre à la question de la topologie, organisation spatiale des éléments du montage d'usinage selon leur définition dans le modèle CAO. Pour cela, nous proposons un recalage automatique de ce dernier sur l'image réelle, afin de localiser les régions d'intérêt délimitant

les zones des éléments de la scène. Ce recalage s'appuie sur une mise en correspondance à partir de l'information image utilisée, soit des points 2D/3D caractérisant des primitives 2D/3D extraites de l'image réelle et du modèle CAO respectivement.

Le modèle CAO est une représentation "*métrique*" de la relation entre les éléments de bridage. Pour établir automatiquement une mise en correspondance entre l'image réelle et le modèle CAO nous proposons d'utiliser les graphes de proximité construits à partir des centres des cercles représentant les trous de fixation dans le montage d'usinage et les intersections des droites représentant les contours rectilignes. Les sections 3.3 et 3.4 développent la procédure de sélection des primitives 2D et 3D respectivement.

Étant donné les problèmes dans l'extraction des primitives liés, d'une part, aux conditions photométriques de la scène, i.e. bruit, faible contraste, etc. et, d'autre part, aux difficultés liées à notre application, évoquées dans le premier chapitre, nous proposons de simplifier le problème par une étape de détection des contours dans l'image réelle et la création d'une image théorique synthétisant la topologie donnée par le modèle CAO. En effet, l'image théorique obtenue doit présenter les arêtes des surfaces en respectant les occultations par rapport au point de vue de la caméra théorique. La section 3.4.1 présente les étapes de la création de l'image théorique. Ces arêtes trouvent leurs correspondants dans l'image réelle à partir des contours. Une nouvelle méthode a été mise en œuvre pour la détection des contours à partir de la modélisation explicite de l'histogramme du module des gradients. Nous avons comparé cette méthode suivant quelques indicateurs de performances utilisés dans la communauté de la vision par ordinateur, donnés par les courbes : Précision *vs* Rappel (*PR*), la courbe *Receiver Operating Characteristic* (*ROC*), la courbe d'exactitude (*E*), etc. Ce point est développé dans la section 3.2.

Le chapitre est structuré comme le montre l'organigramme de la figure 3.1.

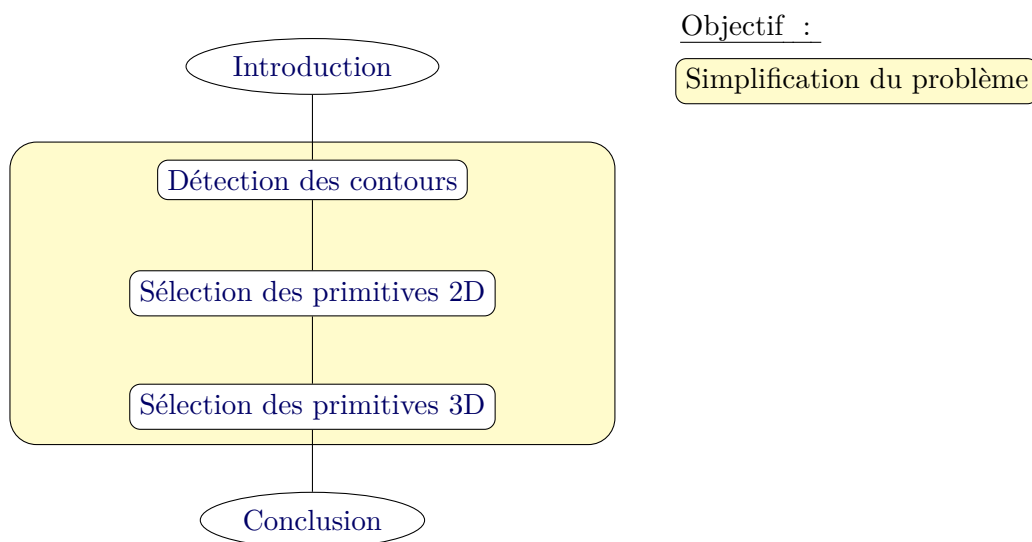


Figure 3.1 – Organigramme du chapitre 3 et objectif de la partie.

3.2 Détection des contours

Cette partie présente le travail mené pour la détection des contours dans l'image de montage d'usinage. Nous avons proposé dans un premier temps une approche par binarisation des modules des gradients. Le calcul des gradients par l'opérateur de Prewitt peut être confié à un circuit FPGA de composition simple permettant d'optimiser le temps de calcul, les autres éléments de l'algorithme peuvent être implantés sur un DSP. Cette méthode est une amélioration d'une méthode présentée dans [Simon 1994].

D'après la structuration des méthodes de binarisation proposée par *G. Papari et N. Petkov* dans leur état de l'art [Papari 2011], cette méthode peut être classée dans le premier des trois groupes de classification. Ces trois groupes sont caractérisés par l'information image utilisée :

- **Méthodes basées sur des caractéristiques locales** : les méthodes dérivatives suivies d'un seuillage plus ou moins sophistiqué [Ku 2005], [Debashis 2010].
- **Méthodes globales ou contextuelles** : parmi lesquelles on trouve les contours actifs [Kass 1988].
- **Méthodes multi-résolution** : par exemple les filtres de Gabor [Namuduri 1992].

La figure 3.2 présente la classification et la place de notre méthode par rapport aux méthodes de seuillage automatique.

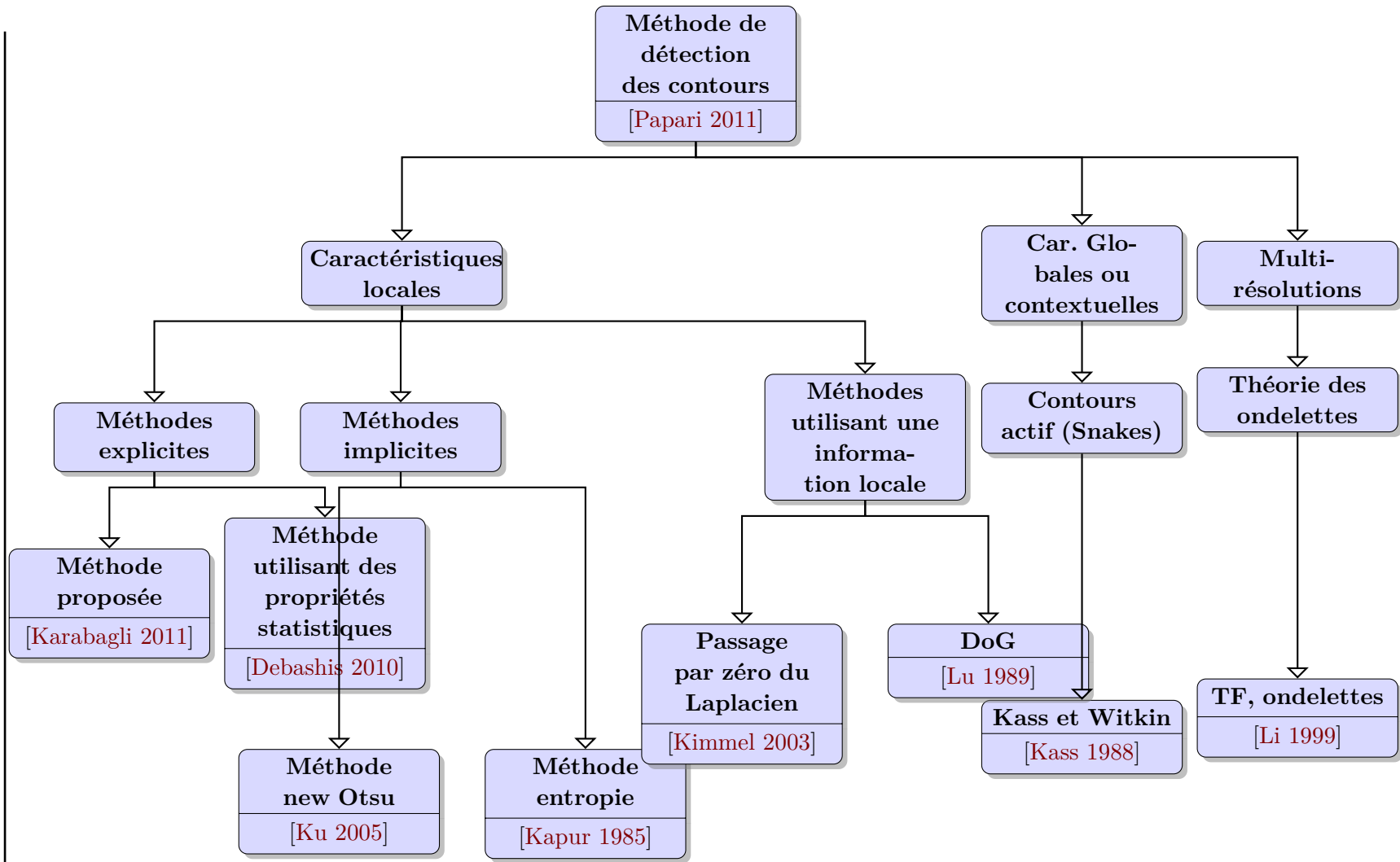


Figure 3.2 – Classification des méthodes de binarisation selon G. Papari et N. Petkov [Papari 2011], et le positionnement de la méthode proposée

3.2.1 Modélisation de l'histogramme du module des gradients

De nombreuses méthodes de détermination automatique du seuil de binarisation utilisent la forme de l'histogramme du module des gradients [Ku 2005] [Debashis 2010]. Dans le cas d'images ne présentant pas de textures très marquées on constate une forme bimodale de l'histogramme. Un mode existe pour les niveaux de gris faibles des gradients représentant des zones homogènes bruitées ou de petites zones texturées. Le second mode, à des niveaux de gris élevés, représente les discontinuités de luminosité plus marquées, probablement les contours et le bruit "poivre et sel".

Pour déterminer le seuil de binarisation, deux solutions sont envisageables, soit construire un modèle de l'histogramme et en déduire une valeur de séparation entre les deux modes à partir des paramètres du modèle, soit considérer la forme de l'histogramme comme un mélange de deux lois de probabilité, probabilité d'appartenance d'un niveau de gris à une classe de contour ou "non-contour". Le principe est de construire une fonction coût séparant au mieux les 2 lois.

Debashis Sen et Sankar K. Pal [Debashis 2010] ont proposé un modèle composé de la somme d'une fonction de Rayleigh et d'une fonction gaussienne, et construisent une règle de séparation à partir des valeurs des coefficients de dissymétrie et d'aplatissement pour déterminer la région d'intérêt (ROI) dans l'histogramme afin d'y placer un seuil de binarisation. Les coefficients statistiques de dissymétrie et d'aplatissement donnés par les moments d'ordre 3 et 4 sont constants pour ces lois. Ce travail montre l'intérêt permanent pour la mise en œuvre de la détermination automatique de seuil de binarisation et l'intérêt de l'utilisation de deux modèles dans la modélisation de l'histogramme du module des gradients.

Afin d'obtenir une grande variété de formes, notamment présentant des dissymétries, nous proposons¹ de modéliser l'histogramme du module des gradients par la somme de 2 fonctions **Gamma**. Ces deux méthodes s'inscrivent dans la première catégorie suivant la classification donnée en figure 3.2.

Pour établir une comparaison pour l'évaluation des performances, nous avons testé deux méthodes s'inscrivant dans la 2^{ème} catégorie : la méthode très connue d'Otsu améliorée par [Ku 2005] dont la fonction coût à maximiser est calculée à partir de la variance interclasse.

$$\sigma_b^2(t) = w_c(t)w_{nc} \left[\mu_c^2(t) - \mu_{cn}^2(t) \right] \quad (3.1)$$

avec $t \in [0, 255]$ le seuil de séparation des deux classes. Les quantités $w_c(t)$ et $w_{cn}(t)$ représentent les proportions relatives des deux classes contours w_c et non-contours w_{cn} ($w_{cn}(t) = 1 - w_c(t)$). μ_c et μ_{cn} représentent la moyenne de la classe $w_c(t)$ et $w_{cn}(t)$ respectivement.

Or,

$$w_c(t) = \sum_{i=0}^t P(i) \text{ et } w_{cn}(t) = \sum_{i=t+1}^{255} P(i) \quad (3.2)$$

1. Le détail sera donné aux paragraphes suivants, sections 3.2.2 et 3.2.3

Le seuil est le niveau de gris donnant la plus forte distance interclasse. La méthode de Kapur *et al.* [Kapur 1985] se repose sur la maximisation de l'entropie de l'histogramme en considérant 2 classes, contours et non-contours à partir de la corrélation entropique donnée par :

$$T_{opt} = \operatorname{argmax} [H_f(T) + H_b(T)] \text{ , avec} \tag{3.3}$$

$$H_f(T) = - \sum_{i=0}^T \frac{p(i)}{P(T)} \log \frac{p(i)}{P(T)} \text{ et } H_b(T) = - \sum_{i=T+1}^{255} \frac{p(i)}{P(T)} \log \frac{p(i)}{P(T)}$$

Le seuil T_{opt} est la valeur du niveau de gris pour lequel la somme des deux entropies H_f et H_b pour le fond et les contours, respectivement, est maximale.

Les trois méthodes entropie [Kapur 1985], la nouvelle méthode d'Otsu [Ku 2005] et [Debashis 2010] considèrent le seuil de binarisation comme étant la valeur du niveau de gris séparant au mieux les deux classes, contours et non contours. Nous nous intéressons dans la section suivante à la méthode proposée qui s'appuie sur un principe de modélisation géométrique de l'histogramme du module des gradients.

3.2.2 Méthode initiale

La méthode initiale proposée [Simon 1994] utilise la forme constante de l'histogramme du module des gradients modélisée par l'expression (3.4), somme de deux fonctions de Gamma pour marquer l'asymétrie du mode, où $x \in [0, 255]$ le niveau de gris, α , β , n et m des paramètres à identifier :

$$H(x) = \frac{(1 - \lambda)}{\Gamma(n + 1)} x^n \cdot e^{-\alpha x} \alpha^{n+1} + \frac{\lambda}{\Gamma(m + 1)} x^m \cdot e^{-\beta x} \beta^{m+1} \tag{3.4}$$

Placer un seuil de binarisation lié à la largeur du mode a été la solution envisagée. Si la valeur de λ est petite, cela conduit à une forme unimodale très marquée. Le modèle est réduit alors à l'expression (3.5)

$$H(x) = \frac{(1 - \lambda)}{\Gamma(n + 1)} x^n \cdot e^{-\alpha x} \alpha^{n+1} \tag{3.5}$$

L'intersection de la tangente au point d'inflexion de la plus grande abscisse avec l'axe des abscisses donne une bonne évaluation de cette largeur.

Cette intersection peut être calculée à partir des paramètres du modèle par l'expression :

$$W = \frac{(n + \sqrt{n})^2}{\alpha n} \tag{3.6}$$

Pour retrouver ces paramètres, nous calculons les moments d'ordre $i \in [0, 2]$, définis par :

$$m_i = \int_0^\infty x^i H(x, n, \alpha) dx \tag{3.7}$$

où x le niveaux de gris et H l'histogramme donné en fonction des paramètres du modèle.

Soit les moments m_0 , m_1 et m_2 calculés, pour le cas continu, par :

$$\begin{aligned}
 m_0 &= \int_0^\infty x^0 H(x, n, \alpha) dx = \int_0^\infty x^0 \frac{x^n e^{-\alpha x} \alpha^{(n+1)}}{\Gamma(n+1)} dx \\
 m_0 &= \frac{\alpha^{(n+1)}}{\Gamma(n+1)} \int_0^\infty x^0 x^n e^{-\alpha x} dx = \frac{\alpha^{(n+1)}}{\Gamma(n+1)} \frac{\Gamma(n+1)}{\alpha^{(n+1)}} = 1 \\
 m_1 &= \frac{\alpha^{(n+1)}}{\Gamma(n+1)} \int_0^\infty x^1 x^n e^{-\alpha x} dx = \frac{\alpha^{(n+1)}}{\Gamma(n+1)} \frac{\Gamma(n+2)}{\alpha^{(n+2)}} \\
 m_2 &= \frac{\alpha^{(n+1)}}{\Gamma(n+1)} \int_0^\infty x^2 x^n e^{-\alpha x} dx = \frac{\alpha^{(n+1)}}{\Gamma(n+1)} \frac{\Gamma(n+3)}{\alpha^{(n+3)}}
 \end{aligned} \tag{3.8}$$

Puisque $\Gamma(n+1) = n\Gamma(n)$ nous en déduisons que :

$$\begin{aligned}
 m_2 &= \frac{m_1(n+2)}{\alpha} \\
 m_1 &= \frac{m_0(n+1)}{\alpha}
 \end{aligned} \tag{3.9}$$

Par conséquent, les paramètres n et α sont calculés à partir de m_0 , m_1 et m_2 par :

$$n = \frac{2m_1^2 - m_2}{m_2 - m_1^2}, \quad \alpha = \frac{m_0}{m_1}(n+1) \tag{3.10}$$

pour le cas où n est nul, cas des images comportant de grandes zones saturées au noir ou au blanc, le modèle devient une exponentielle décroissante, et peut être exprimé par l'équation (3.11). Un exemple est illustré sur la figure 3.3 :

$$H(x) = \alpha e^{-\alpha x} \tag{3.11}$$

Le seuil calculé est obtenu par l'intersection de la tangente à l'origine avec l'axe des niveaux de gris, sa valeur est : $W = \frac{1}{\alpha}$.

Le seuil W ainsi appliqué donne une image avec des contours très bruités. Une comparaison avec des seuils proposés par des experts a conduit à l'augmentation d'une proportion moyenne de 2.8 pour un intervalle de variation comprise entre 2 et 3. Cette

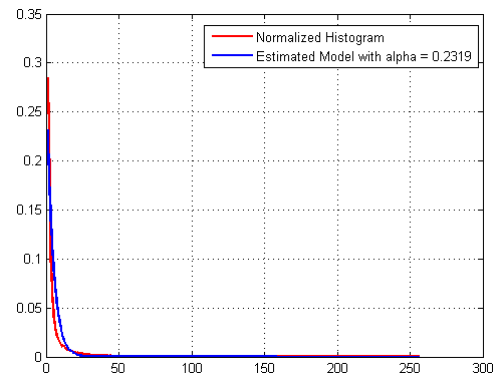


Figure 3.3 – Modélisation de l’histogramme du module des gradients par une exponentielle décroissante avec, $\alpha = 0.2319$

comparaison s’appuie sur le test de Spearman², permettant une étude de l’indépendance entre les seuils donnés par les experts et le seuil calculé par la méthode. Le tableau suivant regroupe les seuils de corrélation calculés.

Tableau 3.1 – Corrélation des experts

	A	B	C	D	E	F	G
A	1	0.92	0.93	0.84	0.93	0.96	0.93
B	0.92	1	0.89	0.90	0.97	0.91	0.95
C	0.93	0.89	1	0.80	0.85	0.95	0.86
D	0.84	0.90	0.80	1	0.90	0.84	0.90
E	0.93	0.97	0.85	0.90	1	0.90	0.95
F	0.96	0.91	0.95	0.84	0.90	1	0.86
G	0.93	0.95	0.86	0.90	0.95	0.86	1

Nous remarquons que le test de Spearman rejette l’indépendance et le coefficient de corrélation est assez fort, montrant ainsi une relation linéaire probable entre les seuils calculés et les seuils des experts. Le seuil optimal est donc donné par $S = k \times W$, avec $k = 2.8$ la pente de la droite exprimant cette relation linéaire.

La méthode donne de bons résultats mais ne tient pas compte de l’étalement du deuxième mode. La figure 3.4 présente des exemples de résultats sur trois images de la base *Berkeley Segmentation DataSet and Benchmarks, BSDS300*³.

2. Le coefficient de corrélation de Spearman indique le degré de dépendance existant entre les rangs des données.

3. <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html>

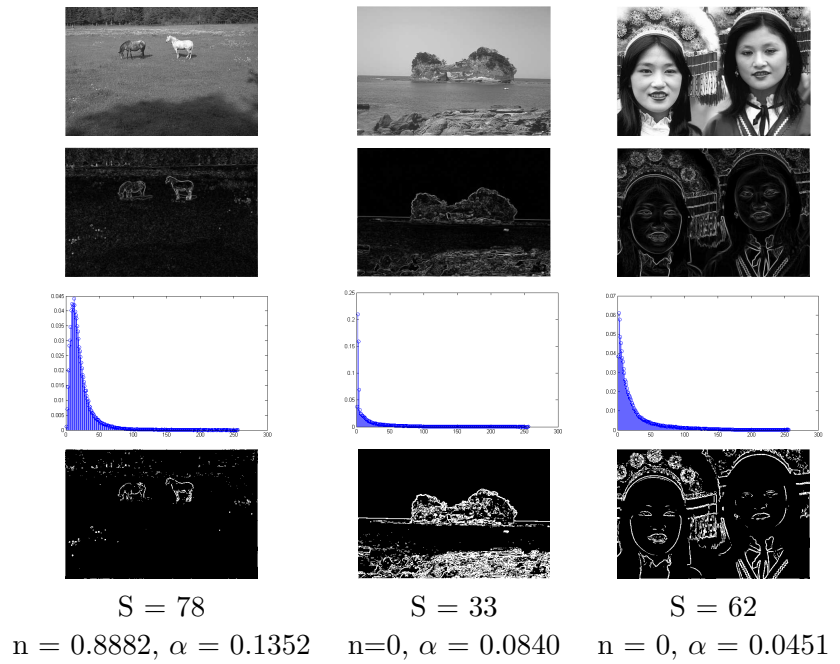


Figure 3.4 – Résultat du seuillage par la méthode initiale sur des images de la base BSDS300.

3.2.3 Amélioration proposée

L'amélioration proposée consiste à tenir compte du deuxième mode et d'estimer sa largeur. Si on applique la méthode de modélisation initiale sur l'histogramme de la première modélisation, le résultat n'est pas satisfaisant. La forme du mode étant plus bruitée, les paramètres estimés ne donnent pas un bon modèle, car la modélisation de la forme du second mode est écartée. Compte tenu de l'asymétrie de ce dernier, nous proposons de prendre un développement limité de la forme initiale, soit l'expression (3.12) pour privilégier une forme parabolique du second mode :

$$H(x) = \frac{(1-\lambda)}{\Gamma(n-1)} x^n \cdot e^{-\alpha x} \alpha^{n+1} + \underbrace{\lambda(ax^2 + bx + c)}_{H_d} \quad (3.12)$$

Nous considérons l'histogramme comme une densité de probabilité (probabilité d'apparition d'un niveau de gris par rapport à l'ensemble des pixels) ce qui implique une renormalisation, les paramètres a , b , c et λ seront calculés pour respecter : $\int_0^{255} H_d(x) dx = 1$

Le triplet (a, b, c) contient les coefficients du polynôme de la forme parabolique. Ils sont déterminés par approximation régressive aux moindres carrés. La figure 3.5 présente l'histogramme du module des gradients d'une image issue de la base BSDS300 et le modèle associé.

La détermination des coefficients du modèle parabolique est réalisée sur un intervalle de niveau de gris $x \in [\varepsilon_1, \varepsilon_2]$. La valeur ε_1 est calculée comme la première valeur de la densité réduite (dans le sens des niveaux de gris croissant), valeur absolue de la

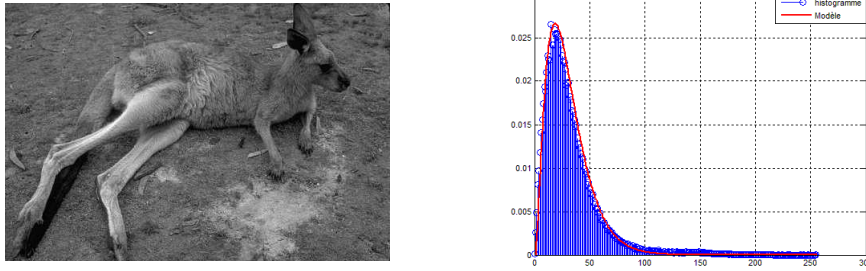


Figure 3.5 – Modélisation de l’histogramme par la méthode proposée. Image de la base *BSDS300*.

différence entre l’histogramme expérimental complet et son modèle unimodal, supérieur à un seuil significatif, seuil fixé heuristiquement à 10^{-4} . La borne supérieure ε_2 est obtenue par la même technique mais dans le sens des niveaux de gris décroissants.

Le seuil est calculé en prenant l’abscisse du point d’intersection des deux tangentes D1 et D2 (figure 3.6.a), respectivement au point d’inflexion du modèle Gamma, et au point d’abscisse la plus petite du modèle polynomial (figure 3.6.b).

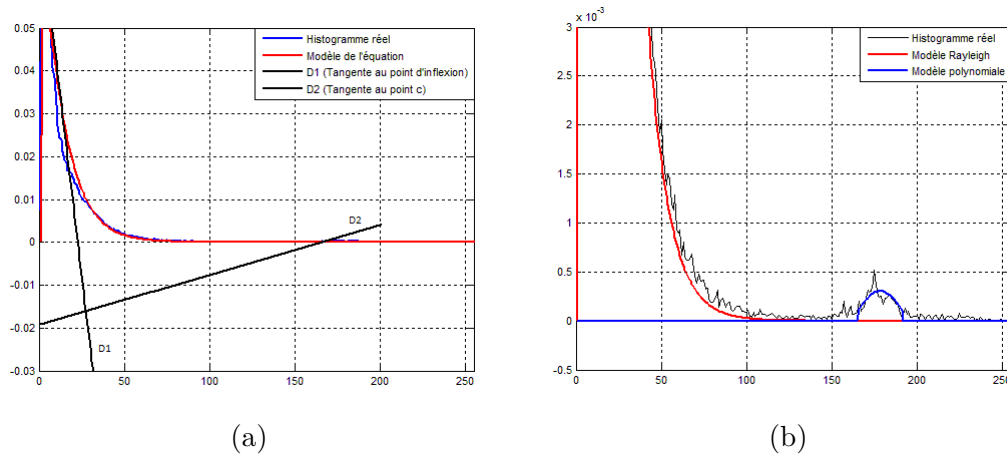


Figure 3.6 – Seuil de binarisation. (a) : point de croisement des tangentes D1 et D2. (b) : Zoom avant sur le deuxième mode.

Cette technique de binarisation automatique donne de meilleurs résultats par rapport aux méthodes de même catégorie. Dans la section suivante, nous l’évaluons par rapport aux critères de performance évoqués précédemment, e.g. courbes (*PR*), (*ROC*), (*E*), etc.

3.2.4 Résultats et comparaison

Afin de comparer notre méthode par rapport aux méthodes citées précédemment, et classées dans la même catégorie, nous proposons de reprendre les indicateurs de

performance proposés par le groupe de vision de l'Université de Berkeley. La binarisation est considérée comme une classification des pixels en deux classes. Celle des "contours" (Classe A) et celle des "non contours" (Classe \bar{A}). Cela suppose qu'il existe une référence pour contrôler la classification. Cette référence a été produite à partir de propositions d'experts.

Ces indicateurs sont résumés par deux courbes. La courbe ROC qui caractérise la qualité de discrimination du processus de classification, et la courbe (PR) de la précision (P) en fonction du rappel (R). Enfin la mesure (F) ou F - score proposée par [Huijsmans 1993] et [Swets 1996] évalue l'efficacité de la méthode par rapport à l'exactitude (E), fraction des pixels bien classés. La mesure F exploite la précision et le rappel combiné par l'expression suivante :

$$F = \frac{2 \times P \times R}{P + R} \quad (3.13)$$

Le rappel est le nombre de pixels correctement attribués aux contours par rapport au nombre total de pixels des contours. La précision est le nombre de pixels correctement attribués aux contours rapporté au nombre de pixels attribués aux contours.

La courbe ROC exploite la sensibilité (Sen) qui mesure la capacité de trouver les pixels des contours et la spécificité (Sep), capacité d'écarter les pixels des "non contours".

Ces cinq indicateurs, P , R , E , Sen , Sep sont calculées à l'aide de la matrice de confusion, donnée par le tableau suivant :

Tableau 3.2 – *Table de confusion*

		Référence	
		Classe A	Classe \bar{A}
Prédiction	Classe A	VP	FP
	Classe \bar{A}	FN	VN

VP (vrais positifs), est le nombre de pixels de l'image correctement attribués aux contours, FP (faux positifs) le nombre de pixels de l'image incorrectement attribués aux contours, FN (faux négatifs) le nombre de pixels "non contours" attribués aux contours et VN (vrais négatifs) les pixels de "non contours" correctement attribués. Nous avons alors :

$$P = \frac{VP}{VP + FP}$$

$$R = \frac{VP}{VP + FN}$$

$$E = \frac{VP + VN}{VP + FP + VN + FN}$$

$$Sen = \frac{VP}{VP + FN}$$

$$Sep = \frac{VN}{FP + VN}$$

Avant de procéder à l'évaluation de la méthode et la comparaison par rapport aux méthodes de même catégorie, le résultat de la détection des contours est illustré sur les figures 3.7 et 3.8 qui donnent des exemples de résultats. Dans la première ligne un extrait de la base de référence définie par des experts, la deuxième ligne les contours de référence, la troisième ligne les contours obtenus par notre méthode, la quatrième par la méthode d'entropie et la cinquième par la méthode Otsu.

L'étude des performances de la méthode est réalisée sur une centaine d'images tirées aléatoirement de la base *BSDS300*.

La figure 3.9 présente les courbes (*PR*) de ces méthodes placées dans le graphe des méthodes testées par le groupe vision de l'Université de Berkeley dans le travail de thèse de M.R. Maire [Maire 2009], le *F - score* de notre méthode est précisé. Le point vert indique la mesure *F* des experts. Nous constatons que notre méthode, marquée en bleu continu, se situe au niveau de la méthode d'entropie et légèrement supérieure à celle d'Otsu avec un *F - score* honorable de 0.56.

La figure 3.10 présente la courbe *ROC* pour chaque méthode de calcul automatique du seuil de binarisation pour les quatre méthodes. La courbe *ROC* de notre méthode produit son maximum pour une valeur de spécification plus petite que la méthode d'Otsu, mais plus grande que la méthode initiale et d'entropie, et se situe bien au-dessus de cette dernière en sensibilité. La courbe *ROC* idéale vaut 1 pour toute spécificité.

La figure 3.11 présente l'évaluation de l'exactitude pour les quatre méthodes de binarisation. On retrouve en abscisse la méthode utilisée et en ordonnée la valeur moyenne de *E* et l'écart-type moyen du processus de calcul d'exactitude sur la même centaine d'images de la base *BSDS300*.

L'étude de l'exactitude confirme le bon classement de notre méthode par rapport aux autres méthodes de même catégorie (classification donnée en figure 3.2), avec une valeur moyenne autour de 68% et un écart-type moyen de 37%, alors que la méthode d'entropie et la méthode initiale ont leur valeur autour de 25% et 59% et un écart-type moyen de 39% et 43%, respectivement. Cette expérience montre la capacité de chaque méthode à bien détecter les pixels des contours en écartant les pixels des non-contours.

Pour l'ensemble des images nous évaluons aussi le temps moyen de calcul enregistré sur un calculateur Xeon 2.53 *Ghz* avec 4 *Go* de mémoire vive. Le résultat est présenté sur la figure 3.12 sous forme d'un diagramme à barres. L'axe des abscisses porte le nom de l'algorithme utilisé et l'axe des ordonnées le temps de calcul moyen du seuil exprimé en secondes. Ces résultats montrent que la méthode proposée est beaucoup plus rapide que celle d'Otsu et d'entropie.

Cette méthode répond aux besoins de disposer d'un détecteur de contours rapide. La figure 3.13 donne les résultats obtenus par l'application des trois méthodes et la méthode proposée sur l'image de notre montage d'usinage. L'image des contours obtenue va être utilisée pour détecter et sélectionner les primitives 2D.

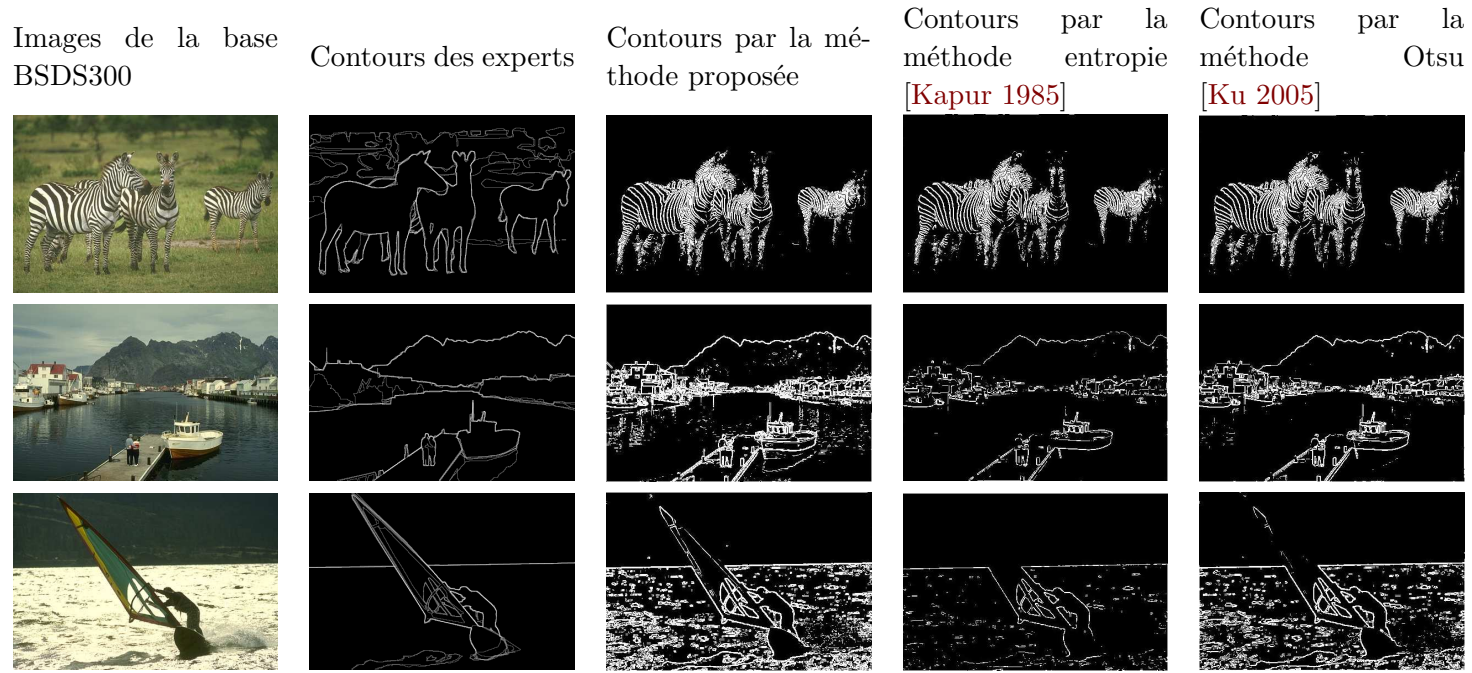


Figure 3.7 – Résultat d'un extrait de la base BSDS300

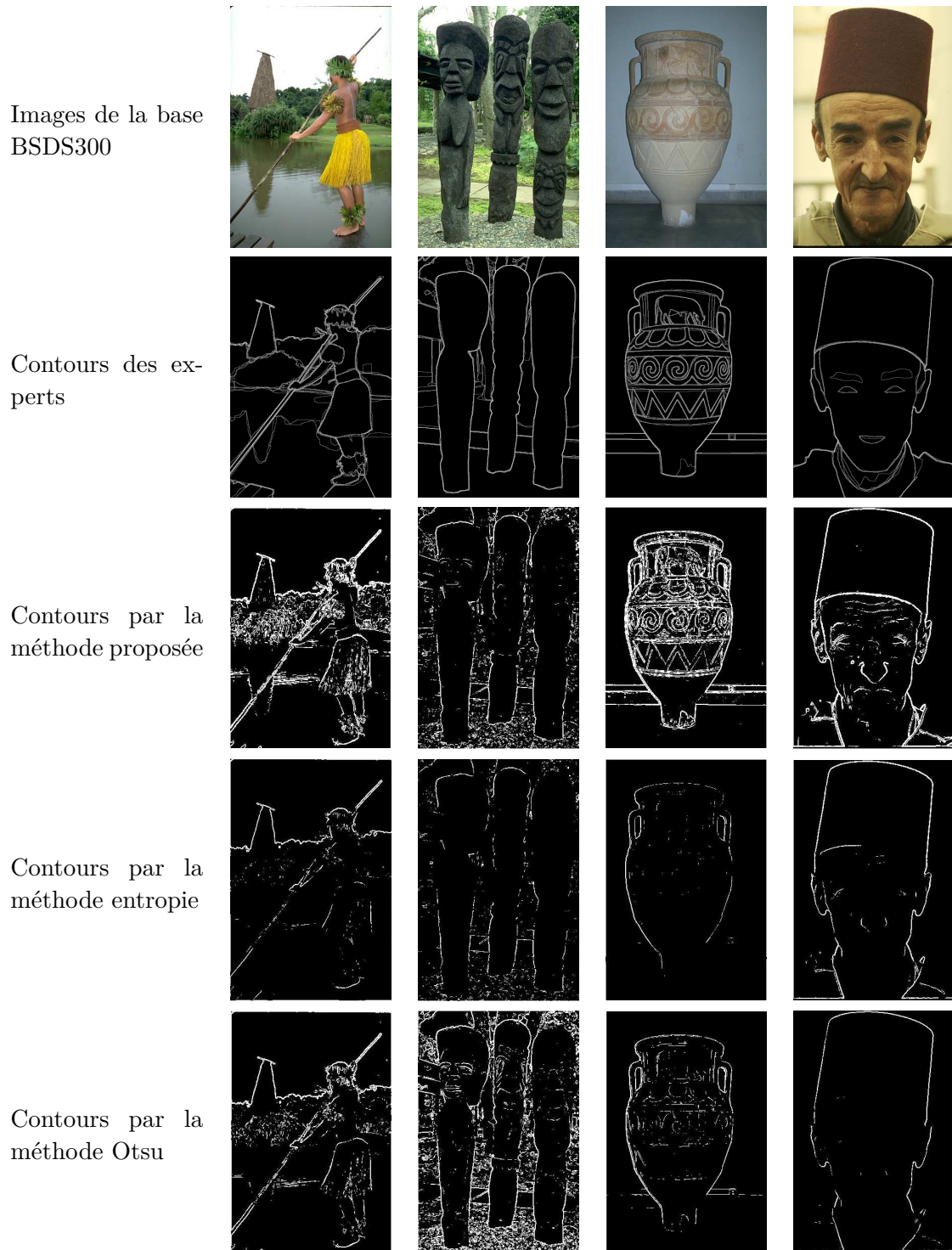


Figure 3.8 – Résultat d'un extrait de la base BSDS300

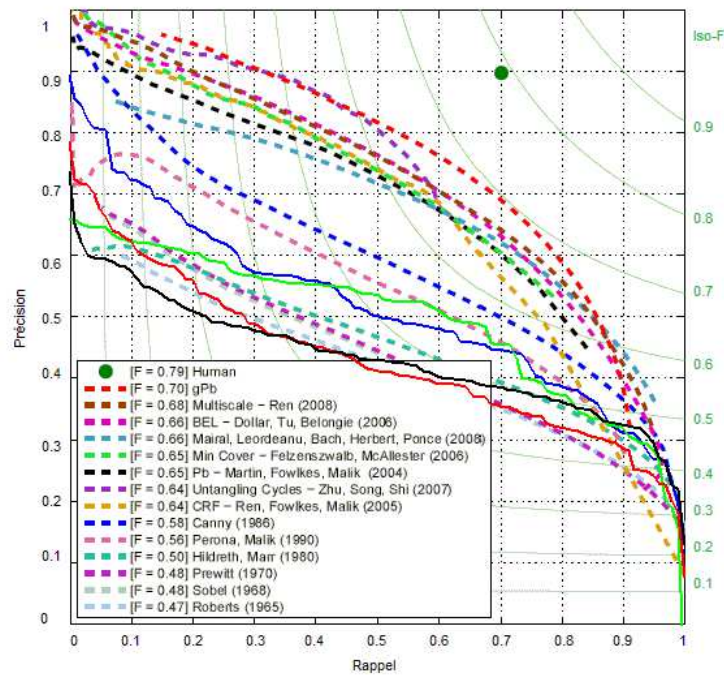


Figure 3.9 – Courbes PR : en rouge continu la méthode Otsu [Ku 2005] avec F – score = 0.50, en vert continu, la méthode d’entropie [Kapur 1985] avec F – score = 0.57, en noir continu la méthode initiale avec F – score = 0.48 et en bleu continu, la méthode proposée avec F – score = 0.56

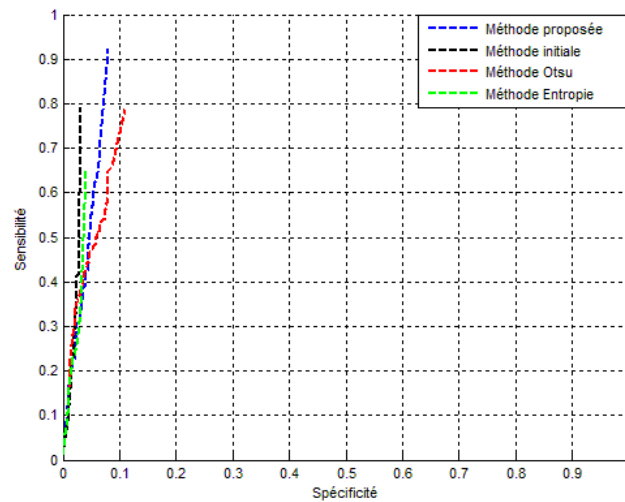


Figure 3.10 – Courbes ROC : en rouge la méthode Otsu [Ku 2005], en vert la méthode d’entropie [Kapur 1985], en noir la méthode initiale et en bleu la méthode proposée.

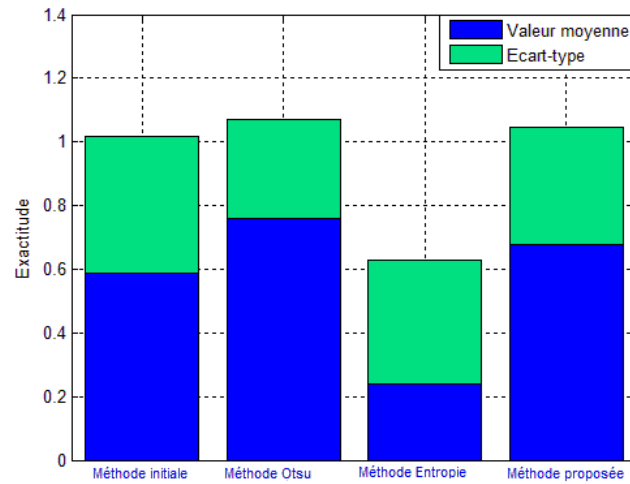


Figure 3.11 – Diagramme d’exactitude : en bleu la valeur moyenne d’exactitude, et en vert l’écart-type correspondant. Le calcul a été fait sur des centaines d’images tirées aléatoirement de la base *BSDS300*.

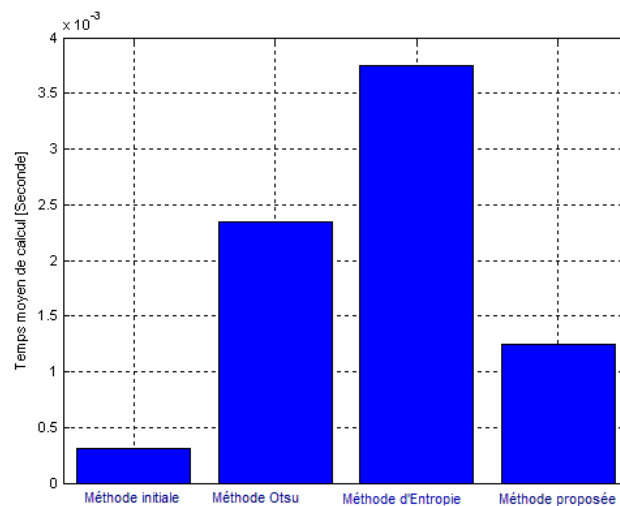


Figure 3.12 – Évaluation du temps de calcul (*CPU Time*) de seuil de binarisation pour chaque méthode

3.3 Sélection des primitives 2D

Cette section décrit la première partie de l’algorithme proposé pour la construction des graphes représentant la relation topologique de la scène, afin de déterminer l’appariement entre les points 2D extraits de l’image réelle et leurs correspondants 3D dans le modèle CAO.

Le montage d’usinage est un dispositif constitué essentiellement par des éléments

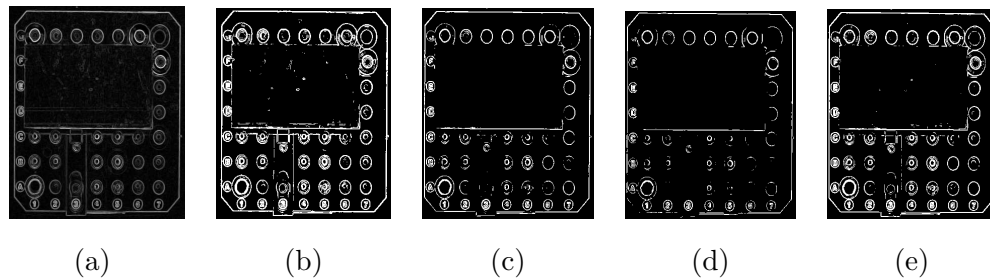


Figure 3.13 – *Seuillage automatique : (a) : image du module des gradients calculée par l'opérateur Prewitt, (b) : résultat de binarisation par la méthode initiale, (c) : résultat de binarisation par la méthode Otsu, (d) : résultat de binarisation par la méthode entropie et (e) : résultat de binarisation par la méthode proposée.*

polyédriques, cylindriques et sphériques comme le montre la figure 3.14⁴ présentant les différents outils de bridage ainsi que les supports des montages d'usinage.

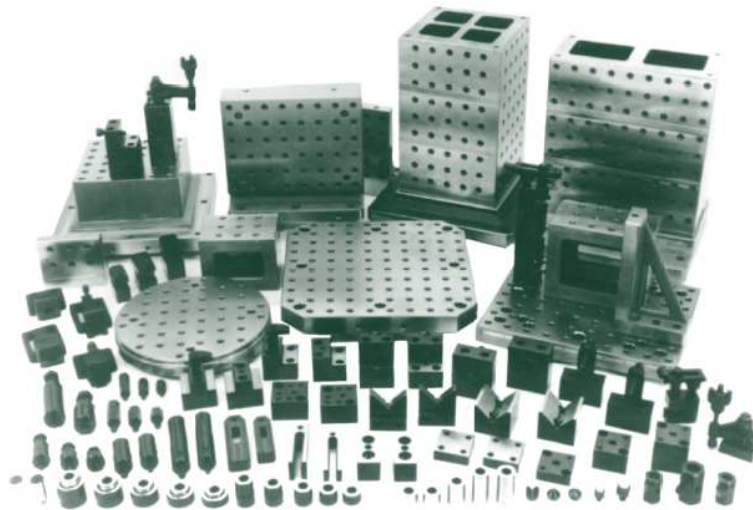


Figure 3.14 – *Différentes éléments de bridage (Catalogue CARRLANE)*

Pour des raisons de complexité évoquées précédemment, nous avons choisi les cercles comme primitives pertinentes. Ces cercles sont présents dans la quasi totalité des montages d'usinage. Ils représentent les trous de fixation des éléments de bridage. La détection de ces cercles pertinents est effectuée en utilisant la transformée de Hough appliquée sur l'image des contours. La configuration du système de vision évoquée dans le chapitre 2, section 2.6, et les paramètres de ce système de vision garantissent une déformation modérée des cercles en ellipses n'impliquant pas un déplacement des centres. Ces cercles portent une information significative de l'organisation topologique du montage. La section 3.3.1 développe le processus de traitement pour la détection des cercles.

4. CARR LANE MODULAR FIXTURING CATALOGUE 1997 CARR LANE MANUFACTURING CO. 8304-897

Les contours d'objets dans leur partie rectiligne enrichissent l'information de la topologique de la scène. Nous avons calculé les segments de droite portés par les contours en utilisant l'algorithme *Line Segment Detector (LSD)* proposé par Rafael Grompone von Gioi *et al.* [Grompone von Gioi 2010] pour sa capacité à assembler localement des points dans leur alignement en utilisant l'information du gradient. Les points d'intersection des droites donnent une information pertinente sur la scène. Nous avons exploité ces deux informations : centres des cercles, points d'intersection des droites, pour la construction des graphes réel et théorique issus respectivement de l'image réelle et théorique du montage d'usinage.

3.3.1 Sélection des cercles

Nous utilisons la transformée de Hough (TH) pour la détection des cercles. La transformée de Hough est une méthode classique pour la détection des formes spécifiques à partir de l'image des contours, elle tente d'accumuler des évidences sur l'existence d'une forme particulière, e.g. circulaire, dans l'espace des paramètres décrivant cette forme. Le choix de l'espace des paramètres est une étape importante dans la représentation de Hough. En effet, représenter un cercle dans l'espace cartésien $r^2 = (x - a)^2 + (y - b)^2$ peut sembler un choix *judicieux*, mais pour pouvoir traiter complètement l'ensemble des cercles possibles, a et b doivent varier entre $]-\infty, +\infty[$ ce qui conduit à une difficulté de quantification de l'espace des paramètres, e.g. une grande taille de la matrice d'accumulation. En général, cette étape est délicate, si la forme circulaire à détecter est affectée par le bruit, ce qui est souvent le cas en traitement d'image.

Nous avons choisi une représentation plus adaptée et bornée. La forme circulaire est donc, caractérisée par le rayon r et l'angle θ dans le plan polaire. Ce changement de repère est calculé en utilisant les relations données par l'équation (3.14)

$$\begin{cases} x &= x_0 + r \cos(\theta) \\ y &= y_0 + r \sin(\theta) \end{cases} \quad (3.14)$$

Or θ est bornée entre $-\pi$ et π , $r \in [r_1, r_2]$ avec $r_1 \leq r_2$. x_0, y_0 sont les coordonnées du centre du cercle.

Le résultat de la détection est illustré par la figure 3.15

Les cercles représentés par des centres (points jaunes) reflètent les cercles pertinents calculés en utilisant le nombre de rayons des cercles détectés donnés par l'équation (3.15), avec ρ le taux d'acceptation fixé ≥ 0.8 .

$$\rho = \left\{ \rho_i \mid \rho_i = \frac{N_{inlier}}{N_{inlier} + N_{outlier}} \right\} \quad (3.15)$$

Avec N_{inlier} le nombre des cercles possédant le même rayon r_{inlier} et $N_{outlier}$ le nombre des cercles possédant des rayons différents à r_{inlier} . Avec $N = N_{inlier} + N_{outlier}$ et $r_1 \leq r_{inlier} \leq r_2$.

Ces cercles sont un élément d'appui pour représenter l'organisation topologique des éléments du montage d'usinage au sens où, si un élément de bridage est présent dans une région, ceci implique une disparition des cercles de cette région, e.g. la région

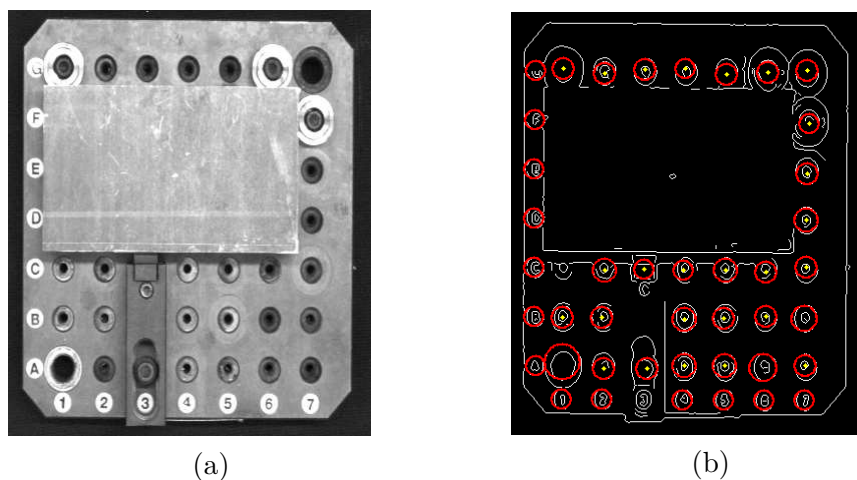


Figure 3.15 – Détection des cercles : (a) image des niveaux de gris, (b) : détection des cercles sur l'image des contours. Les cercles mentionnés avec un centre jaune reflètent les cercles pertinents

du brut. Néanmoins cette supposition n'est pas toujours valide car, dans certains cas, l'élément de bridage peut avoir certaines formes circulaires, ce qui pourrait déclencher des fausses détection, e.g. bride de fixation, présence d'un trou, etc. Nous allons étudier dans le chapitre 4, suivant, l'impact de ce problème dans la mise en correspondance des graphes.

3.3.2 Sélection des droites

Les segments des droites donnent une information importante sur l'organisation topologique de l'image du montage d'usinage. La détection des segments de droites est un problème ancien en vision par ordinateur [Frei 1977] mais qui reste d'actualité [Grompone von Gioi 2010]. Certaines méthodes classiques utilisent l'information des contours pour détecter une forme rectiligne. Elles s'appuient sur la forme de la courbure de contour pour en extraire une forme droite [Deschenes 2004], mais elles sont *gourmandes* en temps de calcul.

Une méthode récente s'appuyant sur l'information de gradient, module et direction proposée par Rafael Grompone von Gioi *et al.* [Grompone von Gioi 2010] a été utilisée pour notre application. Cette méthode est basée sur la méthode de Burns *et al.* [Burns 1986] et une approche *a contrario* [Beveridge 1996] pour la validation d'un segment. Elle s'appuie sur l'information de la direction du gradient. En pratique, à partir de l'image de gradient, nous prenons un pixel et nous regardons dans un voisinage de 3×3 la présence d'un autre pixel dont la valeur du gradient est maximum avec une direction comparable à celle du pixel de référence, avec une certaine tolérance τ^5 , permettant de définir une région des pixels connexes. Chaque région est approximée par un rectangle englobant la région. La taille du rectangle est choisie comme l'enveloppe

5. Pour nous la précision est fixée à $0.11rad$

convexe de la région. Le principe du rectangle caractérisant la région et le résultat de calcul des régions sur l'image du montage d'usinage sont illustrés sur la figure 3.16.

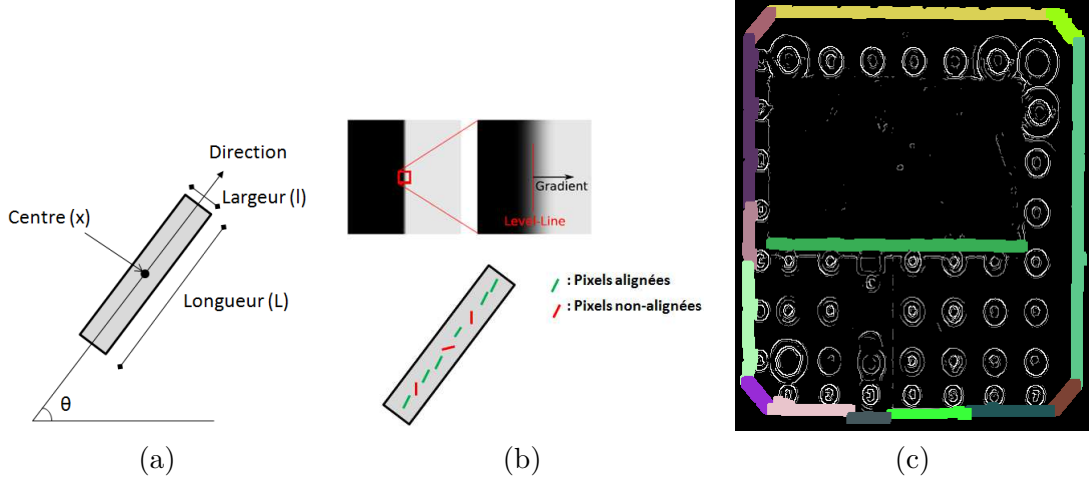


Figure 3.16 – Calcul des régions portant les segments : (a) principe de rectangle approximatif de la région [Grompone von Gioi 2010], (b) test $k(r, x)$, (c) résultat de calcul de régions sur l'image du montage d'usinage

Enfin, pour valider un rectangle, un indice de confiance est calculé sur la direction des gradients par le quotient du nombre de pixels dans une direction, appelés *pixels alignés*, dont le nombre est donné par k , par rapport au nombre total des pixels dans ce rectangle, donnés par n . Ce test est appelé le test $k(r, x)$ exprimant le nombre des pixels alignés dans le rectangle r par rapport à l'image localisée en x . Si $k(r, x) \geq k_r$, avec k_r un seuil dépendant de chaque rectangle (longueur, largeur et orientation), le rectangle est accepté. Selon Desolneux *et al.* [Desolneux 2000] k_r doit être fixé de telle sorte qu'il minimise les pixels non-alignés ($n - k$), appelés les fausses alarmes (**NFA**). Dans le cas de détection de segment de droite, le nombre de pixels alignés doit être supérieur au nombre des pixels non-alignés, et le (**NFA**) est donné par l'équation (3.16).

$$NFA(r, x) = N_{test} \cdot \mathbb{P}_{H_0} [k(r, X) \geq k(r, x)] \quad (3.16)$$

On pose : H_0 l'hypothèse que l'image délimitée par le rectangle r ne contient pas de pixels alignés, N_{test} le nombre de test⁶. Nous pouvons prendre P_{H_0} la probabilité qu'un pixel, sous l'hypothèse H_0 , soit aligné avec le rectangle r . Cette probabilité est donnée par $p = \frac{\tau}{\pi}$ pour une distribution uniforme entre $[0, 2\pi]$. Ce qui conduit que $\mathbb{P}_{H_0} [k(r, X) \geq k(r, x)]$ peut être écrite sous la forme d'une loi binomiale, donnée par l'équation (3.17).

$$\mathbb{P}_{H_0} [k(r, X) \geq k(r, x)] = B(n(r), k(r, x), p) \quad (3.17)$$

6. Pour une image de taille $N \times M$ il y a $(N \times M)^{5/2}$ rectangles possibles

Où,

$$B(n, k, p) = \sum_{j=k}^n \binom{n}{j} p^j (1-p)^{n-j} \quad (3.18)$$

Pour plus de détails, le lecteur peut se référer au travail de Rafael Grompone von Gioi *et al.* [Grompone von Gioi 2010].

Cette méthode donne de meilleurs résultats en écartant les segments isolés. A chaque rectangle sont associés un segment et une droite support ce qui nous permet de trouver des points d'intersection significatifs de la topologie de la scène. En exploitant la totalité des segments, donc des droites, le nombre de points d'intersection générés est trop important et ne tient pas compte d'une proximité d'orientation et de distance des droites support. Pour réaliser un filtrage afin de prendre les points les plus significatifs de la géométrie, nous proposons d'appliquer un nouvel algorithme.

Notre objectif est d'intégrer l'information de la direction du segment pour regrouper et réaliser une meilleure approximation des formes droites, afin d'optimiser les points d'intersection, que nous exploitons par la suite pour la construction des graphes. Un exemple de cas est montré dans la figure 3.17.a par une ellipse jaune. Ces régions mentionnées portent la même direction θ avec la tolérance τ_θ . L'ensemble des droites représentées dans la figure 3.17.b de cette région peut être réduit à une droite plus significative, comme le montre l'exemple de la figure 3.17.b. Rappelons que les images des montages d'usinage présentent souvent des formes polyédriques : rectangle, carré, polygone, ce qui conforte ce choix.

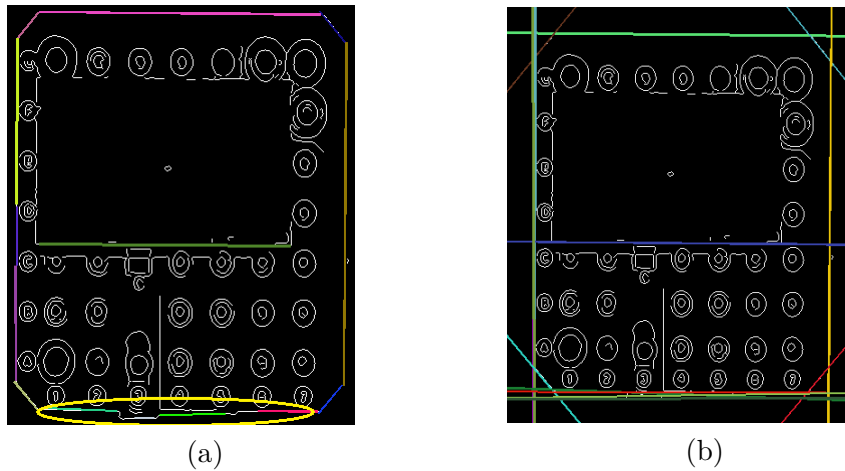
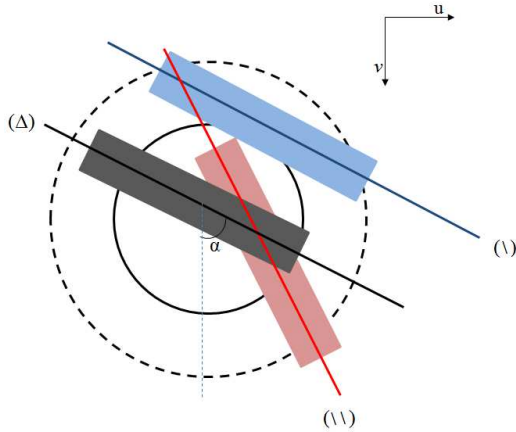


Figure 3.17 – Détection des segments de droites : (a) les segments ; l'ellipse jaune représente une région possédant des segments qui pourraient être représentés par une seule droite (b) : les droites portées par les segments.

Les principes de critères d'orientation et de distance du rectangle sont donnés sur les figures 3.19 et 3.18, respectivement.

Un segment centré en x est fusionné avec le rectangle R_Δ si la direction est égale à la direction du R_Δ , et si sa distance $d \leq 3l$, avec (l) la largeur du rectangle portant



Un segment porté par un rectangle est défini par sa direction α par rapport à l'axe v et sa distance à l'origine.

Dans cette exemple, le rectangle portant le segment (\setminus) est regroupé avec le rectangle du segment (Δ) . Cependant, le rectangle du segment $(\setminus\setminus)$ est écarté, car son orientation $\theta_{(\setminus\setminus)} \notin [\alpha - \tau_\alpha, \alpha + \tau_\alpha]$.

Figure 3.18 – Critère d'orientation (α)

le rectangle R_Δ . La distance $d(\overline{AB}, x)$ entre les 2 rectangles est calculée par rapport aux droites support, avec $\overline{AB} = L$, la longueur du rectangle et x le point centre du rectangle. p la projection du point x sur l'axe porté par le segment \overline{AB} . Un exemple est donné sur la figure 3.19.

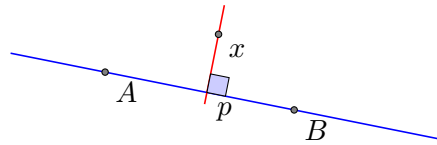


Figure 3.19 – Critère de distance (d)

Le nombre de rectangles R_i est donné, initialement, selon le nombre de régions. Après fusionnement, pour chaque rectangle R_j avec $j \leq i$, l'ensemble des droites dans chaque rectangle est remplacé par une seule droite en résolvant les équations des droites au sens des moindres carrés. Ce processus a permis de sélectionner la meilleure droite représentant les rectangles délimitant les régions des segments.

Reprenons l'exemple de la figure 3.18. L'objectif est de fusionner le rectangle portant le segment (\setminus) et/ou $(\setminus\setminus)$ avec le rectangle de référence R_Δ donné par le segment Δ . On tire aléatoirement un rectangle présentant un segment appelé Δ défini par ses terminaux x_i, x_j donnés par les longueurs du rectangle $x_c - \frac{L}{2}$ et $x_c + \frac{L}{2}$ et son orientation θ . Pour comparer les autres segments par rapport à R_Δ la procédure est la suivante :

- Calculer l'orientation θ de la droite (\setminus) par rapport à l'axe v . Si $\alpha > \pi \Rightarrow \theta = \theta - \pi$.
- Calculer la distance d du point x par rapport à l'axe porté par le segment \overline{AB} , comme le montre la figure 3.19.
- Si la distance d et la direction θ respectent les deux conditions :

- La distance $d \leq 3l$
- La direction $\alpha - \tau_\alpha \leq \theta \leq \alpha + \tau_\alpha$

le rectangle portant la droite (\setminus) est fusionné avec le rectangle $R_{(\Delta)}$.

Pour chaque région, si le nombre des segments dans cette région est supérieur à 1, une droite est calculée au sens de moindres carrés par résolution des équations des droites, afin de remplacer toutes ces formes droites par une seule droite approximative et plus significative.

Ensuite, les points d'intersections sont identifiés. Si l'intersection est en dehors du rectangle englobant le montage d'usinage dans l'image, le point n'est pas pris en compte, sinon le point est considéré. Le processus de regroupement est décrit dans l'algorithme suivant.

Algorithme 1 Regroupement des segments de droites

Précondition : $N > 0$: nombre de classes. l, L, α et θ

Postcondition : $\alpha \in [0, \pi]$, m : classe de référence.

```

n ← 1
pour tout n ≤ N faire
  si  $\alpha_n - \tau_\alpha \leq \theta_n \leq \alpha_n + \tau_\alpha$  alors
    si  $d_m \leq 3l$  alors
       $R_m \leftarrow (\setminus)_n$ 
retour les droites de la classe m

```

Le résultat du regroupement et les points d'intersection des droites sont illustrés sur la figure 3.20. Ce sont ces points qui vont servir pour la construction du graphe de proximité.

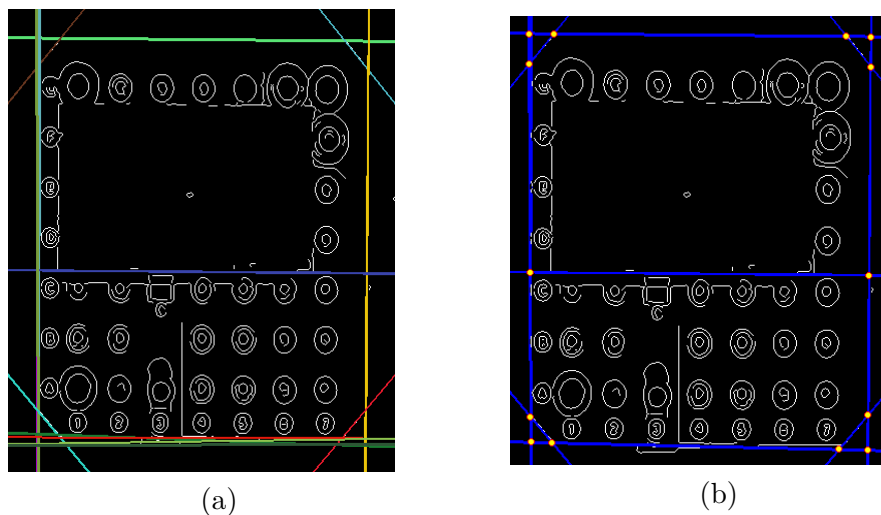


Figure 3.20 – Sélection des coins : (a) les droites des segments, (b) résultat de calcul obtenu après fusion des régions. Les intersections des droites sont mentionnées en jaune.

Les points d'intersection des droites support n'appartiennent pas obligatoirement aux contours des objets, ils restent cependant significatifs de la topologie. Sur la figure 3.20 on peut remarquer les 3 points des coins du support d'usinage dont 1 n'est pas sur l'objet.

Nous avons vu, dans cette section 3.3, la description des processus de détection et de sélection des primitives 2D. Dans la section suivante, section 3.4, nous nous intéressons à la sélection des primitives 3D à partir du modèle CAO.

3.4 Sélection des primitives 3D

Nous supposons ici que les positions des éléments du montage d'usinage sont connues *a priori*. Cette connaissance est donnée par les transformations géométriques reliant les éléments du montage d'usinage définis dans leurs propre repère, i.e. R_{pion1} , R_{pion2} , R_{brut} , R_{bride} , etc. avec le repère du montage d'usinage choisi par le concepteur ($R_{Montage}$), comme le montre la figure 3.21. Nous utilisons ici la description surfacique donnée par le fichier *.IGES *Initial Graphics Exchange Specification* pour sa capacité descriptive et l'exactitude de sa représentation par des surfaces NURBS. Nous proposons de créer une image théorique de cette représentation qui va nous servir de support 2D de référence.

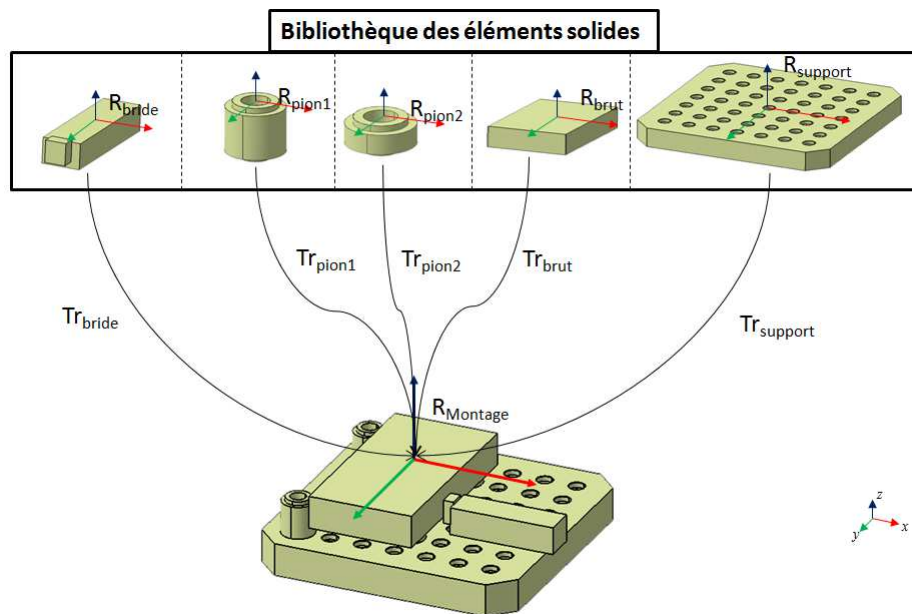


Figure 3.21 – Reconnaissance des éléments par transformation géométrique

3.4.1 Création de l'image théorique

Cette étape de création de l'image théorique est importante pour l'établissement d'une correspondance entre l'image réelle du montage d'usinage issue de la caméra

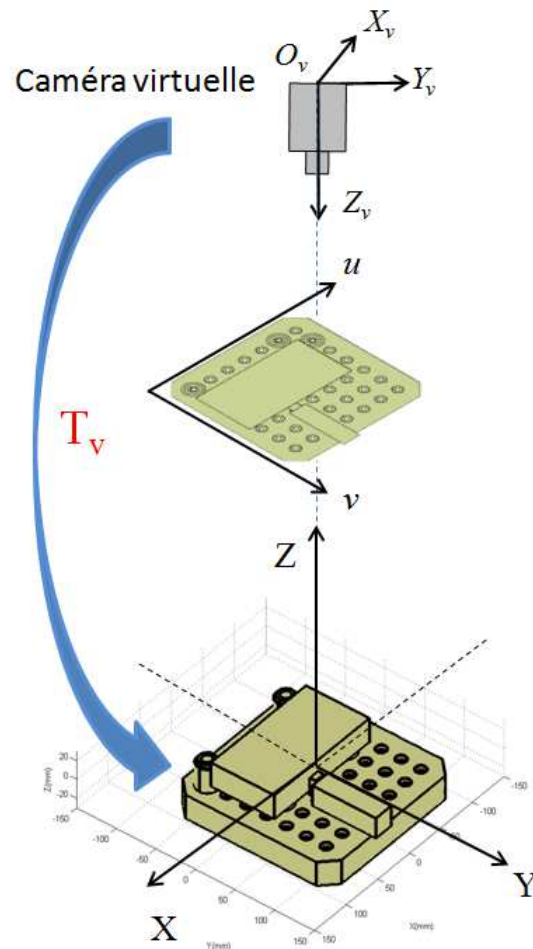


Figure 3.22 – Pose virtuelle donnée par la matrice $[T_v]$

placée sur le porte outil, et son modèle CAO 3D. De plus les deux chaînes réelle et théorique subissent les mêmes traitements d'images, e.g. détection et sélection des primitives, ce qui permet de réduire le développement logiciel et les lignes de code. Comme nous l'avons abordé auparavant, nous proposons, afin d'assurer une meilleure visibilité de la scène, de placer la caméra au dessus du montage d'usinage selon la transformation donnée par les angles entre la broche et le plateau de la machine, section 2.6. **La même hypothèse est posée pour créer l'image théorique.** Une caméra théorique est placée au dessus du modèle numérique 3D du montage d'usinage permettant une projection de ce dernier sur le plan (uv) de l'image synthétique. Cette relation entre les repères de la caméra théorique et le modèle CAO 3D du montage d'usinage est donnée par la *pose théorique* T_v , comme le montre la figure 3.22. En réalité, T_v n'est qu'une matrice de projection de taille (4×4) en coordonnées homogènes. Cette transformation va permettre d'établir un lien direct entre l'image 2D et le modèle CAO 3D.

En effet, nous devons créer une image dont la transformation perspective doit

inclure 3 paramètres [Klette 1998] :

- Un point 3D P donné par le modèle CAO doit être projeté sur le plan image selon la transformation projective donnée par la matrice de projection T_v , exprimant la relation géométrique entre la caméra et l'objet.
- Un point 2D p projeté sur le plan image doit être exprimé par rapport aux coordonnées de l'image selon la relation entre repère image et repère plan (uv).
- Enfin, les points 2D projetés sur l'image sont supposés non distordus. Les deux distorsions, tangentielle et radiale ne sont pas considérées, car l'image réelle issue de la caméra est corrigée en distorsion après une étape d'estimation hors ligne des paramètres intrinsèques de la caméra. Autrement dit, la caméra est supposée calibrée.

Soit P_s un point 3D donné par une arête du modèle CAO dont les coordonnées homogènes sont $(X_s, Y_s, Z_s, 1)^T$ dans le repère lié au montage (R_m). Le point P_c donné par ses coordonnées $(X_c, Y_c, Z_c, 1)^T$ dans le repère caméra (R_c).

La relation entre P_s et P_c est donnée par la pose virtuelle T_v .

$$P_c = T_v P_s \quad (3.19)$$

La matrice de pose T_v est composée par une matrice de rotation $Rot(3 \times 3)$ et un vecteur de translation $Tr(3 \times 1)$, comme le montre l'équation (3.20). Dans la conception mécanique des montages d'usinage, le plan de référence $Z = 0$ est souvent confondu avec le repère du brut à usiner, car le calcul de programme d'usinage (FAO) s'effectue souvent sur le brut.

$$\begin{bmatrix} P_c \end{bmatrix} = \begin{bmatrix} Rot_{(3 \times 3)} & Tr_{réf(3 \times 1)} \\ 0_{(1 \times 3)} & 1 \end{bmatrix} \cdot \begin{bmatrix} P_s \end{bmatrix} \quad (3.20)$$

Afin de respecter l'occultation dans la création de l'image théorique, nous utilisons la définition surfacique du modèle CAO, avec comme plan de référence le plan $Z_{max} = max(Z)$. Les coordonnées des points projetés P_s^N après calcul de la transformation T_v sont données par l'équation (3.21) pour $P_{im}(x_{im}, y_{im}, 1)$ le point projeté dans le plan image (uv).

$$\begin{bmatrix} P_{im} \end{bmatrix} = A \cdot \begin{bmatrix} Rot_{(3 \times 3)} & Tr_{réf(3 \times 1)} \\ 0_{(1 \times 3)} & 1 \end{bmatrix} \cdot \begin{bmatrix} P_c \end{bmatrix} \quad (3.21)$$

Avec $A(4 \times 4)$ la matrice des paramètres intrinsèques. Ils sont donnés par la matrice A de la caméra réelle. Les monographies de Klette R. *et al.* [Klette 1998] et Hartley R. [Hartley 2004] présentent plus en détails la géométrie de l'image et les différentes applications dans le domaine de la vision par ordinateur. Le résultat de la création de l'image théorique est illustré sur la figure 3.23.

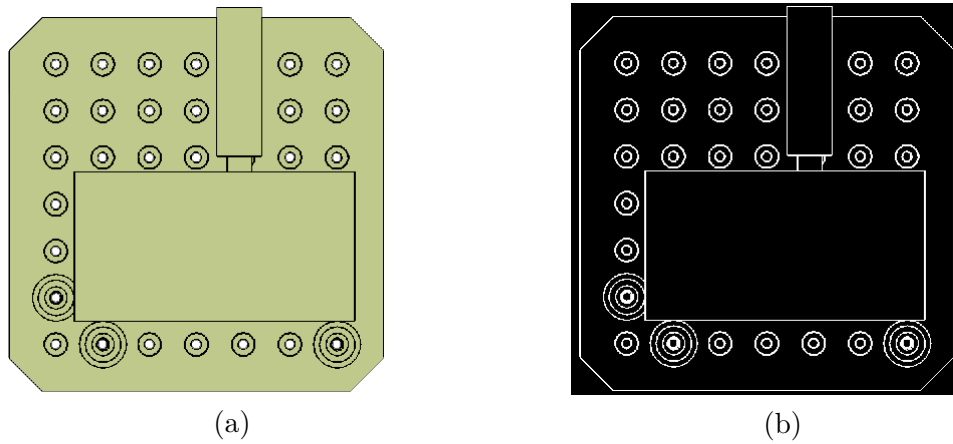


Figure 3.23 – Image théorique du modèle CAO, (a) image du montage d'usinage, (b) image des contours

Le traitement de la création de l'image théorique a été opéré avant le calcul de la pose réelle et l'établissement de la relation géométrique entre la caméra réelle et le montage d'usinage installé sur la machine, ce qui pourrait introduire un problème du facteur d'échelle. Pour contourner cette dernière difficulté, nous proposons de **normaliser indépendamment** les coordonnées des primitives sélectionnés, avant de procéder à la mise en correspondance des graphes. Cette étape est développée dans le chapitre 4.

3.4.2 Extraction des centres des cercles 3D

Les traitements opérés sur l'image théorique peuvent être relativement alourdis si nous utilisons le modèle 3D CAO, raison indispensable dans l'étape de calcul des centres des cercles. Le même processus de sélection des primitives 2D dans l'image réelle est appliqué pour l'image théorique. Nous utilisons la transformée de Hough (TH) pour le calcul des cercles sur l'image théorique. Ces cercles trouvent leurs correspondants dans le modèle CAO 3D à partir de la relation (3.21). La figure 3.24 montre la correspondance entre les cercles 2D calculés sur l'image théorique et leurs correspondant 3D dans le modèle 3D CAO.

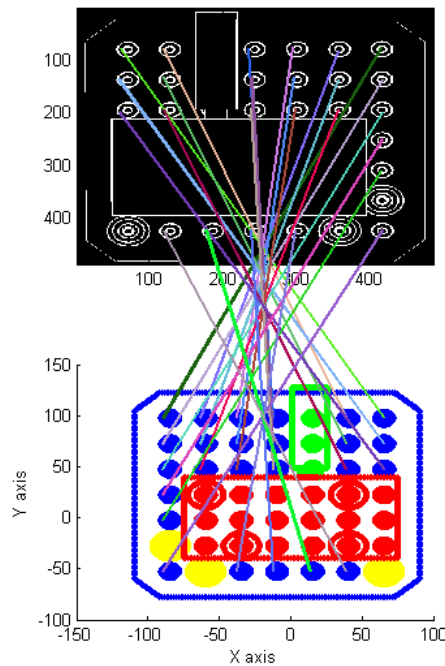


Figure 3.24 – *Extraction des centres des cercles 3D à partir de l'image théorique 2D et leurs correspondants dans le modèle CAO.*

3.4.3 Extraction des points 3D d'intersection des droites

Les points d'intersection sont donnés par la résolution des équations des droites 2 à 2 pour chaque plan Z_i dans le modèle CAO. Pour le calcul des points d'intersection des droites sur l'image théorique, nous utilisons l'algorithme *LSD* [Grompone von Gioi 2010]. Les correspondances sont données par la transformation T_v entre les intersections 2D calculés à partir de l'image théorique et les points d'intersections 3D donnés par le modèle CAO, comme le montre la figure 3.25.

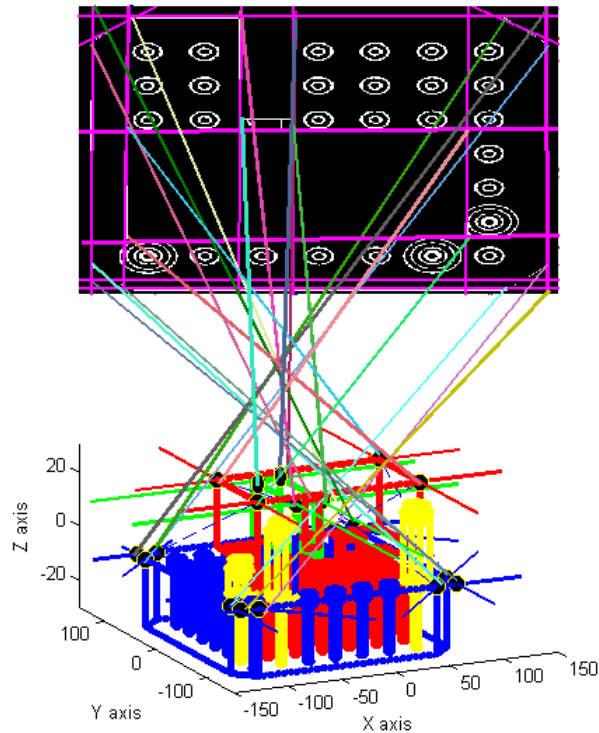


Figure 3.25 – *Extraction des points d'intersection 3D à partir de l'image théorique 2D et leurs correspondants dans le modèle CAO.*

3.5 Conclusion

Ce chapitre a présenté notre travail mené pour le calcul de l'image des contours et la sélection des primitives 2D/3D données par les centres de cercles et les intersections des droites, pour la construction des graphes.

L'image des contours est calculée à partir de la modélisation explicite de l'histogramme du module des gradients. Cette méthode a montré son efficacité par rapport à quelques indicateurs de performance utilisés dans la communauté de la vision par ordinateur. Pour sélectionner les primitives pertinentes, nous avons proposé des critères de sélection à la fois pour la sélection des cercles et des droites. Ces critères de sélection ont permis de calculer les points caractérisant au mieux la topologie de la scène.

Enfin, nous proposons d'utiliser les graphes de proximité comme élément de représentation de la scène. En effet, les graphes construits vont être utilisés pour mettre en correspondance les primitives sélectionnées pour le calcul de pose et le recalage du modèle CAO sur l'image du montage d'usinage réalisé, afin de déterminer les zones d'intérêt représentant les emplacements attendus des éléments de bridage. Le chapitre 4 présente le travail mené.

Bibliographie

- [Beveridge 1996] J.R. Beveridge, C. Graves et C. Leshner. Some Lessons Learned from Coding the Burns Line Extraction Algorithm in the Darpa Image Understanding Environment. Technical Report CS-96-125, Computer Science Dept. Colorado State Univ., 1996. (Cité en page 53.)
- [Burns 1986] J.B. Burns, A.R. Hanson et E.M. Riseman. Extracting Straight Lines. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 8, no. 4, pages 425–455, 1986. (Cité en page 53.)
- [Debashis 2010] S. Debashis et K. Pal Sankar. Gradient histogram : thresholding in region of interest for edge detection. Image and vision computing journal, vol. 28, pages 677–695, 2010. (Cité en pages 37, 38, 39 et 40.)
- [Deschenes 2004] F. Deschenes, D. Ziou et M.-F. Auclair-Fortier. Detection of lines, line junctions and line terminations. International Journal of Remote Sensing, vol. 25, no. 3, pages 511–535, feb 2004. (Cité en page 53.)
- [Desolneux 2000] A. Desolneux, L. Moisan et J.-M. Morel. Meaningful Alignments. International Journal of Computer Vision, vol. 40, no. 1, pages 7–23, 2000. (Cité en page 54.)
- [Frei 1977] W. Frei et C. Chen. Fast boundary detection : A generalization and a new algorithm. IEEE Trans.on Computer, vol. 26, no. 10, pages 988–998, 1977. (Cité en page 53.)
- [Grompone von Gioi 2010] R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel et G. Randall. LSD : A Fast Line Segment Detector with a False Detection Control. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 32, no. 4, pages 722–732, 2010. (Cité en pages ix, 52, 53, 54, 55 et 62.)
- [Hartley 2004] R. I. Hartley et A. Zisserman. Multiple View Geometry in Computer Vision. 2004. (Cité en page 60.)
- [Huijsmans 1993] D.P. Huijsmans et N. Sebe. Extended performance graphs for cluster retrieval. IEEE Conference on Computer Vision and Pattern Recognition (CV-PR'01), feb Dec.1993. (Cité en page 45.)
- [Kapur 1985] J. N. Kapur, P. K. Sahoo et K. C. Wong. A new method for gray-level picture thresholding using the entropy of the histogram. Graph. Models Image Process., vol. 29, pages 273–285, 1985. (Cité en pages viii, 38, 40, 47 et 49.)
- [Karabagli 2011] B. Karabagli, T. Simon et Orteu J.-J. Détermination automatique du seuil de binarisation des modules des gradients par modélisation de leur histogramme. In Proceedings of ORASIS 2011, Praz-sur-Arly, France, pages 2–8, juin 2011. (Cité en page 38.)
- [Kass 1988] M. Kass, A. Witkin et D. Terzopoulos. Snakes : Active contour models. International Journal of Computer Vision 1, pages 321–331, 1988. (Cité en pages 37 et 38.)

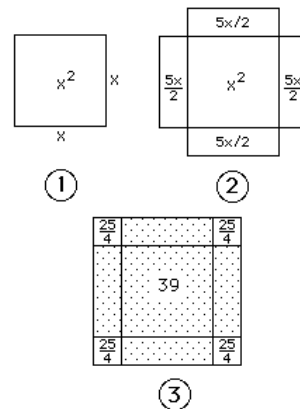
- [Kimmel 2003] R. Kimmel et AM. Bruckstein. Regularized Laplacian Zero Crossings as Optimal Edge Integrators. *International Journal of Computer Vision*, vol. 53, no. 3, pages 225–243, 2003. (Cité en page 38.)
- [Klette 1998] R. Klette, K. Schlûns et A. Koschan. COMPUTER VISION : Three-Dimensional Data from Images. page 392, 1998. (Cité en page 60.)
- [Ku 2005] C.L. Ku. On improvement of the computation speed of Otsu's image thresholding. *J. Electron. Imaging.*, vol. 14(2)023011-023011-12., 2005. (Cité en pages viii, 37, 38, 39, 40, 47 et 49.)
- [Li 1999] Chang-Tsun Li et Lou Der-Chyuan. Edge detection based on the multiresolution Fourier transform. *IEEE Workshop on Signal Processing Systems, SiPS 99.*, pages 686–693, 1999. (Cité en page 38.)
- [Lu 1989] Y. Lu et R.C. Jain. Behavior of edges in scale space. *IEEE T-PAMI*, vol. 11 (4), pages 337–356, 1989. (Cité en page 38.)
- [Maire 2009] Michael Randolph Maire. Contour Detection and Image Segmentation. PhD thesis, UNIVERSITY OF CALIFORNIA, BERKELEY., 2009. (Cité en page 46.)
- [Namuduri 1992] K. R. Namuduri, R. Mehrotra et N. Ranganathan. Edge Detection Models Based on Gabor Filters. *International Conference on Image, Speech and Signal Analysis*, vol. 3, 1992. (Cité en page 37.)
- [Papari 2011] G. Papari, N. Petkov et W. C. Regli. Edge and line oriented contour detection : State of the art. *Image and Vision Computing*, vol. 29, pages 79–103, 2011. (Cité en pages viii, 37 et 38.)
- [Simon 1994] T. Simon, M Khoudeir, B. Heit et J. Bernard. Image segmentation by setting thresholds for the magnitude of gradient. *ICECS'94-IEEE Electronics, Circuits and Systems*, pages 40–45, 1994. (Cité en pages 37 et 40.)
- [Swets 1996] J.A. Swets. Signal detection theory and ROC analysis in psychology and diagnostics. *Collected papers*. Mahwah, NJ : Erlbaum., 1996. (Cité en page 45.)

ESTIMATION AUTOMATIQUE DE POSE PAR APPARIEMENT DE GRAPHES

To solve the equation $x^2 + 10x = 39$, El Kwarismi writes^a : "... a square and 10 roots are equal to 39 units. The question therefore in this type of equation is about as follows : what is the square which combined with ten of its roots will give a sum total of 39 ? The manner of solving this type of equation is to take one-half of the roots just mentioned. Now the roots in the problem before us are 10. Therefore take 5, which multiplied by itself gives 25, an amount which you add to 39 giving 64. Having taken then the square root of this which is 8, subtract from it half the roots, 5 leaving 3. The number three therefore represents one root of this square, which itself, of course is 9. Nine therefore gives the square." Al-Khwarizmi (780 - 850)

^a. F Rosen (trs.), Muhammad ibn Musa Al-Khwarizmi, Algebra (London, 1831). http://www-groups.dcs.st-andrews.ac.uk/~history/HistTopics/Arabic_mathematics.html

al-Khwarizmi completes the square



Sommaire

4.1 Introduction et problématique	68
4.2 Notions sur les graphes	71
4.2.1 Définition d'un graphe	71
4.2.2 Matrice d'adjacence	72
4.2.3 Isomorphisme	72
4.3 Construction des graphes	73
4.4 Appariement de graphes pour le calcul de pose : Notre proposition	77
4.4.1 Calcul des sous graphes	78
4.4.2 Détermination de la relation géométrique Φ	79
4.5 Calcul de pose et extraction des régions d'intérêts	87
4.5.1 Modélisation de la caméra	87
4.5.2 Détermination des paramètres intrinsèques de la caméra	88
4.5.3 Calcul des paramètres extrinsèques de la caméra	89
4.6 Évaluation de la méthode	92
4.6.1 Évaluation de la méthode par rapport au bruit	96
4.6.2 Évaluation de la méthode par rapport à la topologie des graphes	98
4.7 Vérification de la pose et cas limites	101
4.7.1 L'impact de la matrice de rotation	101
4.7.2 L'impact du vecteur de translation	103
4.7.3 Cas du montage totalement différent du modèle numérique	104
4.8 Conclusion	105
Bibliographie	106

Mots clés : Graphe de proximité, Mise en correspondance automatique des graphes, Calcul automatique de pose, Extraction des ROI.

4.1 Introduction et problématique

Ce chapitre est structuré en 2 grandes parties, comme le montre le diagramme de la figure 4.1. La première partie traite du problème de la mise en correspondance automatique entre les primitives 2D calculées dans l'image et les primitives 3D du modèle CAO. La seconde partie, s'intéresse au calcul de pose de la caméra par rapport au montage d'usinage, afin de projeter le modèle de référence sur l'image du montage réalisé. Nous produisons ainsi une **réalité virtuelle** en projetant le modèle CAO sur l'image. L'objectif de cette étape est de pouvoir segmenter l'image en régions d'intérêt représentant les endroits "*prévus*" des éléments de bridage selon leur organisation topologique donnée en CAO. En effet, l'estimation 2D/3D de pose et la segmentation d'image sont deux domaines complémentaires, la connaissance de l'une peut servir

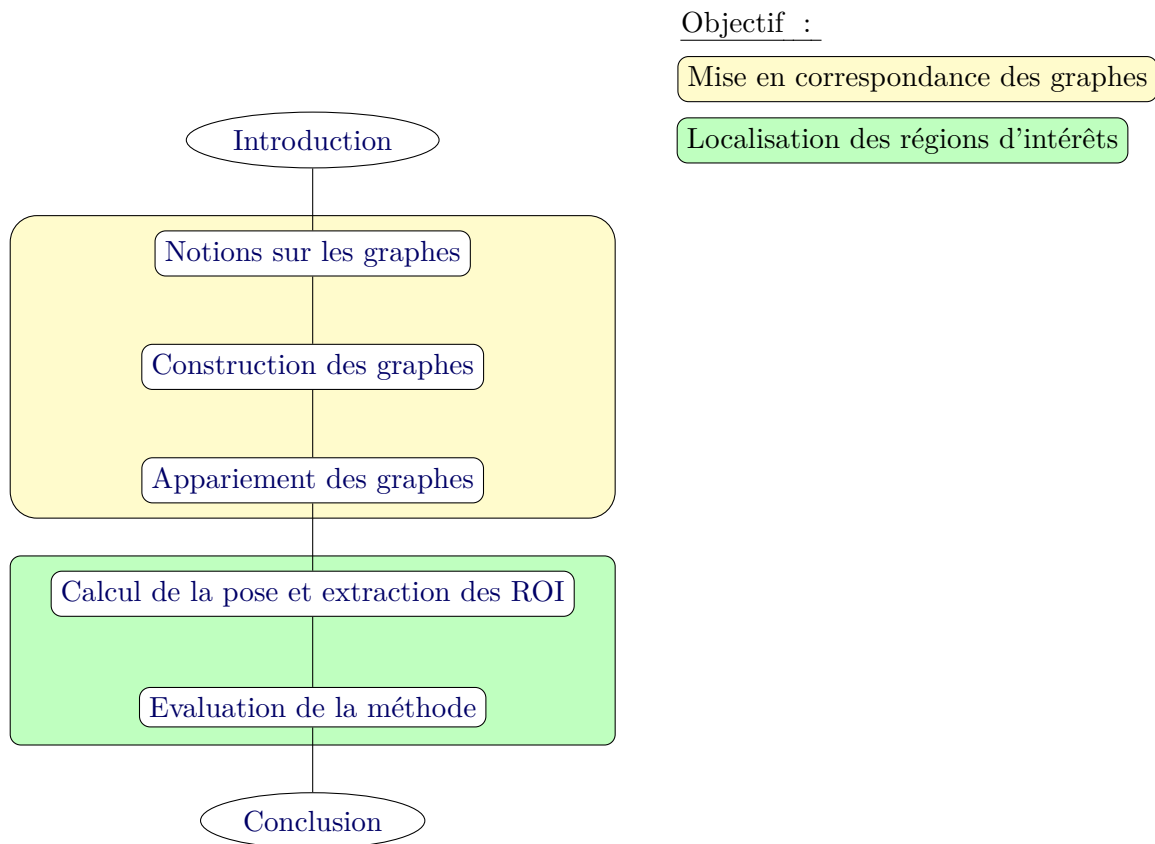


Figure 4.1 – Organigramme du chapitre 4 et objectif de chaque partie.

pour déterminer l'autre. Les travaux de Sandhu R. *et al.* [Sandhu 2011], et Brox T. *et al.* [Brox 2005] en donnent des exemples.

La mise en correspondance théorique/réel, 2D/3D en utilisant les graphes est un élément clef de la stratégie de l'analyse de conformité des montages d'usinage. Certains travaux de mise en correspondance peuvent être regroupés, **selon l'information image utilisée**, en méthodes utilisant l'information de luminance [Lowe 2004], [Schmid 1997], méthodes utilisant l'information spatiale [Caelli 2004], [Zheng 2006] et méthodes, plus récentes, utilisant la combinaison des deux [Duchenne 2011] pour la reconnaissance des objets déformables.

La méthode de mise en correspondance des graphes que nous proposons est fondée sur une approche topologique, car, d'une part, les montages d'usinage sont des objets rigides qui ne peuvent pas subir une déformation mécanique importante, et d'autre part, l'image théorique calculée ne peut pas donner une information de luminance comparable à l'image issue de la caméra. Seuls les contours peuvent être exploités. Ces contours portent l'ensemble de l'information topologique, géométrique et dimensionnelle, et ils trouvent leurs correspondants dans la définition CAO par les arêtes des surfaces d'objets utilisées dans la conception du montage.

Les travaux de mise en correspondance **utilisant les graphes** en vision par ordi-

nateur peuvent être scindés en deux groupes : 1- la mise en correspondance des graphes dits "exactes", et 2- la mise en correspondance des graphes dits "inexactes". Le premier groupe englobe les méthodes répondant aux problèmes de mise en correspondance avec les propriétés liées aux isomorphismes [Gori 2005], [Dong-Qing 2004],[Wu 1993]. Généralement, cela revient à déterminer une fonction $f : V_g \rightarrow V_h$ entre le graphe $G_g = (V_g, E_g)$ et le graphe $G_h = (V_h, E_h)$, avec $|V_g| = |V_h|$ ¹, tel que $(u,v) \in E_g$ ssi $(f(u),f(v)) \in E_h$. $|\cdot|$ signifie la taille du graphe, donnée par le nombre de nœuds. Le second groupe traite de la mise en correspondance quand $|V_g| \neq |V_h|$, [Bunke 1983] et [Roberto M. 2005].

Nous utilisons les graphes exacts pour la détermination des emplacements des éléments de la scène, et les graphes inexactes pour la mise en correspondance des primitives pour le calcul de la pose, donnés par les graphes de proximité, *Graphes de Gabriel* [Gabriel 1969].

La détermination de la relation géométrique entre les primitives 2D/3D est définie par un calcul de pose. Le calcul de pose consiste à trouver la rotation et la translation de la caméra par rapport à la scène observée, en utilisant l'image 2D et son modèle 3D défini en CAO. Nous nous intéressons dans ce travail à l'estimation de la pose à partir d'**une seule acquisition d'image**. Cette relation géométrique est exprimée par une matrice de projection P de dimension (4×4) en coordonnées homogènes.

De nombreux algorithmes existent pour le calcul de pose, et se distinguent, selon la méthode de calcul utilisée, en méthodes linéaires, de décomposition matricielle [Bouguet 2004], [Tsai 1986], ou méthodes non linéaires [Dementhon 1995], [Zhang 2000] proposant de multiples solutions, selon le nombre de points mis en correspondance [Cox 2005], [Quan 1999]. Ces méthodes non-linéaires, proposent souvent une approche itérative. L'algorithme **POSIT** de [Dementhon 1995], par exemple, s'appuie sur deux étapes. La première étape, **POS : Pose from orthography and Scaling** utilise une approche linéaire en approximant la projection perspective par une projection orthogonale, afin de déterminer la rotation et la translation, en résolvant un système d'équations linéaires. La seconde partie **POSIT : POS with Iteration** utilise à chaque itération les coordonnées des points estimés pour que la projection orthographique se rapproche d'une projection perspective. L'inconvénient majeur de cette méthode, est qu'elle diverge à partir d'un nombre important de points mis en correspondance, pour nos essais à partir d'une vingtaine de points. Le calcul de la pose et l'évaluation de la méthode proposée sont présentés dans les sections 4.5 et 4.6.

Avant de décrire notre méthode de mise en correspondance des graphes, nous proposons de commencer ce chapitre par la définition de quelques notions sur les graphes qui seront utilisées par la suite.

1. Dans certains cas, même si $|V_g| = |V_h|$ cela n'implique pas l'existence de l'isomorphisme, et le problème est traité par des algorithmes d'appariement des graphes inexactes.

4.2 Notions sur les graphes

L'image est une représentation 2D d'une scène photométrique ou théorique dont le contenu peut être décrit par un graphe dans lequel les nœuds représentent les primitives pertinentes et les arêtes les relations topologiques que l'on souhaite mettre en évidence. Un des principaux intérêts d'une telle représentation est la simplification de la scène induisant des temps de calcul plus faibles par la diminution de la complexité et de la quantité d'informations à traiter.

4.2.1 Définition d'un graphe

On appelle graphe $G = (V, E)$ dans sa forme de base, la donnée d'un ensemble dont les éléments sont appelés nœuds, notés V et une partie $E \subset V \times V$ symétrique* dont les éléments sont appelés arêtes, notés $e(i, j)$, reliant le nœud i avec le nœud j .

* Attention ! Cette propriété n'est pas légitime pour un graphe ayant un retour de nœud (*self-loop*) i.e. arc $e(i, i)$

L'ordre d'un graphe est donné par le nombre de nœuds de G , et représenté par $|G|$.

Deux nœuds dans G , nommés $u, v \in V$ sont appelés **Adjacents** ou **voisins** s'ils sont connectés par une arête $e \in E$ dont les extrémités sont u et v .

Graphiquement, les nœuds peuvent être représentés par des points ou des cercles, et les arêtes par des traits ou des arcs reliant les nœuds.

Les nœuds des graphes peuvent posséder une information simple d'étiquetage, labélisation binaire, pour nous, 0 : nœuds représentant des cercles, et 1 : nœuds représentant une intersection de droites. Dans ce cas, le graphe est appelé **graphe d'étiquetage**. Ce type de graphe est présenté dans la section 4.4. Ou bien, les nœuds contiennent plus d'informations, vecteur des caractéristiques, forme, dimension, etc. et le graphe dans ce cas est appelé **graphe des attributs**, *Attributed Relational Graph (ARG)*. Ce type de graphe est présenté dans la section 4.5.

Pour certaines applications les arêtes ont des directions. Dans ce cas, le graphe est appelé **graphe orienté**. De manière réciproque, le terme **arc** est utilisé pour remplacer l'arête quand le graphe est orienté. Le terme **arête** est utilisé quand le graphe est non-orienté.

Dans notre travail, les graphes non-orientés sont utilisés. Les graphes d'étiquetages² sont utilisés pour l'appariement des points de correspondances 2D/3D et le calcul de pose, les graphes des attributs sont utilisés pour la détermination des régions d'intérêt pour l'analyse locale.

Nous nous intéressons dans les parties suivantes, section 4.2.2 et section 4.2.3, à la définition de quelques propriétés des graphes que nous utilisons par la suite pour la mise en correspondance.

2. Pour la suite, le graphe d'étiquetage, utilisé pour l'appariement 2D/3D, est appelé tout simplement graphe.

4.2.2 Matrice d'adjacence

Considérons un graphe $G = (V, E)$ non-orienté, dont le nombre des nœuds est N . La matrice d'adjacence A du graphe G est une matrice $N \times N$ dont chaque élément $[A]_{ij}$ est égal à 1 si et seulement si $(i, j) \in E_A$ et égal à 0 ailleurs. La figure 4.2 illustre l'exemple d'un graphe et la matrice d'adjacence associée.

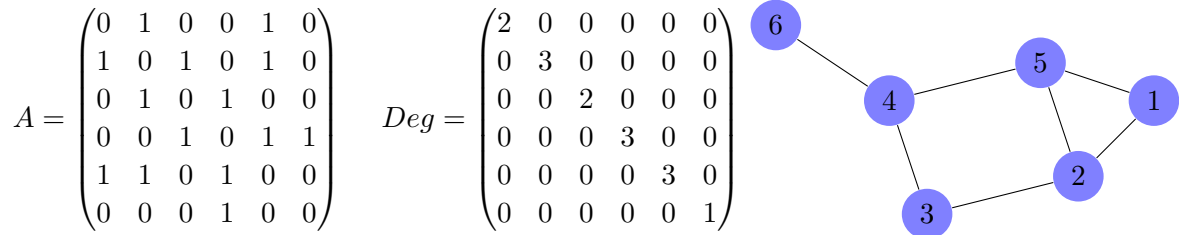


Figure 4.2 – Matrice d'adjacence et matrice de degré de connectivité des nœuds.

La matrice d'adjacence est un outil pertinent pour décrire le voisinage entre les nœuds et la topologie de la scène. Cependant, elle n'est pas suffisante pour certaines applications nécessitant une quantification du critère de similarité, i.e. le voisinage. Ce qui conduit naturellement, à une mesure quantitative [Medasani 2001] exprimée souvent par une distance, ou un poids compris entre $[0, 1]$, reflétant le score du voisinage donné.

D'autres propriétés peuvent aussi être utiles pour la mise en correspondance des graphes, notamment le **degré de connexité** d'un nœud, i.e. la matrice Deg de la figure 4.2. Le degré de connexité est défini comme étant le nombre d'arêtes reliant un nœud avec ces voisins. Un nœud avec un degré de connexité élevé reflète un nœud intéressant, qui pourrait définir une zone importante dans la scène. Nous reviendrons sur cette propriété dans la section 4.4.2 pour l'appariement des graphes.

4.2.3 Isomorphisme

Si l'isomorphisme existe entre deux graphes, cela signifie qu'ils sont **structurellement** semblables. Plus formellement, deux graphes sont *isomorphes* s'il existe une fonction bijective entre l'ensemble des nœuds, tel que si deux nœuds sont connectés dans le même graphe il restent connectés par transformation bijective. De plus, si les graphes G_a et G_b sont des graphes d'étiquetage ou d'attributs, l'isomorphisme entre eux doit garder les propriétés de chaque nœud, vecteur caractéristique et arc. La figure 4.3 montre un exemple de deux graphes isomorphes.

Il est à noter que l'isomorphisme entre les graphes de proximité issus des images réelle et théorique, n'est pas applicable pour l'appariement, car les données peuvent être manquantes due à la mauvaise détection des primitives 2D/3D ou à une définition incomplète en CAO. Nous devons donc prendre en compte ces manques dans le calcul de la mise en correspondance. Par ailleurs, les graphes des attributs, doivent être isomorphes, les éléments constituant le montage d'usinage doivent être géométriquement les mêmes et placés au bon endroit selon leurs représentations numériques données en

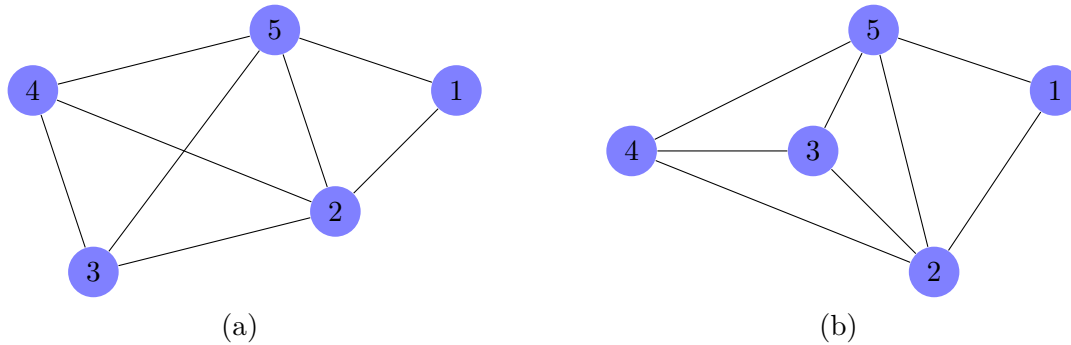


Figure 4.3 – Graphes isomorphismes : (a) Graphe non-orienté, (b) Graphe isomorphe au graphe de la figure (a).

CAO. Cette propriété peut être exploitée pour déterminer la précision de la pose, le paragraphe 4.7 présente un élément de solution.

Toutes ces propriétés reflètent la diversité de l'utilisation des graphes en vision par ordinateur, partant de la segmentation [Shi 2000] jusqu'à la reconnaissance de formes [Duchenne 2011], [Roberto M. 2005].

Pour nous, les graphes sont utilisés pour l'appariement et le calcul de pose. L'objectif est de trouver la relation géométrique Φ permettant d'assurer le meilleur appariement au sens de la distance minimale par rapport au nombre des nœuds appariés.

Pour construire les graphes, nous proposons une approche utilisant l'information spatiale dans laquelle le montage d'usinage décrit par ses primitives simples est représenté par un graphe qui définit les relations topologique entre ces derniers. La section 4.3 présente la construction des graphes, à la fois pour les primitives de l'image réelle et les primitives données par l'image théorique.

4.3 Construction des graphes

A partir des primitives sélectionnées, nous construisons les graphes de proximité représentant la scène, dans laquelle les centres des cercles et les intersections des droites représentent les nœuds du graphe. Il existe plusieurs critères de proximité, e.g. distance minimale, cercle de référence, etc. , permettant de caractériser les graphes de proximité, par exemple :

- **Graphe de sphère d'influence**, Sphere influence Graph (SIG), introduit par [Toussaint 1988]. Il est construit à partir de la notion d'intersection des cercles. En effet, pour chaque nœud $d_i(x, y)$ dans le graphe, on calcule sa distance par rapport aux autres nœuds dans le graphe. Le nœud le plus proche est donné par une distance minimale r_{min}^i désignant le rayon du cercle du nœud d_i . Pour chaque nœud d_i , nous traçons le cercle centré sur (x, y) avec le rayon r_{min}^i . Pour chaque paire de nœuds, si l'intersection des cercles existe, les deux nœuds sont considérés voisins. Ce voisinage est marqué par une arête dans le graphe. La figure 4.4.a montre le principe.

- **Graphe de Delaunay**, introduit par [Delaunay 1928], appelé triangulation de Delaunay, figure 4.4.b. Son principe s'appuie sur la triangulation entre les nœuds. Le voisinage est défini par un triangle qui a comme sommets trois nœuds, tel que son cercle circonscrit ne contienne aucun autre nœud. L'assemblage de tels triangles de Delaunay pour les N nœuds forme le graphe de Delaunay. Une telle triangulation est une solution unique, sous condition qu'aucun des trois nœuds ne soient aligné, et que les quatre nœuds ne se trouvent pas sur le même cercle.

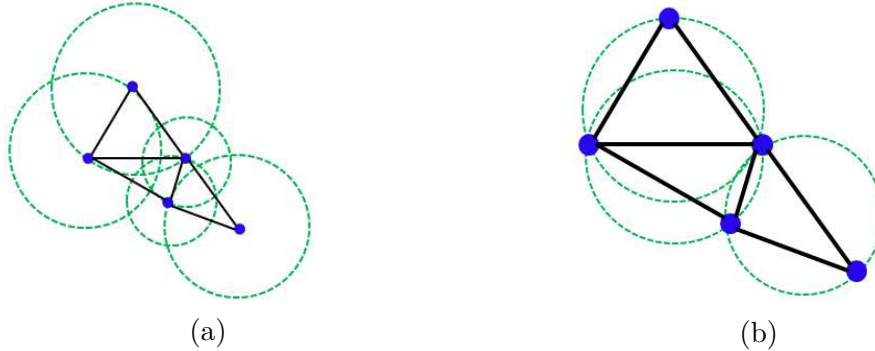


Figure 4.4 – Graphes de proximités : (a) principe de calcul du graphe SIG, (b) principe de calcul du graphe de Delaunay

- **Graphe de voisinage relatif**, *Relative Neighborhood Graph (RNG)* introduit par Lankford [Lankford 1969]. Il s'appuie sur la définition du voisinage relatif. Deux nœuds x_i, x_j sont définis relativement proches, si :

$$d(x_i, x_j) \leq \max d(x_i, x_l), d(x_j, x_l) \text{ pour } l \neq i, j$$

Avec $x_l, d(x_i, x_j) \leq \max d(x_i, x_l), d(x_j, x_l)$, si x_l ne se situe pas à l'intérieur de l'intersection des deux cercles centrés sur x_i et x_j avec les rayons $d(x_i, x_j)$. L'intersection est appelée voisinage relatif, et deux nœuds sont relativement voisins si leur voisinage relatif est vide.

- **Graphe de Gabriel**, *Gabriel Graph (GG)* [Gabriel 1969] est un sous graphe de la triangulation de Delaunay [Matula 1980], pour lequel deux nœuds x_i, x_j , sont considérés voisins, si un cercle passant par x_i, x_j dont le rayon est donné par $d(x_i, x_j)/2$ ne contient pas de nœud à l'intérieur du cercle. La figure 4.5 montre le principe du graphe de Gabriel et sa relation avec la triangulation de Delaunay.

Il existe d'autres types de graphes de proximité, i.e. *Closest Pair (CP)*, le voisin le plus proche *Nearest neighbors graph (NNG)*, etc. Le principe est souvent le même, mais ils se distinguent dans la définition du critère de voisinage, i.e. distance euclidienne, cercle de voisinage, etc.

Nous avons adopté les principes de construction des graphes inspirés des travaux de Gabriel *et al.*. En effet, pour chaque nœud, nous traçons un cercle dont le centre est donné par les coordonnées du nœud. Pour déterminer le voisinage, nous agrandissons le rayon du cercle autour de ce nœud jusqu'à ce qu'un autre nœud apparaisse à l'intérieur du cercle. Les deux nœuds sont alors considérés voisins, i.e. proches, et ils sont liés entre eux par une arête. Le processus se répète pour l'ensemble des nœuds.

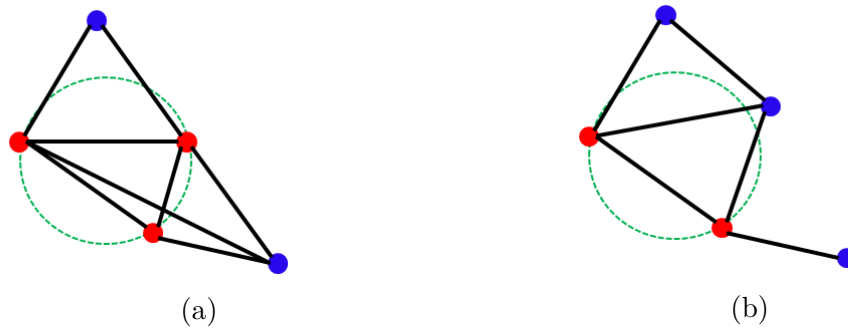


Figure 4.5 – Principe de calcul du graphe Gabriel : (a) graphe de Gabriel, (b) graphe de Delaunay

Le résultat du calcul des graphes de proximité est illustré sur la figure 4.6 pour le cas du montage Norelem conforme au modèle CAO.

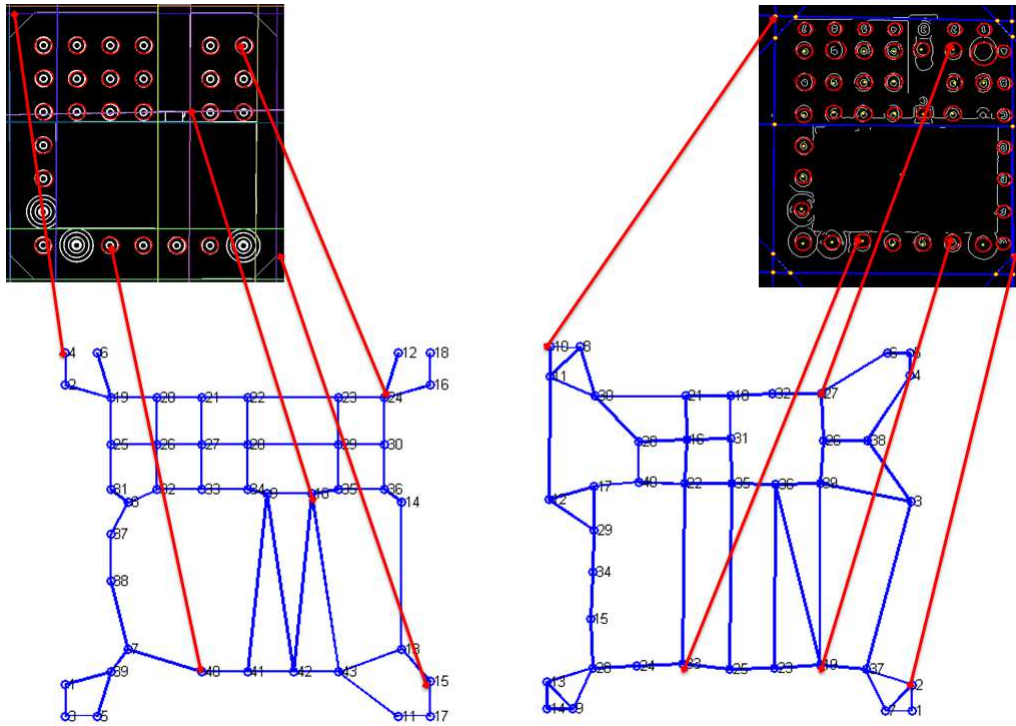


Figure 4.6 – Construction automatique des graphes : cas du montage conforme.

Pour différentes configurations non-conformes le résultat de construction des graphes est illustré sur la figure 4.7

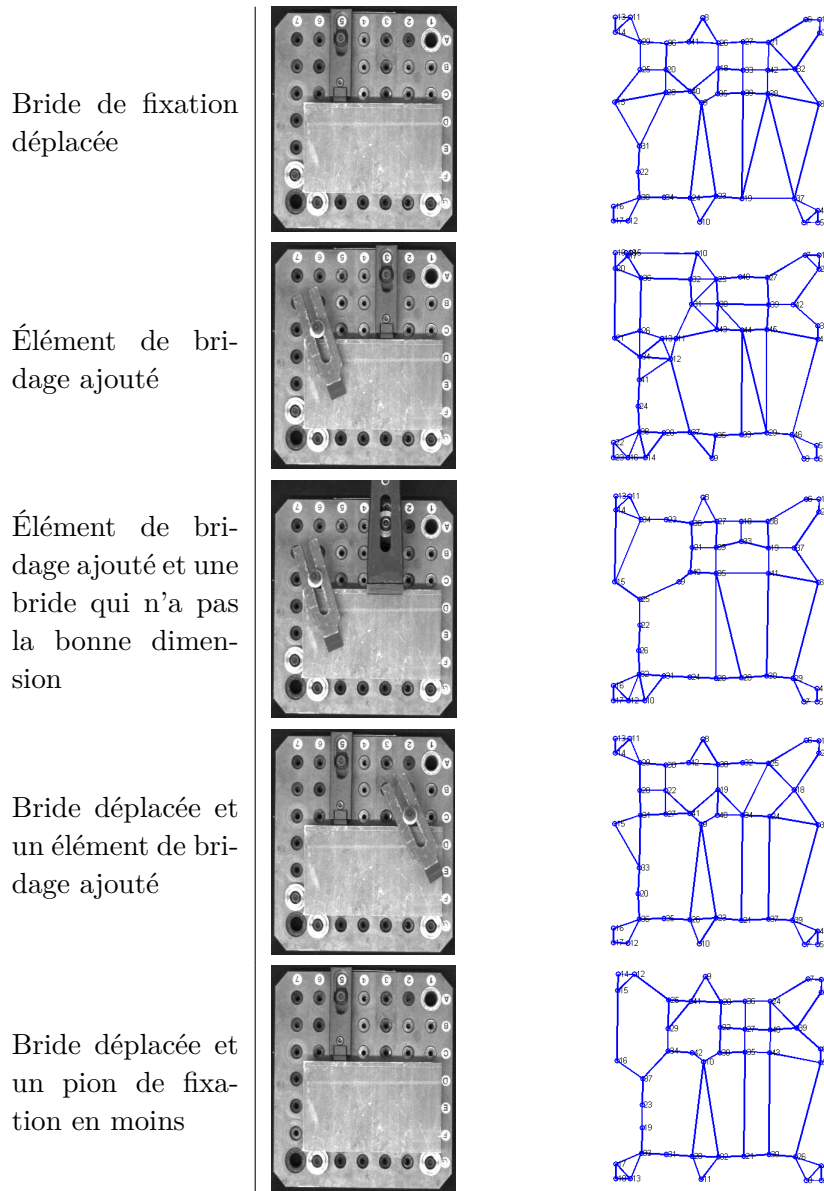


Figure 4.7 – Construction automatique des graphes pour différentes configurations non-conformes du montage d'usinage

4.4 Appariement de graphes pour le calcul de pose : Notre proposition

Différents indices peuvent être exploités à partir de l'image pour établir une mise en correspondance entre deux ensembles de points extraits de l'image. Cela nécessite une représentation significative des données. Nous nous intéressons dans cette section à l'algorithme proposé pour l'appariement entre le graphe (G) issu de l'image réelle du montage d'usinage, appelé **graphe réel**, et le graphe (H) issu de l'image théorique, appelé **graphe théorique**.

Certaines notions sont aussi importantes à prendre en compte dans la mise en correspondance. L'unicité, est un critère important, car les nœuds appariés vont servir pour le calcul de pose et le recalage du modèle CAO sur l'image. Ils ne peuvent donc être appariés qu'une seule fois. Néanmoins, les difficultés de l'extraction et la sélection de ces primitives, et l'incomplétude des ensembles de ces derniers pour la construction des graphes, i.e. fausse détection, etc. due aux performances des algorithmes de détection et sélection des primitives, ainsi que les conditions photométrique de la scène, nécessite de prendre en compte ce cas de figure dans la mise en correspondance. De plus, l'ensemble des primitives sélectionnées peut être différent pour une même configuration, cas du montage conforme donné par la figure 4.6, ce qui entraîne une contrainte supplémentaire dans le processus de la mise en correspondance. Par exemple, la bride présentant un trou au milieu peut déclencher une détection du cercle autour de ce trou, alarmes de type faux positif ou faux négatifs, ce qui n'est pas le cas dans le modèle CAO, car la surface supérieure de la bride est représentée par une surface pleine, cachant tout détail intérieur. Par conséquent, même si le montage réalisé est conforme au modèle CAO, les graphes construits pourraient ne pas être isomorphe. Un exemple est donné sur la figure 4.8. Les colonnes (a-b) présentent, successivement, une bride réelle et sa définition numérique donnée en CAO. Les colonnes (c-d) illustrent, successivement, le résultat de la détection de cercles à l'intérieur de la bride pour les deux cas, e.g. réel et théorique.

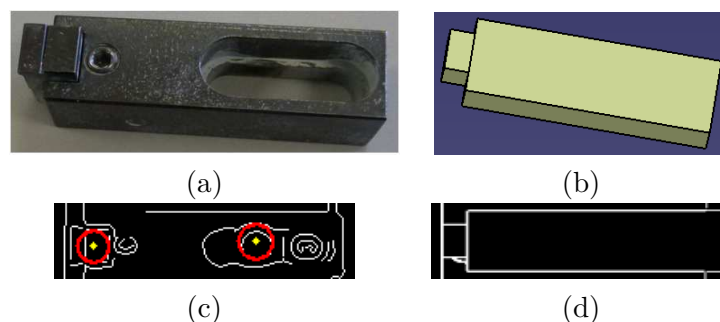


Figure 4.8 – L'impact de "l'infidélité" du modèle CAO à la réalité dans la conception du montage d'usinage, (a) image de la bride réelle, (b) modèle numérique de la bride correspondante à (a), (c) contours sur la zone de la bride sur l'image réelle du montage d'usinage, (d) contours sur la zone de la bride dans l'image théorique.

De façon générale, nous proposons que l'ensemble des appariements retenus vérifie

deux propriétés :

- Premièrement, les deux nœuds appariés doivent disposer de caractéristiques similaires, ressemblance mesurée sur la nature des nœuds. Cette information est donnée par les arêtes de partitionnement.
- Deuxièmement, l'ensemble des appariements doit respecter la relation géométrique entre les nœuds. Cette information est portée par la fonction de ressemblance (S) que nous présenterons ultérieurement.

Par conséquent, l'algorithme proposé se déroule en deux étapes :

- La première étape consiste à réaliser des coupures sur les graphes à partir de l'information sur la nature de nœuds. Ce point est développé dans la section 4.4.1.
- La mise en correspondance est effectuée alors sur les plus grands sous-graphes, indicé "gsg", en tenant compte de la relation géométrique Φ entre les deux graphes. La section 4.4.2 en présente le détail.

4.4.1 Calcul des sous graphes

Considérons un graphe $G = (v, D, p)$ calculé à partir de l'image réelle (**de même pour l'image théorique** $H = (w, E, q)$). Le vecteur $v = \{v_i, \forall i \in \mathcal{I}\}$ ($w = \{w_j, \forall j \in \mathcal{J}\}$ **pour le graphe théorique**) représente les indicateurs des nœuds, où $\mathcal{I} = 1 \dots |v|$ ($\mathcal{J} = 1 \dots |w|$ **pour le graphe théorique**). Le vecteur d'étiquettes $p = \{p = (x_i, y_i, \lambda_i), \forall i \in \mathcal{I}\}$ (**pour le graphe de l'image théorique** $q = \{q = (x_j, y_j, \lambda_j), \forall j \in \mathcal{J}\}$) représente les coordonnées spatiales pour chaque nœud et l'étiquette associée, tel que pour un nœud représentant un centre de cercle, λ_i (λ_j pour H) vaut 0, et pour un nœud représentant une intersection des droites, λ_i (λ_j pour H) vaut 1. La matrice d'adjacence D (**E pour le graphe H**) contenant l'information de proximité, l'indice de voisinage entre les paires des nœuds (a,b) est donné pour D_{ab} (**de même pour** E_{ab}) par :

$$D_{ab} = \begin{cases} 1 & \text{ssi } v_a \text{ et } v_b \text{ sont voisins} \\ 0 & \text{ailleurs} \end{cases} \quad (4.1)$$

Avant de procéder au calcul des sous-graphes, rappelons que l'algorithme proposé normalise les deux graphes. La normalisation est une étape importante pour l'appariement. En effet, la projection opérée pour obtenir l'image théorique, section 3.4.1, n'est pas dans la dimension réelle de la scène car **la pose n'est pas encore connue**. Il faut se mettre dans une situation unidimensionnelle comparable. Le but est de s'affranchir de cette différence en normalisant les deux graphes, selon l'équation (4.2), pour les ramener dans le même espace de variation $[a, b]$, e.g. pour le graphe G (de même pour le graphe H).

$$\forall (X, Y) \in \mathcal{I} \times \mathcal{I} \Rightarrow \begin{cases} N_i^x = \left(\frac{x_i - \min(X)}{\max(X) - \min(X)} \right) \times (b - a) + a \\ N_i^y = \left(\frac{y_i - \min(Y)}{\max(Y) - \min(Y)} \right) \times (b - a) + a \end{cases} \quad (4.2)$$

4.4. Appariement de graphes pour le calcul de pose : Notre proposition 79

Avec X et Y les vecteurs des coordonnées x_i et y_i respectivement. La nature du nœud doit être oubliée pour cette normalisation. Ce qui impose un premier traitement.

Ce premier traitement consiste à réaliser des coupures de graphe pour séparer le graphe des cercles et le graphe de l'intersection des droites. Les coupures sont obtenues par une analyse de la nature des arêtes, donnée par les vecteurs d'étiquetages des nœuds, (p) pour le graphe (G) et (q) pour le graphe (H), et effectuée sur les arêtes entre un cercle et une intersection des droites. Ces arêtes sont appelées **arêtes de séparation**, *Cut edge*. La figure 4.9 montre le principe du critère de partitionnement de graphe et l'algorithme associé.

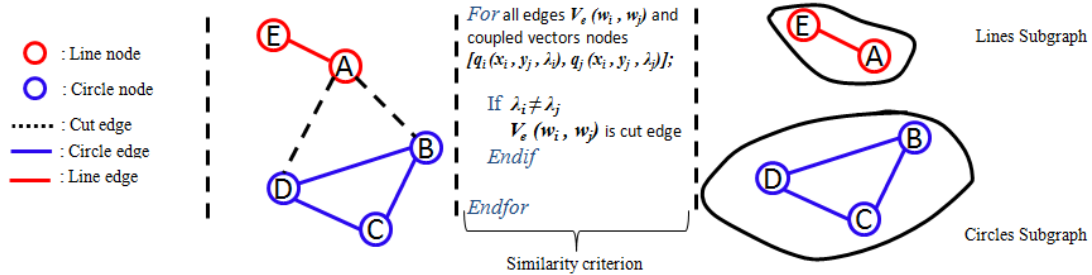


Figure 4.9 – Calcul de la coupe par analyse de la nature d'arêtes : Critère de partitionnement.

Cette coupe permet de réorganiser les deux jeux de données (primitives) et simplifier l'appariement sur les plus grand sous-graphes, tout en conservant les caractéristiques structurelle de la scène, ce qui conduit à une optimisation du temps de calcul. La figure 4.10 montre le résultat de l'application de ce critère de partitionnement sur des images de différentes configurations réalisées du montage d'usinage. Les centres des cercles sont représentés par des nœuds bleus et les arêtes bleues représentent la relation de voisinage entre deux primitives circulaires. De même pour les intersections des droites, avec la couleur rouge. Les lignes pointillées noires, illustrent les arêtes de séparations, reliant une primitive circulaire avec une primitive rectiligne, e.g. intersection des droites. Dans le cas de notre exemple "fil rouge" de la plaque Norelem[®], les primitives circulaires présentent une information riche de la topologie de la scène, sous graphe bleu, noté G_{gsg}^b vis-à-vis des intersections de droites, sous graphe rouge, noté G_{gsg}^r .

4.4.2 Détermination de la relation géométrique Φ

Les coupures obtenues par séparation des nœuds produisent des sous graphes similaires au sens du critère de la nature des nœuds. Notre objectif, maintenant, consiste à déterminer la fonction objectif/coût permettant de maximiser/minimiser le critère de ressemblance/différence entre les graphes. Ce critère de ressemblance (S) est donné par l'équation suivante :

$$S = \underset{N}{\operatorname{argmax}} \{ \min F(G|H, \hat{\Phi}) \} \quad (4.3)$$

Avec, Φ la relation géométrique permettant d'assurer un maximum de nœuds appariés, N le nombre des nœuds appariés.

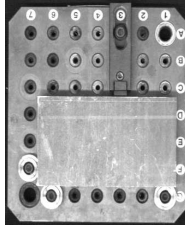
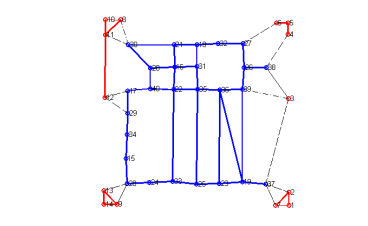
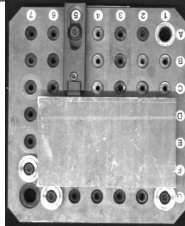
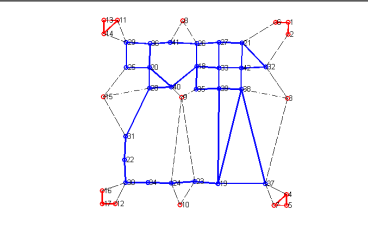
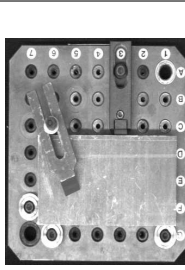
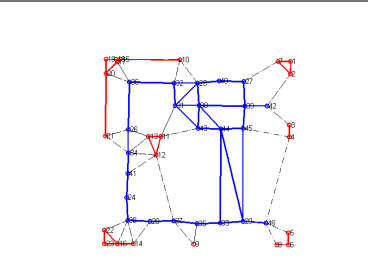
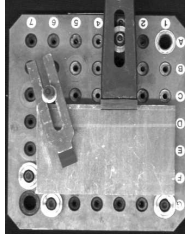
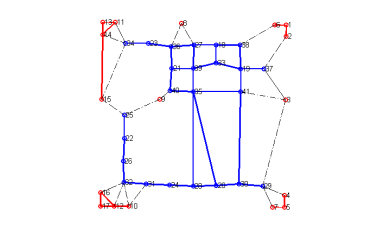
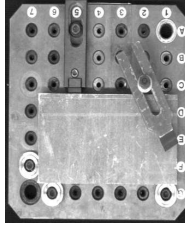
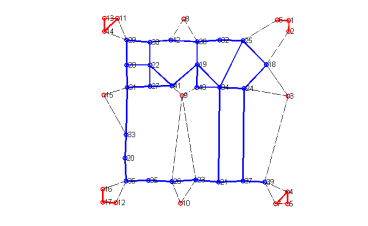
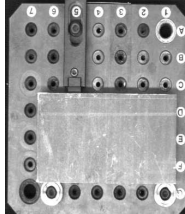
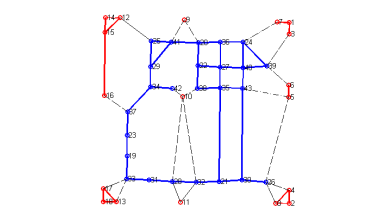
Montage conforme au modèle CAO			
Bride déplacée			
Élément de bridage ajouté			Montages non-conformes au modèle CAO
Élément de bridage ajouté et une bride qui n'a pas la bonne dimension			
Bride déplacée et un élément de bridage ajouté			
Bride déplacée et un pion de fixation en moins			

Figure 4.10 – Calcul des sous-graphes pour différentes configurations du montage d'usinage

La méthode proposée doit calculer l'appariement **quelle que soit la configuration réalisée** du montage d'usinage, même si cette dernière est totalement différente

4.4. Appariement de graphes pour le calcul de pose : Notre proposition 81

du modèle numérique du montage d'usinage attendu, car la détection de l'anomalie se fait dans une étape ultérieure, présentée dans le chapitre 5 suivant. Cette étape s'intéresse à l'analyse locale des régions définies par la pose. Ces régions présentent les emplacements attendus des éléments de bridage .

Afin de pallier ce problème, nous proposons d'effectuer un changement de repère. Ce changement de repère est réalisé sur le nœud possédant le degré de connectivité le plus élevé, donné par P_i pour le plus grand sous-graphe réel (G_{gsg}) et le nœud P^{ref} dans le plus grand sous-graphe théorique (H_{gsg}), respectivement, car le degré de connexité de chaque nœud dans le G_{gsg}^b reflète "le degré d'importance d'une région" dans le montage d'usinage. En effet, un nœud représentant un degré de connexité élevé peut indiquer une zone de forte présence de primitives circulaires et rectilignes, i.e. région de brut. Le choix du nœud caractérisé par un degré de connexité élevé dans le graphe théorique (H_{gsg}) augmente la probabilité de trouver son correspondant dans le graphe (G_{gsg}), car le critère de voisinage a été utilisé dans la construction des graphes. Si plusieurs nœuds ont le même degré de connexité le plus élevé, un tirage aléatoire est effectué sur cet ensemble pour sélectionner un nœud. La relation entre les deux nœuds de "changements de repère" sélectionnés doit respecter la condition (4.4) suivante :

$$\|P^{ref} - P_i\| \leq a + \varepsilon. \quad (4.4)$$

Avec ε la valeur de sécurité qui dépend de la distance minimale entre deux nœuds de référence (w_j^1, w_j^2) dans le graphe (H) et donnée par : $\varepsilon = \frac{w_j^1 + w_j^2}{2}$. Le couple (w_j^1, w_j^2) représente les nœuds exprimant cette distance.

Deux conclusions peuvent être tirées de cette condition :

- Si la relation (4.4) donne un résultat proche de b , le changement de repère mesuré est loin de la solution optimale, car les deux nœuds se situent à l'extrémité de l'espace de variation unidimensionnelle $[a, b]$, l'un par rapport à l'autre.
- Sinon et si la relation (4.4) donne un résultat proche de a , le changement de repère calculé est proche de la solution optimale.

La recherche de la relation géométrique optimale Φ_q^{Opt} , avec q le nombre des configurations possibles du graphe (G_{gsg}) centré sur P_i , est effectuée selon l'équation (4.5), et la maximisation de la relation (4.3) revient à minimiser cette dernière :

$$F(G|H, \hat{\Phi}) = \frac{1}{N_{ij}} \sum_{i,j \in \mathcal{I}, \mathcal{J}} Dist(v_i, w_j | S, \Phi_q) \quad (4.5)$$

F appelé la fonction de voisinage, dans laquelle pour chaque paire de nœuds $v_i(x_i, y_i), w_j(x_j, y_j)$, la distance minimale est donnée par :

$$Dist(v_i, w_j | S, \Phi_q) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (4.6)$$

Rappelons que nous avons normalisé les deux graphes, ce qui implique que la transformation Φ n'est qu'une matrice (2×2) représentant une transformation affine invariante aux translations et à l'échelle.

$$\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \quad (4.7)$$

En plus, nous nous intéressons ici à un objet rigide, qui ne peut pas subir une déformation importante, ce qui signifie que la matrice Φ est une matrice orthogonale. La difficulté pour calculer les paramètres de la matrice Φ provient de la méconnaissance des correspondances entre les nœuds. Le problème peut être résolu par un système d'équations linéaires, de type $\|P_i^{ref} - \Phi P_i\| = 0$. Avec $\|\cdot\|$ la norme euclidienne.

Certains travaux ont traité ce problème pour des applications de classification [Zhang 2007a], de reconnaissance de formes [Zhang 2006] et de calcul géométrique [Zheng 2006]. Ces solutions, ont souvent une complexité algorithmique non polynomiale *NP-complete*. Pour éviter le problème du temps de calcul dans la recherche de l'appariement, nous proposons de poser une contrainte supplémentaire. La table de fixation du montage d'usinage au sein de la machine dispose des rainures droites permettant de limiter les orientations en positionnement du montage d'usinage. La figure 4.11 en montre un exemple. Finalement, Φ n'est qu'une matrice d'orientation du plan de la recherche de l'organisation topologique des éléments de bridage, avec 4 orientations possibles $\Phi_q = [0, \pi/2, \pi, 3\pi/2]$. Dans le cas où l'organisation des éléments de bridage est symétrique par rapport aux deux axes $\vec{o}x$ et $\vec{o}y$, le problème n'est pas posé, car les rotations du plan Φ_q donnent le même appariement à chaque itération "q".

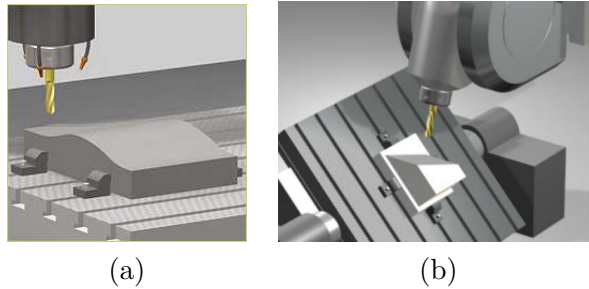


Figure 4.11 – Rainures de fixation sur la table d'usinage³.

Un exemple est donné sur la figure 4.12, avec dans la colonne (a) les deux graphes G et H dans lesquels le sous-graphe bleu représente le graphe G_{gsg} et le sous-graphe vert dans le graphe H représente le graphe H_{gsg} , respectivement. La première ligne illustre le changement de repère réalisé sur le nœud P_3 et les différentes orientations du plan Φ_q , avec $q = 1 \dots 4$ sont données sur la seconde ligne des colonnes (b), (c) et (d). Il est à noter que la configuration Φ_i^{Opt} de la colonne (d) est la configuration optimale.

Enfin, pour déterminer les correspondances, la solution proposée s'appuie sur le principe du voisin proche ou minimum local entre les nœuds v_i et w_j , employé auparavant pour la construction du graphe de proximité [Gabriel 1969]. En effet, pour chaque nœud de référence w_j , un cercle de rayon $r = \min(Dist(v_i, w_j | S, \Phi_q))$ centré sur ce dernier est calculé pour déterminer le voisinage. Un exemple de résultat pour

3. Images extraites à partir de : www.heidenhain.de

4.4. Appariement de graphes pour le calcul de pose : Notre proposition 83

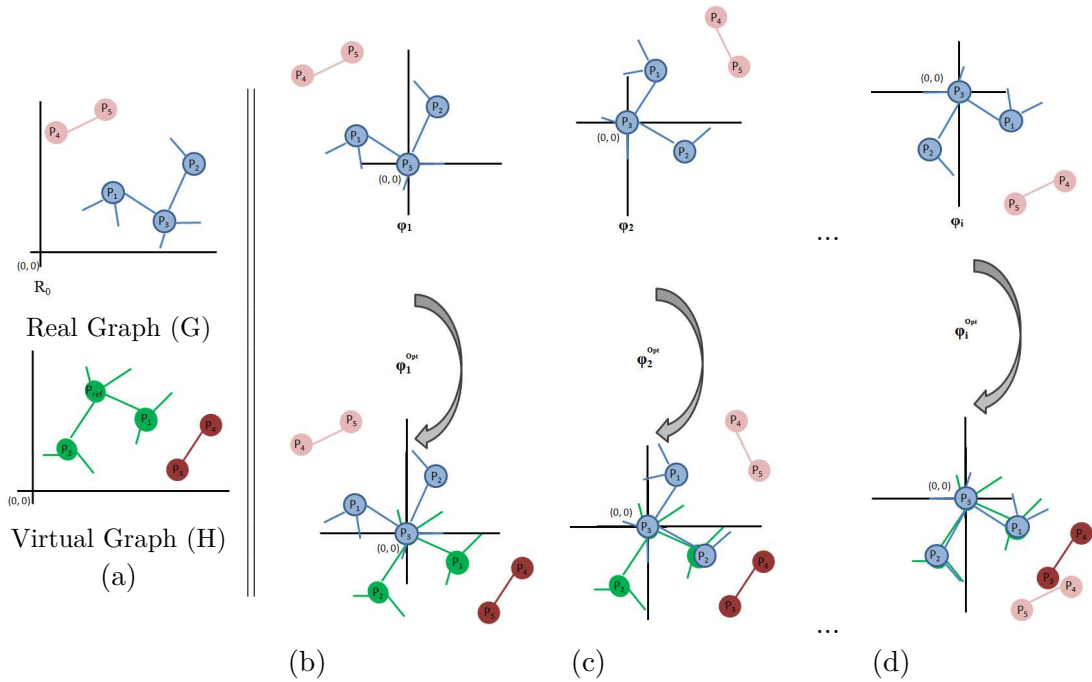


Figure 4.12 – Principe de changement du repère et la recherche de la configuration optimale ϕ : Le nœud P_3 du graphe G et le nœud P_{ref} du graphe H présentent les nœuds de changement de repère. les colonnes (b), (c) et (d) représentent les différentes configurations Φ_i , la configuration optimale est donnée par la colonne (d) ϕ_i^{Opt} .

la cas conforme du montage "fil rouge" est donné sur la figure 4.13. Les cercles jaunes illustrent les cercles de voisinage calculés pour déterminer la correspondance. Pour plus de détails la monographie de Shakhnarovich G. *et al.*[Shakhnarovich 2005] décrit les méthodes de recherche du voisin proche et leur utilisation en vision par ordinateur.

Le résultat de la mise en correspondance pour le montage d'usinage conforme est illustré sur la figure 4.14. La paire de nœuds 16 et 29 dans les graphes (G) et (H) respectivement sont les nœuds de changement de repère. Le tableau (b) indique les coordonnées des nœuds appariés et leurs étiquettes. Le diagramme à barres dans la colonne(c) représente la ressemblance (S) calculée pour chaque configuration et le nombre de nœuds appariés.

Le résultat de l'appariement pour les montages non conformes est illustré sur la figure 4.15. **Nous remarquons, que malgré le fait que les montages ne sont pas conforme au modèle CAO, l'algorithme réussit à appairer les deux graphes et trouve les correspondances entre les nœuds.**

Nous avons vu dans cette partie, la méthode proposée pour le calcul de la mise en correspondance entre les graphes G et H . Le calcul de pose de la caméra par rapport au montage d'usinage est effectué à partir des ensembles d'appariement. La seconde partie, section 4.5, présente les étapes de l'estimation des paramètres de la caméra et le calcul de la matrice de projection pour le recalage du modèle numérique CAO sur l'image du montage réalisé.

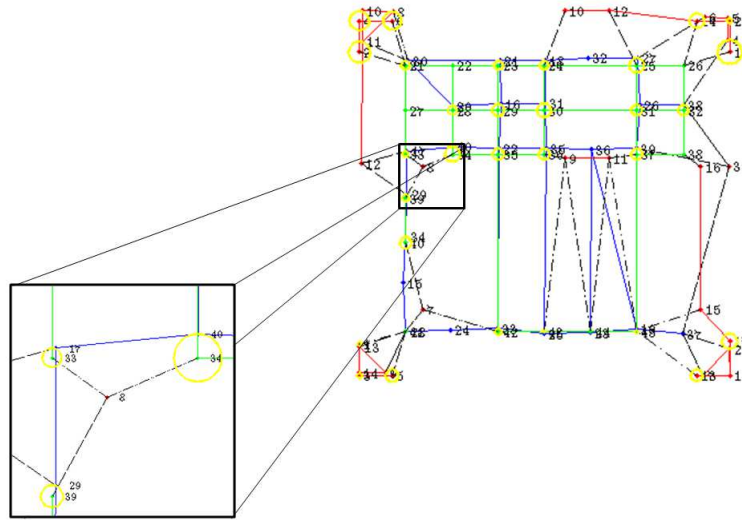


Figure 4.13 – Détermination du voisinage pour la configuration Φ_i^{Opt} de la figure 4.12. Les deux graphes G en bleu et H en vert sont projetés dans le même espace de variation dimensionnelle.

4.4. Appariement de graphes pour le calcul de pose : Notre proposition 85

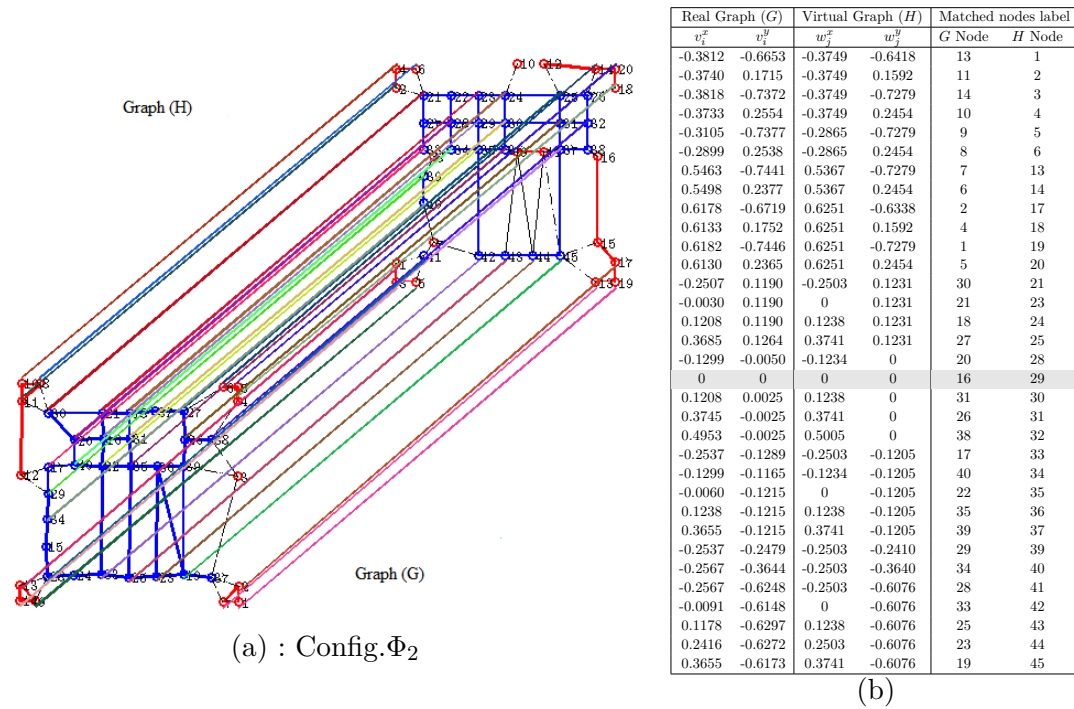
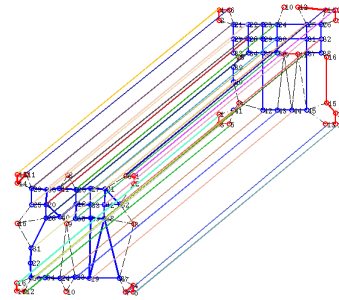
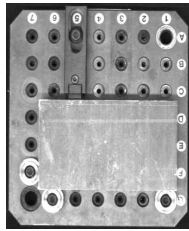
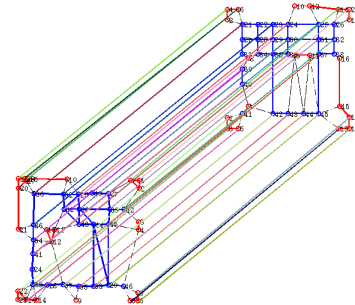
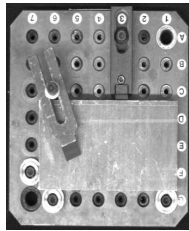


Figure 4.14 – Résultat de la mise en correspondance des graphes réel (G) et théorique (H) : (a) Configuration retenue Φ_2 , (b) : Tableau des appariements, la ligne "grisée" indique les nœuds de changement de repère. (c) : Diagramme des barres pour différentes configuration Φ_q , dans la config.1 correspond à Φ_0 pour laquelle, aucune rotation de plan n'est appliquée. La config.2 correspond à une rotation de plan $\Phi_1 = \pi/2$. Cette configuration donne de meilleurs appariements, e.g. un maximum de nœuds appariés avec une distance minimale. La config.3 correspond à une rotation de plan $\Phi_2 = \pi$. Enfin, la config.4 correspond à une rotation du plan $\Phi_2 = 3\pi/2$

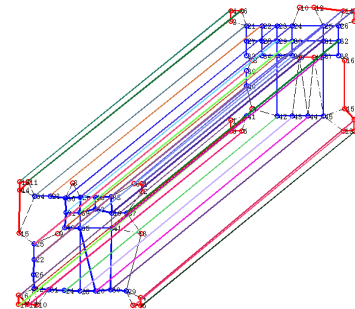
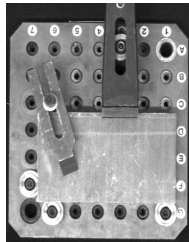
Bride déplacée



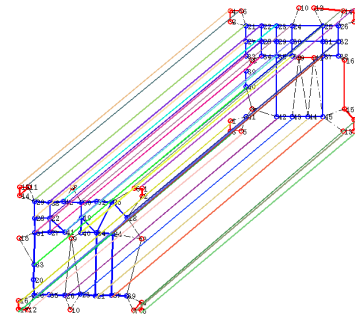
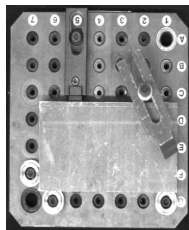
Élément de bridage ajouté



Élément de bridage ajouté et une tige qui n'a pas la bonne dimension



Bride déplacée et un élément de bridage ajouté



Bride déplacée et un pion de fixation en moins

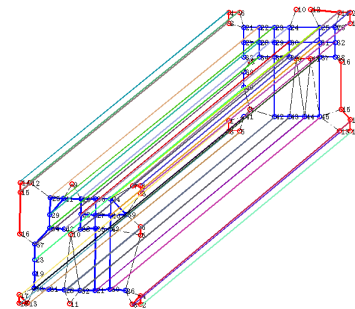
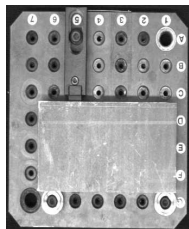


Figure 4.15 – Mise en correspondance automatique des graphes pour différentes configurations non conforme du montage d'usinage

4.5 Calcul de pose et extraction des régions d'intérêts

Rappelons que la caméra est installée au-dessus du montage d'usinage et normale par rapport au table de préparation du montage dans la machine. Les axes Z_c et Z_m des repères caméra et montage d'usinage, respectivement, sont supposés parallèles ou dans le cas idéal, confondus.

4.5.1 Modélisation de la caméra

Le modèle d'une caméra appelé sténopé (*pinhole*), défini par un ensemble de paramètres, modélise un comportement géométrique de formation des images. Dans la projection perspective, le modèle géométrique sténopé est constitué d'un plan dans lequel l'image se forme à travers le passage des rayons de lumière reflétés par l'objet par un trou très petit, appelé point focal.

Il existe deux ensembles de paramètres modélisant la caméra. Le premier ensemble, appelé **intrinsèque** et donné par la matrice $[A]$, modélise la géométrie interne du plan image et les caractéristiques optiques de la caméra. Le second ensemble, modélise la relation géométrique entre le repère de la caméra et le repère de la scène associé à l'objet de référence⁴. Ces paramètres sont appelés paramètres **extrinsèques** ou paramètres **de pose** de la caméra, ils sont donnés par la matrice $[R|t]$. La figure 4.16 en montre la configuration.

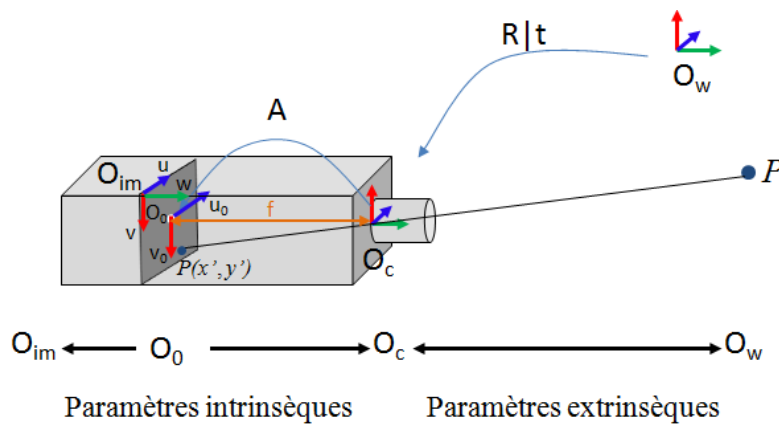


Figure 4.16 – Modèle de caméra "pinhole" et relation entre ses paramètres.

La détermination de la matrice de projection P , donnée par l'équation (4.8), nécessite la connaissance à la fois des paramètres intrinsèques de la caméra et de ses paramètres extrinsèques. Les sections 4.5.2 et 4.5.3 présentent, respectivement, les paramètres intrinsèques et les paramètres extrinsèques.

$$P = A.[R|t] \quad (4.8)$$

4. Nous supposons que le référentiel 3D du modèle CAO est confondu avec le référentiel réel du montage d'usinage.

4.5.2 Détermination des paramètres intrinsèques de la caméra

Les paramètres caractérisant une caméra sont appelés paramètres intrinsèques. Ils sont donnés sous forme matricielle par :

- La focale, donnée par le vecteur $f_c(2 \times 1)$, pour distinguer la focale⁵ en x et en y .
- Les points principaux, caractérisant la projection du centre optique sur le plan image. Ils sont donnés par le vecteur $c_c(2 \times 1)$.
- Le coefficient de l'inclinaison *skew*, défini l'angle proche de $\frac{\pi}{2}$ entre l'axe u et v de l'image. Il est donné par le facteur α_c .

La détermination de ces paramètres est obtenue par le calcul de la transformation entre le repère caméra et le repère lié au plan image $O_c \rightsquigarrow O_{im}$.

Soit un point P donné par ses coordonnées homogènes dans l'espace par rapport au référentiel caméra $P_c = (X_c, Y_c, Z_c, 1)$. La projection de ce point dans le plan image selon l'axe optique est donnée par le point $p = (x', y', z')$. Nous supposons que l'axe z est confondu avec l'axe optique, ce qui conduit à $z' = f$ selon l'équation (4.9).

$$\begin{aligned} x' &= f \frac{X_c}{Z_c} \\ y' &= f \frac{Y_c}{Z_c} \\ z' &= f \end{aligned} \tag{4.9}$$

De plus, les points images sont mesurés en pixel dans le repère (O_{im}) alors qu'ils sont donnés en unité métrique dans le repère (O_c). Afin de pouvoir écrire les relations entre le repère image et le repère caméra, nous devons introduire les paramètres suivants :

- k_u le facteur d'échelle horizontal donné en (pixel/mm).
- k_v le facteur d'échelle vertical donné en (pixel/mm).

La transformation de repère caméra au repère image peut s'écrire, maintenant, sous la forme matricielle suivante :

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} k_u & s & 0 \\ 0 & k_v & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \\ w_0 \end{bmatrix} \tag{4.10}$$

Cette relation est appelée transformation affine représentant un facteur d'échelle et une translation. s est le facteur de skew. La composante w est toujours nulle, la relation (4.10) peut être écrite dans les coordonnées homogènes :

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_u & s & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \tag{4.11}$$

5. Pour certaines caméra, les pixels ne sont pas carrés.

Remplaçons l'équation (4.9) dans (4.11) nous trouvons après simplification :

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha_u & s & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\mathbf{A}} \cdot \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (4.12)$$

Avec ,

$$\begin{aligned} \alpha_u &= f k_u \\ \alpha_v &= f k_v \end{aligned} \quad (4.13)$$

La matrice $[\mathbf{A}]$ est appelée matrice des paramètres intrinsèques : f , k_u , k_v , u_0 et v_0 qui peuvent être calculés à partir d'une étape d'étalonnage hors-ligne en utilisant, par exemple, la bibliothèque de calibration développée par J-Y Bouguet [Bouguet 2004]. Pour plus de détails sur les étapes de calibration d'une caméra, le lecteur peut se référer à ce travail.

4.5.3 Calcul des paramètres extrinsèques de la caméra

La relation qui existe entre le repère monde (O_w) et le repère caméra (O_c) permet de définir les paramètres extrinsèques. Ces derniers déterminent la relation géométrique entre le repère (O_c) et le repère (O_w) donnée par une rotation $R(3 \times 3)$ et une translation $T(3 \times 1)$, définissant **la pose de la caméra** par rapport au montage d'usinage. Rappelons que nous supposons que le référentiel du montage d'usinage est confondu avec le référentiel du modèle numérique CAO. Dans cette étape, ce sont les coordonnées des nœuds appariés qui vont servir à la détermination de la matrice des paramètres extrinsèques, transformation entre le repère (O_w) et (O_c).

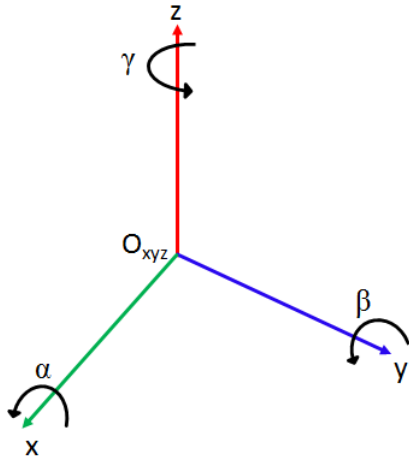
La position du point focal par rapport au repère O_w permet de déterminer le vecteur translation (T). La matrice de rotation $R(3 \times 3)$ permettant de définir la rotation du repère (O_c) par rapport au repère (O_w), elle peut être calculée par le produit matriciel des trois matrices de rotation simple, comme le montre la figure 4.17.

Avec, $R = Rot(Z, \gamma).Rot(Y, \beta).Rot(X, \alpha)$.

$$R = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \quad (4.17)$$

De manière à simplifier la notation de la matrice R , il est plus intéressant de l'écrire en fonction de ses vecteurs d'orientation r_i , dont chacun possède trois composantes. Cette relation est donnée par l'équation (4.18).

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad (4.18)$$



$$Rot(X, \alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \quad (4.14)$$

$$Rot(Y, \beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \quad (4.15)$$

$$Rot(Z, \gamma) = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.16)$$

Figure 4.17 – Définition des paramètres de la matrice de rotation.

De ce fait, la relation (4.18) est habituellement donnée par une seule matrice avec des coordonnées homogènes, comme le montre l'équation (4.19)

$$\begin{bmatrix} P_c \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ 0_{1 \times 3} & 1 \end{bmatrix} \cdot \begin{bmatrix} P_w \\ 1 \end{bmatrix} \quad (4.19)$$

La détermination de ces paramètres à partir de l'ensemble des appariements, permet de trouver la relation géométrique entre la caméra et le montage d'usinage. Le même algorithme [Bouguet 2004] a été utilisé pour le calcul des paramètres extrinsèques. La figure 4.18 illustre un résultat de calcul de la matrice des paramètres extrinsèques $[R|t]$ pour le cas conforme du montage de notre exemple "file rouge".

L'hypothèse de placer la caméra au-dessus du montage d'usinage peut être exprimée physiquement par une matrice de rotation, i.e. la matrice de la figure 4.18.d, dont les valeurs absolues de la diagonale doivent être égales à 1 pour le cas "idéal", la matrice de rotation égale à la matrice d'identité I . C'est le cas des "rotations pures" en robotique, par exemple.

La figure 4.18.a présente la relation entre la caméra et le montage d'usinage exprimée par rapport au repère de ce dernier. Les points bleus représentent les coordonnées des points caractéristiques utilisés dans la mise en correspondance. La figure 4.18.b présente cette même relation dans le repère de la caméra.

A partir de la matrice P , nous projetons le modèle CAO sur l'image réelle, constituant ainsi une forme de **réalité virtuelle**. La figure 4.19.b montre l'exemple pour le cas conforme du montage de la plaque Norelem[©]. Cette projection va permettre de localiser les régions *prévues* des éléments de la scène selon leur définition de référence donnée en CAO.

En effet, l'organisation topologique des éléments constituant le montage d'usinage dans sa définition numérique CAO est exprimée par un graphe des attributs, ARG .

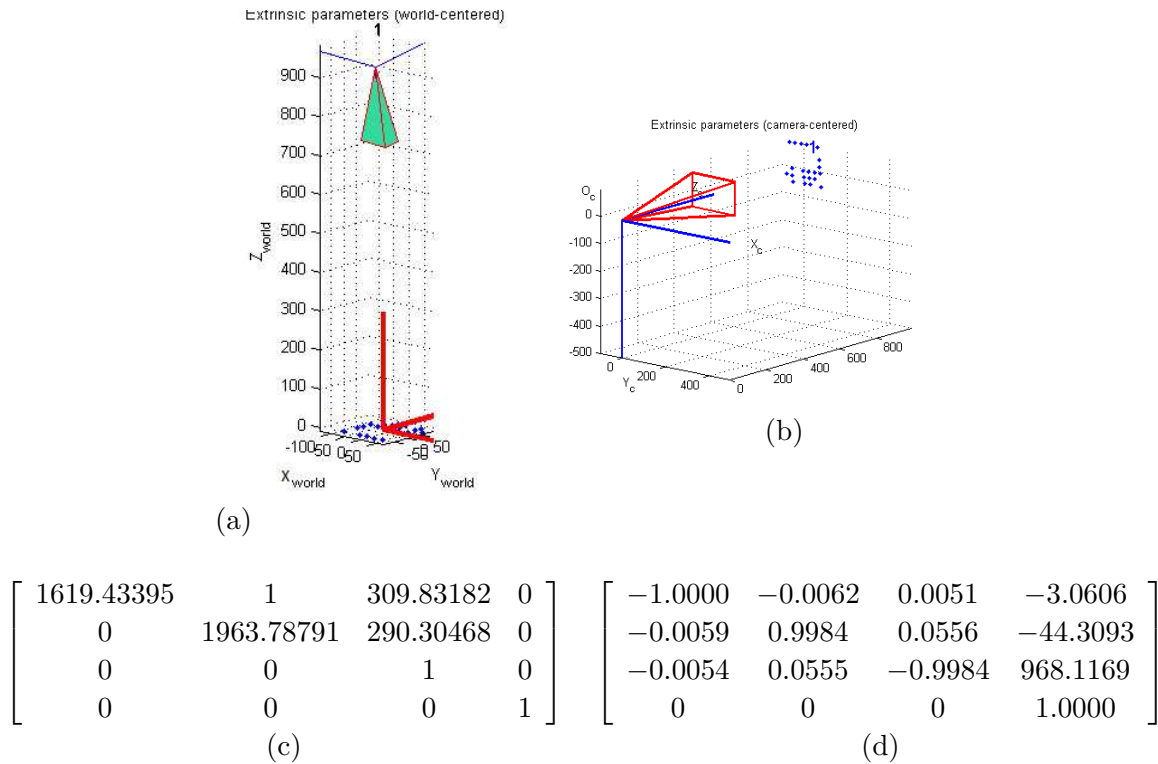


Figure 4.18 – Estimation des paramètres intrinsèques et de la pose de la caméra par rapport au montage d'usinage, dans le cas conforme : (a) La pose de la caméra exprimée par rapport au repère du montage d'usinage, donné par le modèle CAO, (b) la pose de la caméra exprimée par rapport au repère de la caméra, (c) la matrice des paramètres intrinsèques déterminés lors du calibrage hors-ligne et (d) la matrice des paramètres extrinsèques ou de la pose de la caméra par rapport au montage d'usinage.

Celui-ci est calculé à partir du modèle CAO, voir la figure 4.19.a, et projeté sur l'image "réelle" après calcul des paramètres de la matrice de projection, voir la figure 4.19.b. Les arêtes entre les nœuds représentent le voisinage mesuré sur le modèle CAO entre les objets constituant le montage d'usinage. Chaque nœud, dans ce graphe des attributs, voir la figure 4.19.c, est caractérisé par un vecteur caractéristique contenant l'information de dimension, de taille et de la forme⁶ de l'objet en question, ce qui permet de délimiter les zones d'intérêts. La figure 4.19.d en présente un exemple. Une fois la pose estimée, nous recalculons ce voisinage après projection du graphe sur l'image afin de vérifier la précision de la pose.

La figure 4.20 montre le résultat du calcul de pose et la segmentation d'image en régions d'intérêts pour différentes réalisations non-conformes du montage d'usinage. **Il est remarquable que la méthode proposée a réussi à calculer et à recalculer le modèle numérique de référence sur l'image du montage réalisé et à déterminer correctement les emplacements prévus des éléments de bridage, malgré que les montages d'usinage réalisés ne soient pas conformes**

6. Nous utilisons les contours pour décrire la forme de l'objet

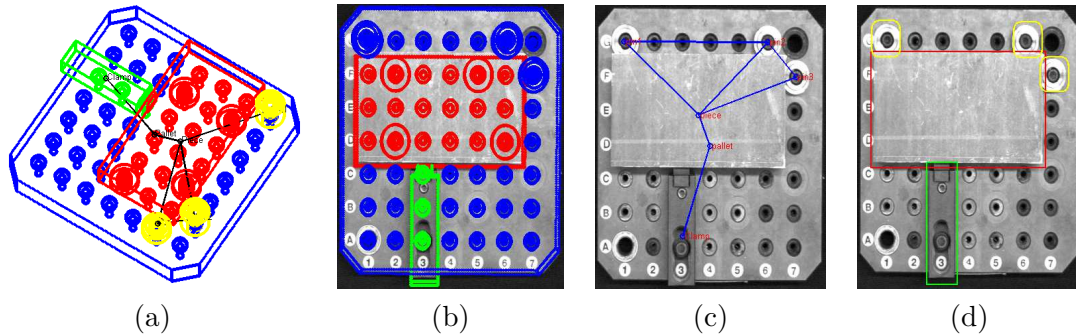


Figure 4.19 – Localisation des éléments de bridage : (a) graphe des attributs (ARG) du modèle CAO, (b) projection du modèle CAO sur l'image réelle du montage d'usinage, (c) Calcul du graphe des attributs après estimation de la pose. (d) localisation des éléments de la scène à partir du ARG projeté.

au modèle CAO.

La figure 4.21 illustre le résultat pour différentes réalisations du montage "Semelle", et de même sur la figure 4.22 pour le montage Minimanche.

Dans la section suivante, nous nous attachons à évaluer la méthode proposée, à la fois par rapport aux paramètres du modèle, e.g. graphe, et les paramètres photométriques de la scène, i.e. éclairage, faible contraste, etc. de façon à ce qu'elle puisse être exploitée dans la majorité des cas. Nous montrerons aussi ses limites.

4.6 Évaluation de la méthode

Les sections 4.4 et 4.5 ont montrées l'intérêt de l'utilisation des graphes dans la mise en correspondance pour déterminer la pose et recaler le modèle numérique sur l'image réelle du montage d'usinage. Ce qui permet aussi de segmenter l'image en régions d'intérêt afin d'analyser chaque région pour la prise de décision sur la conformité du montage réalisé, e.g. chapitre 5. Dans le but de vérifier la robustesse de la méthode, nous avons réalisé deux études principales :

- **Test d'évaluation de la robustesse vis-à-vis du bruit** : il consiste ici à tester l'évaluation de la robustesse de la méthode par rapport au bruit dans les images, pour déterminer sa limite, car le logiciel développé doit répondre aux besoins du cahier des charges sous les conditions photométriques de l'atelier, usine de production⁷. Ce test est développé dans la section 4.6.1.
- **Test d'évaluation de la robustesse par rapport à la topologie des graphes** : ce test consiste à étudier la relation entre les graphes et la méthode de la mise en correspondance. Il s'agit d'analyser l'influence de la structure du graphe, le nombre des nœuds, leur distribution, etc. dans la mise en correspondance. Des travaux existants ont déjà étudié la relation entre la géométrie des points et la précision de la pose [Wrobel 2001]. Certains d'entre eux reposent

7. Le problème de reflet spéculaire sur des pièces réfléchissantes n'est pas considéré dans ce travail.

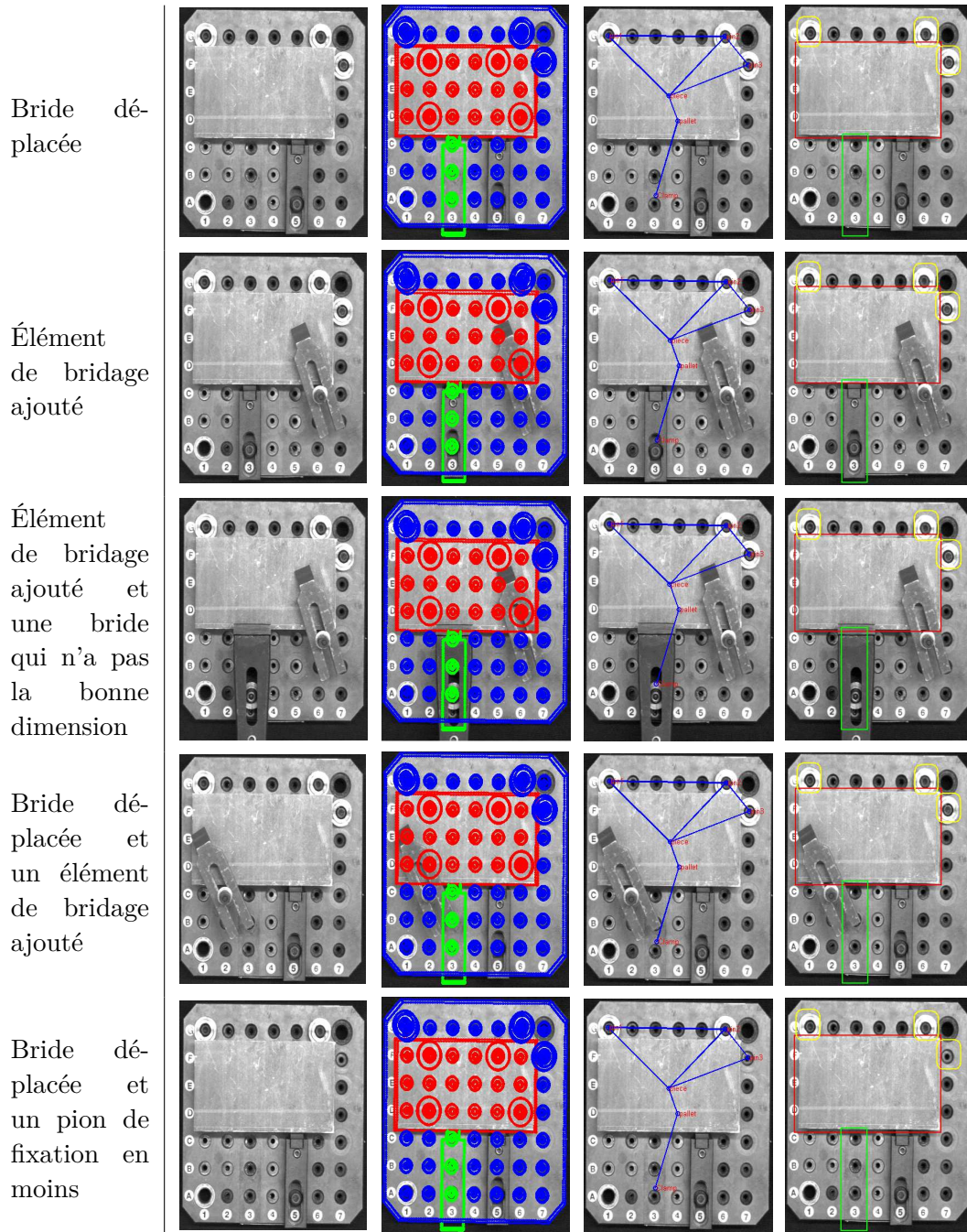


Figure 4.20 – Localisation automatique des régions d'intérêts pour différentes configurations conformes et non conformes du montage Norelem[©].

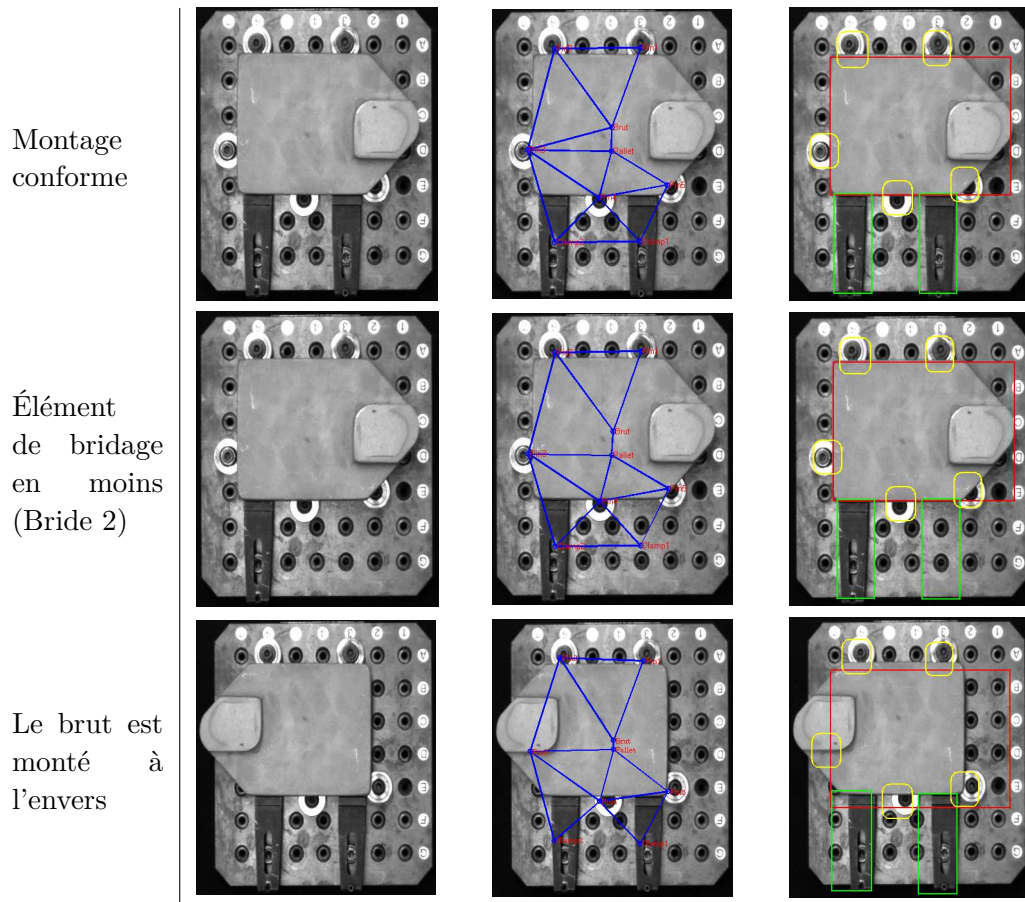


Figure 4.21 – Localisation automatique des régions d'intérêts pour différentes configurations conformes et non conformes du montage semelle.

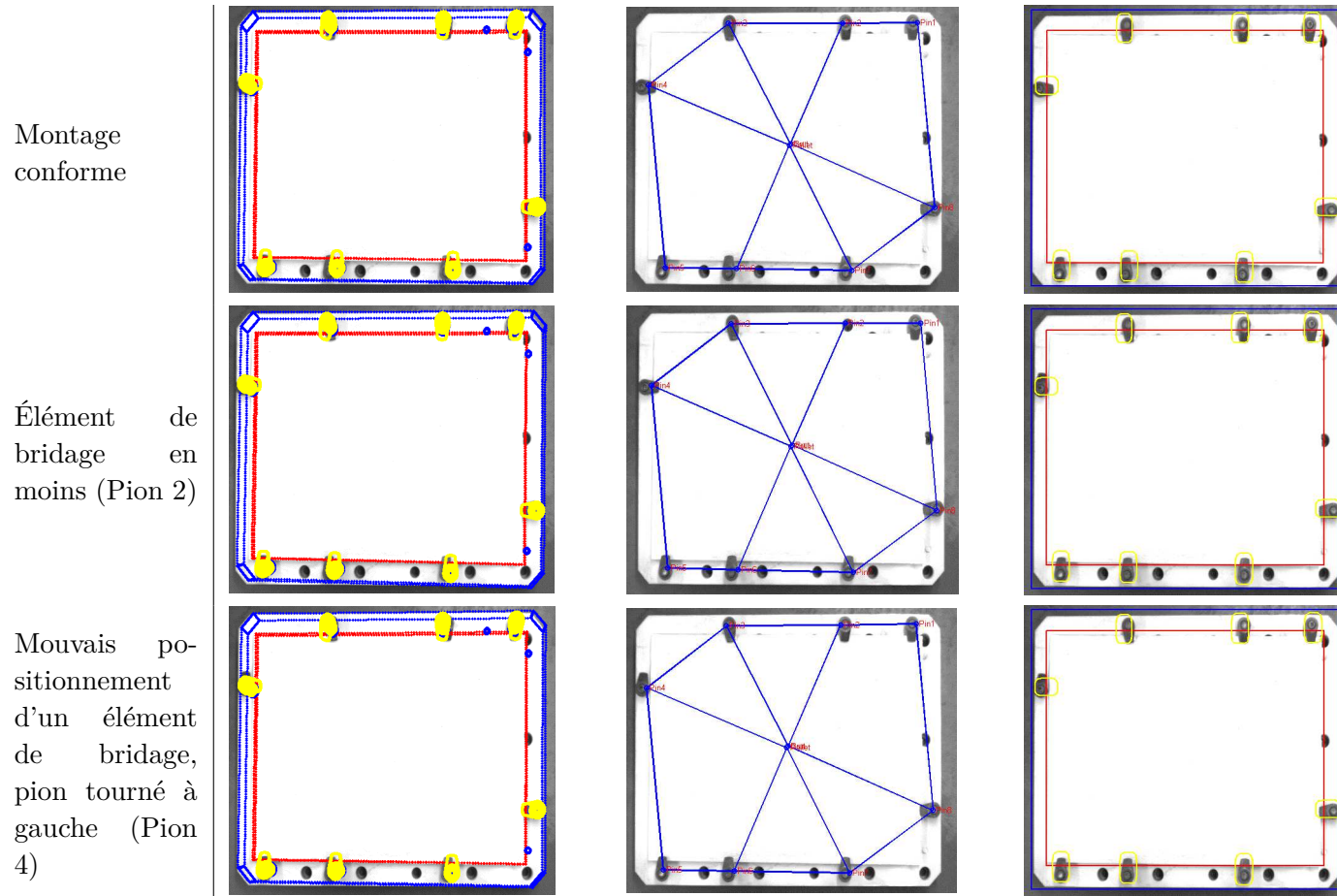


Figure 4.22 – Localisation automatique des régions d'intérêts pour différentes configurations conformes et non conformes du montage Mini-manche.

sur l'analyse des solutions selon le nombre de points utilisés dans le calcul de la pose, i.e. pour une configuration à 3 points [Haralick 1991], ou une configuration à 4 points [Nister 2004], et plus récemment pour une configuration à n points, algorithme PnP [Moreno-Noguer 2007]. L'objectif de notre étude repose sur le calcul de l'erreur moyenne et de l'écart-type de la pose par rapport à chaque configuration réalisée aléatoirement. La section 4.6.2 en présente le détail.

La figure 4.23 montre un schéma simplifié de la relation entre les différentes sections du chapitre et les trois essais menés pour l'étude de la robustesse de la pose.

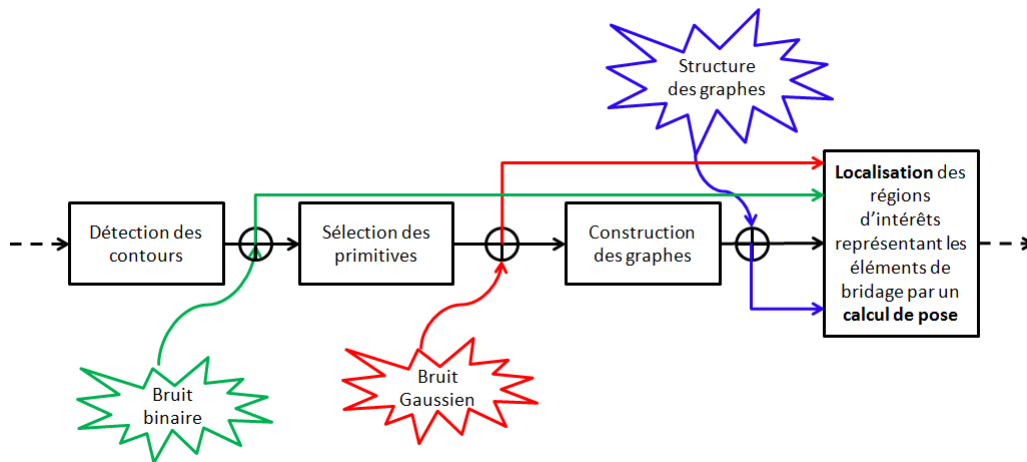


Figure 4.23 – Schéma simplifié de l'étude de la robustesse des paramètres de chaque section du chapitre par rapport à la précision de la pose. La ligne "verte" présente le test de l'évaluation du bruit impulsif, la ligne "rouge" présente le test de l'évaluation du bruit gaussien et la ligne "bleu" illustre le test de l'évaluation de la topologie des graphes.

4.6.1 Évaluation de la méthode par rapport au bruit

Pour étudier l'impact du bruit dans l'image vis-à-vis de la précision de la pose, nous proposons d'utiliser l'image théorique des contours comme image de référence. Nous affectons à chaque itération un bruit binaire d'une densité différente⁸. En effet, le bruit peut affecter la détection des primitives soit par un décalage des points caractéristiques, i.e. centre des cercles et intersection des droites, soit par l'incapacité des algorithmes de détection des primitives à détecter ces dernières, due à la quantité du bruit sur les contours. La figure 4.24 montre le résultat obtenu pour différents taux de bruit. Nous procédons ainsi plutôt que de bruiteur l'image initiale afin de perturber plus directement les contours, car le traitement de sélection des primitives exploite les contours. En effet, le lien entre les perturbations et les conditions photométriques de prise de vue, nous conduisent à une multitude de combinaisons de variation d'éclairément pour obtenir les mêmes degrés de perturbation.

8. Initialisé à 5% au départ

A chaque itération, nous augmentons le taux du bruit, et nous calculons l'erreur moyenne et l'écart-type de la pose en x et y sur l'ensemble des appariements. Nous remarquons qu'à partir d'un taux supérieur à 20%, les algorithmes de détection de primitives perdent leurs robustesse, et la construction des graphes n'est plus possible. Par conséquent, l'appariement n'est plus possible non plus. Pour un taux inférieur, compris entre 5 et 10%, le calcul de la pose n'est pas acceptable car l'erreur moyenne et l'écart-type sont importants, et pour un taux inférieur à 5%, la méthode demeure relativement précise, e.g. erreur moyenne et écart-type inférieur à 3 pixels.

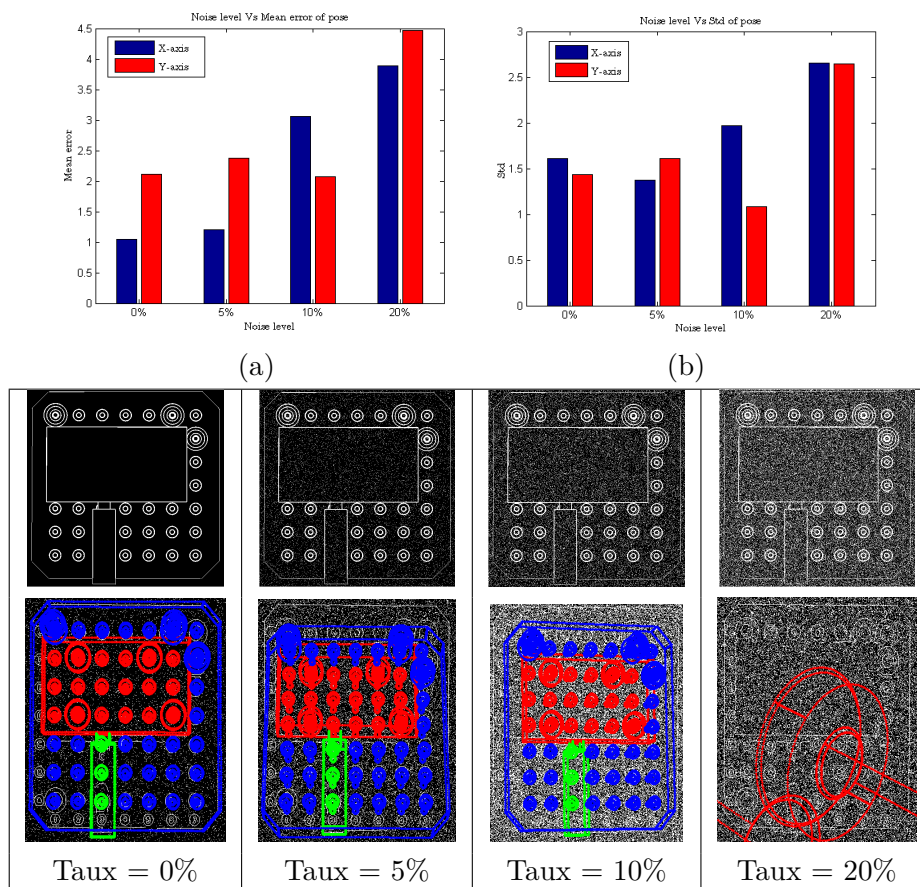


Figure 4.24 – Évaluation de la pose par rapport au bruit : (a) Erreur moyenne en pixels par rapport à l'axe x et y , (a) Écart-type moyen en pixels par rapport à l'axe x et y , (c) Images théoriques utilisées pour le test.

De plus, la précision du calcul des centres des cercles et des intersections des droites donnant les points de mise en correspondance peut affecter aussi la précision de la pose. Nous avons réalisé un autre type de test s'appuyant sur la perturbation à chaque itération i de l'ensemble des points mis en correspondance en utilisant un bruit Gaussien de moyenne nulle et de variance $\sigma_i = i$. Le résultat est donné sur la figure 4.25. Nous remarquons que la variation de l'erreur moyenne et l'écart-type de la pose mesurés sur les points appariés croît exponentiellement à partir d'une valeur de bruit $O(i, \sigma_i)$ avec $i \geq 10$

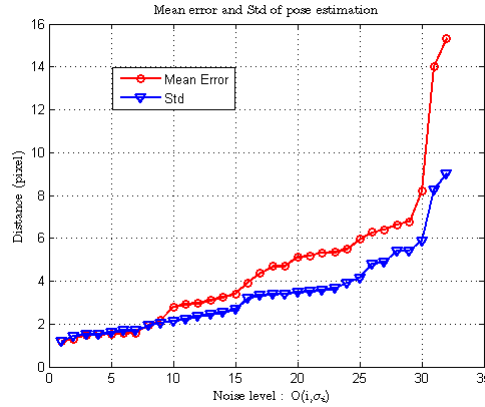


Figure 4.25 – Évaluation de la pose par rapport aux coordonnées des points d'intérêts.

4.6.2 Évaluation de la méthode par rapport à la topologie des graphes

Dans un premier temps, nous nous intéressons au test de l'évaluation de la pose par rapport au nombre de points utilisés dans la mise en correspondance. Le résultat pour le cas conforme est donné sur la figure 4.26.

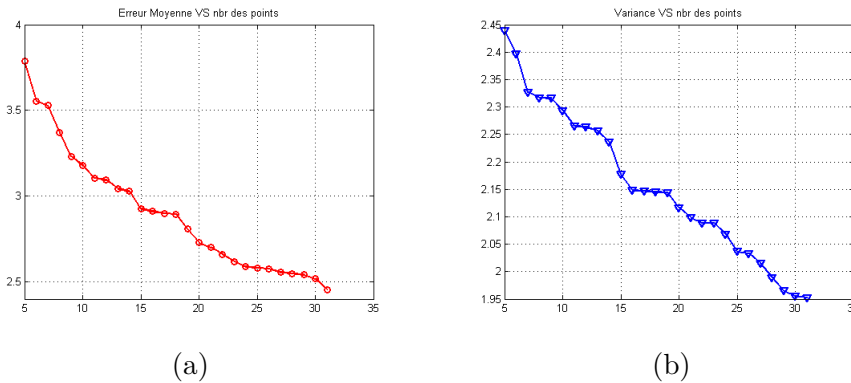


Figure 4.26 – Évaluation de la pose par rapport au nombre des nœuds appariés pour le cas conforme.

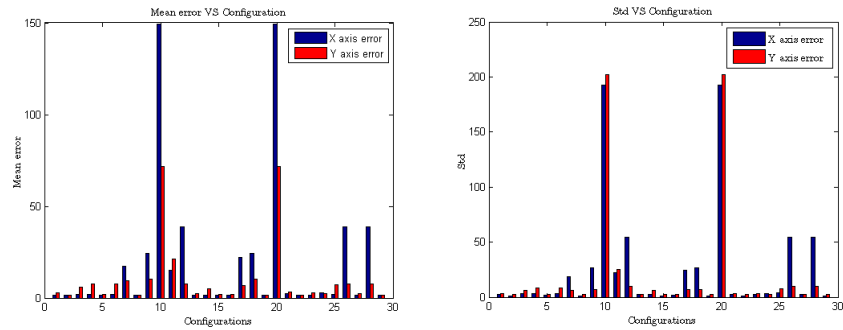
Nous remarquons qu'à partir de 5 points la précision de la pose augmente. L'erreur moyenne et l'écart-type diminuent avec l'augmentation du nombre de points.

Dans l'objectif d'évaluer la robustesse de la méthode pour différentes configurations du graphe, un autre type de test a été réalisé. Dans le travail de Haralick *et al.* [Haralick 1991], les auteurs ont étudié la stabilité des méthodes utilisant 3 points et le problème de redondance. Pour une solution linéaire à partir de 3 points, le problème de calcul de pose possède en général 4 solutions possibles. Pour pallier ce problème, il faut ajouter des données. Par exemple, si l'on ajoute un quatrième point de référence, il existe, en général, une solution unique. La convergence de la solution dépend de la distribution des points dans l'espace par rapport au plan image. Dans certains cas, même un nombre infini de points ne donne pas une solution unique, c'est le cas où tous les points sont à l'infini, la translation de la caméra ne peut pas être estimée.

Si 4 points coplanaires sont distribués pour former une droite dans le plan image, i.e. situations Config.10 et Config.20 de la figure 4.27, la rotation ne peut pas être calculée. Nous constatons que la pose est "incorrecte", l'erreur moyenne et l'écart-type sont supérieurs à 150 et 180 pixels respectivement, voir les figures 4.27.a et 4.27.b qui correspondent à l'erreur moyenne et l'écart-type de chaque configuration de la figure 4.27.c.

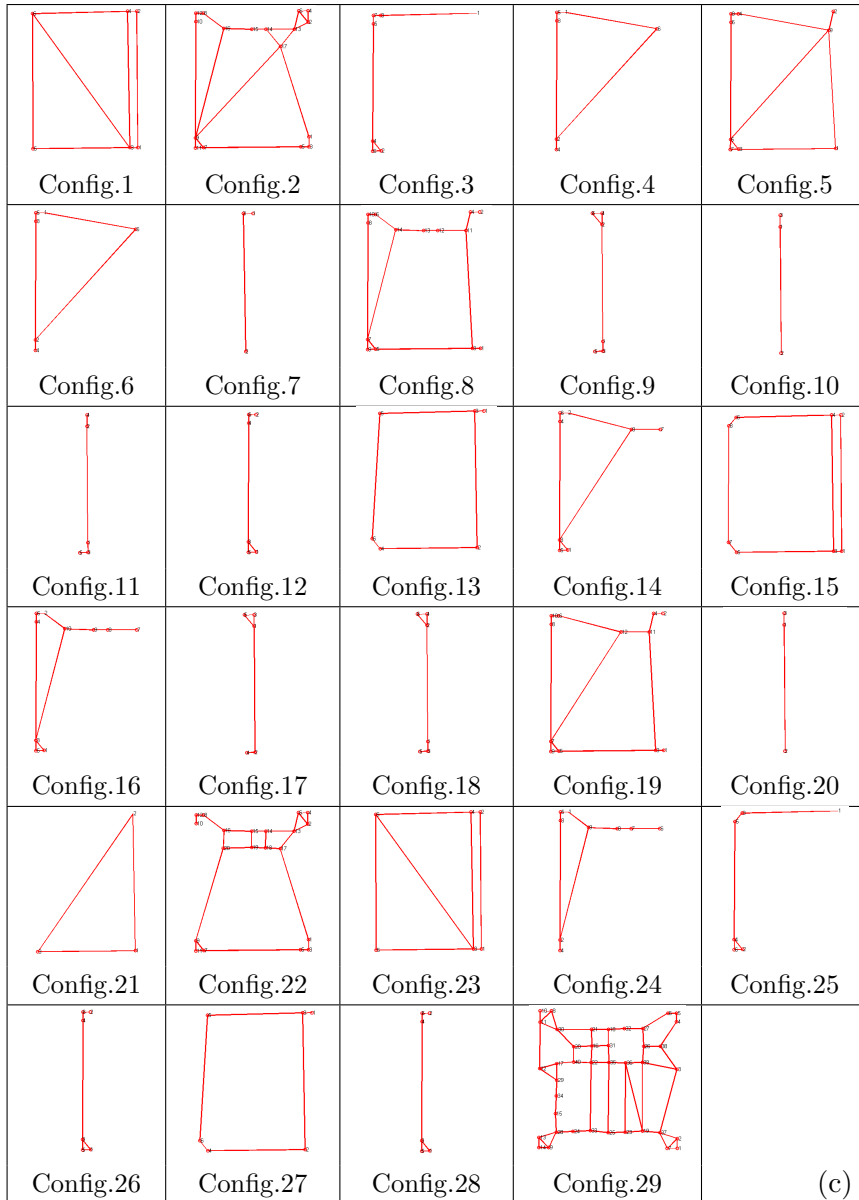
Pour résumer, la répartition des nœuds et leur nombre dans le graphe jouent un rôle important dans la précision de la pose.

Cependant, cette propriété ne peut pas affecter la robustesse de notre méthode, car les graphes sont construits à partir des primitives pertinentes, e.g. circulaires et rectilignes, qui caractérisent les montages d'usinage. Ces derniers sont constitués par des éléments polyédriques, cylindriques et sphériques. La présence de primitives circulaires et rectilignes dans les montages d'usinage est suffisamment riche.



(a)

(b)



(c)

Figure 4.27 – Évaluation de la pose par rapport à la structuration du graphe, i.e. distribution spatiale des nœuds : (a) Erreur moyenne en pixels par rapport à l'axe x et y , (b) Écart-type moyen en pixels par rapport à l'axe x et y , (c) Configurations des graphes correspondants. La configuration 29 représente la configuration dont la valeur moyenne de l'erreur de pose est minimale, avec une moyenne en $Mean(x,y) = (1.2063, 2.3697)$ et un écart-type $std(x,y) = (1.3732, 1.6096)$

4.7 Vérification de la pose et cas limites

La matrice des paramètres extrinsèques est composée d'une matrice de rotation $R(3 \times 3)$ et d'un vecteur de translation $T(3 \times 1)$. Nous allons étudier dans cette section l'impact de chaque composante dans la localisation des régions d'intérêts.

4.7.1 L'impact de la matrice de rotation

Un des moyens pour détecter la précision de la pose est de mesurer l'isomorphisme du graphe ARG calculé sur le modèle CAO et projeté sur l'image réelle. Si l'isomorphisme entre le graphe ARG théorique (ARG_{th}) n'est pas respecté par rapport au graphe ARG réel (ARG_r) projeté, cela signifie que la pose est probablement incorrecte. Ce cas de figure peut être présent dans le cas d'un changement, "non respect" de l'hypothèse de configuration de la caméra par rapport à l'objet, et cela surtout dans le cas d'une inclinaison du plan (xz) ou (yz) ou des deux. Cette anomalie peut être détectée par la matrice d'adjacence de voisinage. En effet l'inclinaison dans le plan (xz) ou (yz) peut déclencher une nouvelle définition d'un voisinage dans (ARG_r) qui n'est pas défini dans le (ARG_{th}). Un exemple est donné sur la figure 4.28, dans laquelle l'arête liant le nœud représentant le "pions3" avec le nœud représentant la "plaque" n'est pas définie dans la représentation CAO. Sa présence dans le graphe réel déclenche une anomalie, e.g. matrices (e) et (f) de la figure 4.28 .

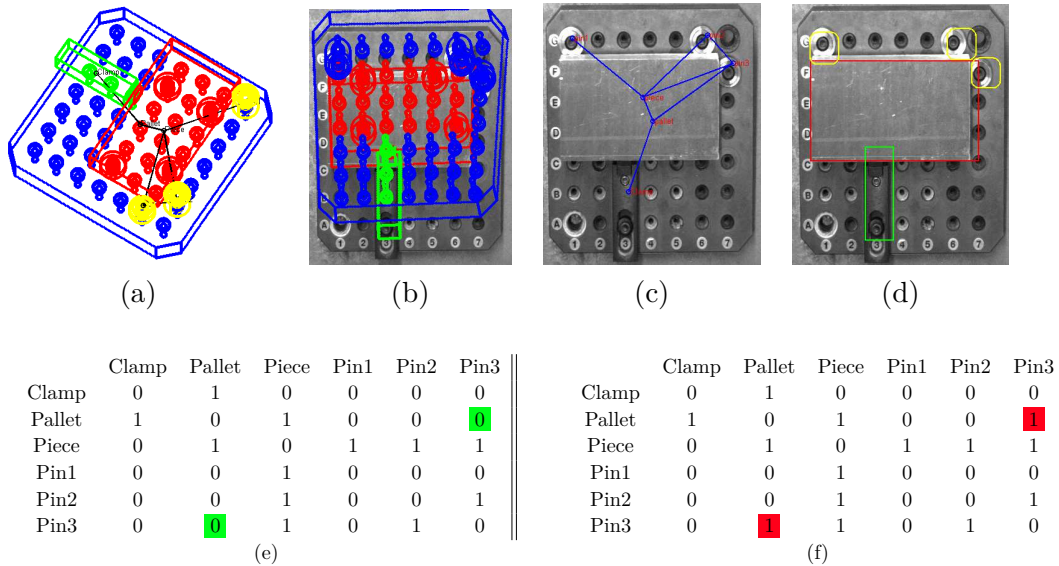


Figure 4.28 – Détection de l'anomalie dans le calcul de pose à partir du graphe ARG : (a) Graphe ARG_{th} selon sa définition en CAO, (b) : projection du modèle numérique CAO sur l'image du montage réalisé, (c) : Graphe ARG_r calculé après estimation de la pose, (d) : Segmentation d'image et extraction des régions d'intérêts, (e) Matrice adjacence du graphe ARG_{th} selon la définition CAO, (f) Matrice adjacence du graphe ARG_r . Les cases rouges dans la matrice (f) représentent l'anomalie, décrit par la définition d'un nouveau voisinage qui n'est pas défini en ARG_{th} .

Une façon de pallier ce problème est de recalculer les paramètres de la matrice de rotation en utilisant les droites données par l'équation (4.20) :

$$aX + bY + cZ = \vec{n} \cdot \vec{F}P = 0 \quad (4.20)$$

Avec X, Y, Z les coordonnées du point P dans le repère du montage d'usinage, dont sa projection est donnée par p dans la plan image. Par définition, une droite est caractérisée par sa direction \vec{v} et un vecteur point \vec{p} . L'équation de cette droite portant le point P dans le repère caméra peut être écrit :

$$\begin{aligned} \vec{P} &= R\vec{p} + \vec{t} \\ \vec{V} &= R\vec{v} \end{aligned} \quad (4.21)$$

Avec R et \vec{t} représentent la matrice de rotation (3×3) et le vecteur de translation (3×1) respectivement. Deux méthodes, linéaire [Horaud 1993] et itérative [Liu 1990] sont envisageables pour déterminer la nouvelle matrice de rotation R . Nous présentons ici **quelques éléments de solution** de chaque méthode et nous montrons les limites par rapport à notre application.

Dans le premier cas, les contraintes de correspondance expriment le fait qu'une droite appartient au plan défini par l'équation 4.20 respecte les relations suivantes :

$$\begin{cases} \vec{n} \cdot \vec{V} = 0 \\ \vec{n} \cdot \vec{P} = 0 \end{cases} \quad (4.22)$$

Ce qui conduit à l'expression,

$$\begin{cases} \vec{n} \cdot (R\vec{v}) = 0 \\ \vec{n} \cdot (R\vec{p} + \vec{t}) = 0 \end{cases} \quad (4.23)$$

A partir de l'équation (4.23), chaque droite fournit deux contraintes. Avec 6 paramètres à déterminer, le problème de la pose peut être résolu à partir de 3 équations de droites. En général, le problème revient à minimiser une fonction coût donnée par l'équation (4.24) pour N droites [Horaud 1993].

$$f(R, t) = \sum_{i=1}^N (\vec{n}_i \cdot (R \cdot \vec{v}_i))^2 + \sum_{i=1}^N (\vec{n}_i \cdot (R \cdot \vec{p}_i + \vec{t}))^2 \quad (4.24)$$

Seule une inclinaison sur la plan (xz) et (yz) peut poser un problème dans le calcul de la pose, car, dans ce cas, l'hypothèse de la création de l'image théorique n'est pas respectée. La rotation sur le plan (xy) ne pose pas de problème car, cette information est portée par les formes rectilignes, e.g. sélection des droites, dans le montage et par conséquent les points d'intersection des droites.

La seconde solution consiste à corriger la pose et recalculer la matrice de rotation $[R]$. Elle vise à exprimer la matrice $[R]$ en fonction de ses 3 paramètres angles α, β et γ . Le calcul de ces angles de rotation est réalisé en ajustant des petites corrections

$\Delta\alpha$, $\Delta\beta$ et $\Delta\gamma$ aux valeurs initiales des angles α , β et γ . Les matrices de rotation sur l'axe x et y sont données par :

$$Rot(X, \alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \quad Rot(Y, \beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \quad (4.25) \quad (4.26)$$

Liu *et al.* [Liu 1990] utilisent une expression linéaire $L(\Delta\alpha, \Delta\beta, \Delta\gamma)$ aux voisinage des valeurs initiales des angles α , β et γ pour retrouver les valeurs de la matrice de rotation.

Pendant, la complexité pour notre application, se trouve au niveau du modèle CAO. En effet, le modèle numérique CAO du montage d'usinage est un assemblage compliqué de formes simples : cubes, sphères, etc. Elles sont souvent exprimées soit par des surfaces⁹, format *.iges ou des triangles, format *.stl, etc. Contrairement à d'autres travaux proposant une solution d'estimation simultanée de la pose des objets relativement "simples"¹⁰ [Dementhon 1995], [Sandhu 2011], pour nous, la diversité des formes constituant le montage d'usinage complique la tâche de définition de paramètre ou d'un vecteur des paramètres, sur lequel le modèle peut s'appuyer pour converger itérativement à la solution optimale, dans le cas de détection de l'anomalie.

4.7.2 L'impact du vecteur de translation

La précision de la pose dépend aussi du vecteur de translation dans la matrice des paramètres extrinsèques. L'inclinaison sur le plan (xz) ou (yz) pourrait affecter la matrice de rotation et par conséquent le vecteur de translation, car les deux composantes sont dépendantes et constituent la matrice de solution d'un système d'équations linéaires.

Pour réaliser un essai de vérification de la translation, nous avons fixé la matrice de rotation, en changeant la composante Z du vecteur translation $T(X, Y, Z)$. Un exemple est donné sur la figure 4.29 pour (a) $Z = 700mm$ et (b) $Z = 1200mm$. Les deux acquisitions d'images ont les mêmes paramètres intrinsèques et de rotation, seule le vecteur translation est différent, e.g. différents points de vue de la caméra.

Nous remarquons dans la figure 4.29.c que même si la pose est correcte, l'utilisation d'un élément qui n'a pas la bonne dimension pourrait affecter la détermination de la région d'intérêt. La bride est plus large que la bride attendue. En s'appuyant sur la taille de la bride de référence donnée par la modèle CAO, la région de la bride déterminée après le calcul de la pose ne reflète pas la taille réelle de la bride utilisée. Ce cas de figure est analysé dans le chapitre suivant de l'analyse locale pour la prise de décision.

En effet, si la pose n'est pas correcte, il est difficile de qualifier le(s) paramètre(s) responsable(s) de cette erreur. D'une part, les paramètres de la matrice de pose ne sont que les solutions d'un système d'équations linéaires. Le changement d'un paramètre

9. Le fichier *.IGS du montage de la plaque Norelem est composé de plus de 600 surfaces. Quant au montage Semelle, il contient plus de 850 surfaces.

10. Objet constitué par moins de dizaines de facettes, i.e. cube 6 facettes.

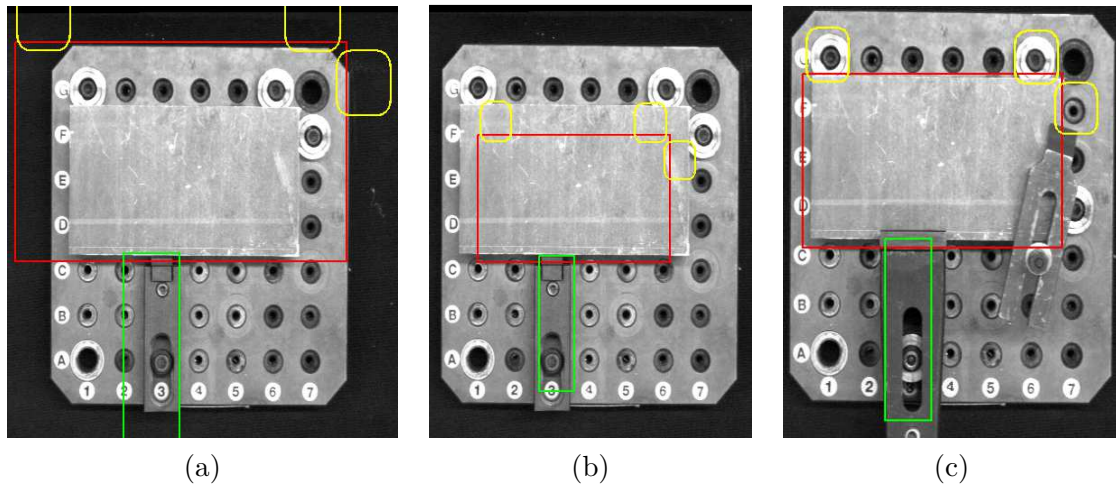


Figure 4.29 – Influence de point de vue dans détermination des régions d'intérêts. (a) la composante Z du vecteur translation veut $Z = 700\text{mm}$, (b) la composante Z du vecteur translation veut $Z = 1200\text{mm}$ et (c) cas d'une pose correcte avec une bride qui n'a pas la bonne dimension

peut ainsi affecter la convergence de la solution globale. D'autre part, notre méthode s'appuie sur la topologie des éléments de bridage pour déterminer la pose. Sous une projection perspective d'un modèle numérique complexe, les légères déformations sur certaines zones représentant les éléments de bridage, peut contribuer à affecter la délimitation de la région de l'élément en question. Autrement dit, les primitives pertinentes sélectionnées pour le calcul de pose ne caractérisent pas la géométrie locale du montage d'usinage. Seule l'analyse des surfaces du modèle numérique CAO complet peut donner cette information. Néanmoins, utiliser l'ensemble des surfaces ou triangles donnés par le modèle CAO du montage d'usinage pour l'appariement n'est pas une solution adaptée, ni optimale.

4.7.3 Cas du montage totalement différent du modèle numérique

Nous avons étudié dans les paragraphes 4.7.1 et 4.7.2 l'impact des paramètres de la matrice de pose sur la localisation des régions d'intérêt. Ce paragraphe se focalise sur l'étude des situations appelées "extrêmes", dans lesquelles l'opérateur peut commettre une erreur sur le choix de réalisation du montage attendu par la machine.

En effet, pour des raisons de diversité des situations "conforme" et "non-conforme", la pose peut être correcte, mais le montage réalisé pas "conforme" au modèle numérique attendu. La figure 4.30 illustre cette situation par un exemple dans lequel le montage attendu est celui sur la **plaque Norelem**[®], tandis que l'opérateur a réalisé le montage **Semelle**. Ce cas de figure, reflète la **robustesse de notre méthode pour des situations extrêmes** et malgré cette confusion de la part de l'opérateur, le calcul de pose demeure correct. Cependant, pour la prise de décision automatique sur la conformité du montage observé par la caméra au sein de la machine par rapport au modèle CAO, la nécessité d'une étape d'analyse locale des régions d'intérêts est importante. Cette partie est développée dans le chapitre 5.

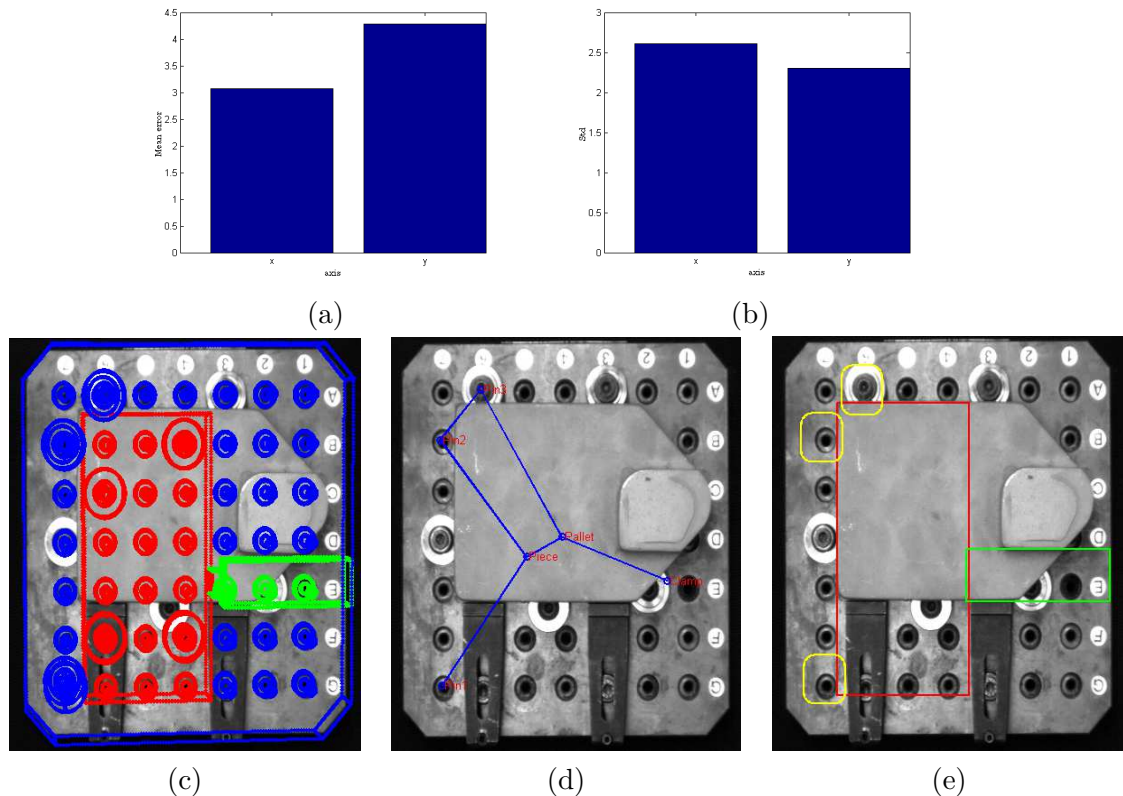


Figure 4.30 – Robustesse de notre méthode pour des cas extrêmes. (a) Erreur moyenne de la pose, (b) écart-type moyen calculé sur les points appariés, (c) modèle numérique du montage Norelem[®] projeté sur l'image du montage Semelle, (d) Graphe des attributs du montage Norelem[®] projeté sur l'image du montage Semelle, (e) localisation des régions d'intérêts

4.8 Conclusion

Ce chapitre a traité le problème de la mise en correspondance des graphes pour le calcul de pose. La méthode d'appariement des graphes que nous proposons repose sur la détermination de la relation géométrique Φ permettant de maximiser la ressemblance entre les graphes réel et théorique. La mesure de ressemblance est donnée par la distance minimale et le nombre des nœuds appariés. Le résultat de l'appariement permet de projeter le modèle CAO sur l'image afin de segmenter, ultérieurement, cette dernière en régions d'intérêt représentant les éléments de bridage. Le recalage du modèle numérique sur l'image se fait par un calcul automatique de pose. Nous avons évalué le calcul de la pose par rapport aux étapes de la chaîne numérique de traitement, partant des contours à l'appariement des graphes, à la fois par rapport aux paramètres d'acquisitions et aux paramètres du modèle. Nous avons montré que notre méthode est robuste, même pour des situations "extrêmes" respectant l'hypothèse de la configuration de la caméra par rapport au montage d'usinage.

Dans le chapitre suivant, nous nous intéressons à l'analyse locale des régions d'intérêt, ce qui va permettre de prendre une décision sur la conformité du montage d'usinage

réalisé en regard du modèle de référence donné en CAO.

Bibliographie

- [Bouguet 2004] J.-Y. Bouguet. Camera calibration toolbox for matlab, mar 2004. (Cité en pages 70, 89 et 90.)
- [Brox 2005] T. Brox, B. Rosenhahn et J. Weickert. Three-Dimensional Shape Knowledge for Joint Image Segmentation and Pose Estimation. In Pattern Recognition, Springer LNCS 3663, pages 109–116, 2005. (Cité en page 69.)
- [Bunke 1983] H. Bunke et G. Allermann. Inexact graph matching for structural pattern recognition. Pattern Recognition, vol. 1, no. 4, pages 245–253, 1983. (Cité en page 70.)
- [Caelli 2004] T. Caelli et S. Kosinov. An eigenspace projection clustering method for inexact graph matching. IEEE Trans. Pattern Analysis and Machine Intelligence., vol. 26(4), pages 515–519, April 2004. (Cité en page 69.)
- [Cox 2005] D. Cox, J. Little et D. O'Shea. Using Algebraic Geometry, Second Edition. Graduate Texts in Mathematics, Springer, vol. 185, pages 364–374, 2005. (Cité en page 70.)
- [Delaunay 1928] B. Delaunay. Sur la sphère vide. volume 1, pages 223–228, 1928. (Cité en page 74.)
- [Dementhon 1995] D. Dementhon et L.S. Davis. Model-based object pose in 25 lines of code. volume 15(1/2), pages 123–141, 1995. (Cité en pages 70 et 103.)
- [Dong-Qing 2004] Zhang Dong-Qing et Chang Shih-Fu. Stochastic Attributed Relational Graph Matching for Image Near-Duplicate Detection. Rapport technique 1-15, 2004. (Cité en page 70.)
- [Duchenne 2011] O. Duchenne, F. Bach, In-So Kweon et J. Ponce. A Tensor-Based Algorithm for High-Order Graph Matching. IEEE Trans. Pattern Analysis and Machine Intelligence., vol. 33, no. Dec. 2011, pages 2383–2395, 2011. (Cité en pages 69 et 73.)
- [Gabriel 1969] Ruben K. Gabriel et Robert R. Sokal. A New Statistical Approach to Geographic Variation Analysis. Systematic Zoology, vol. 18, no. 3, pages 259–278, sep 1969. (Cité en pages 70, 74 et 82.)
- [Gori 2005] M. Gori, M. Maggini et L. Sarti. Exact and Approximate Graph Matching Using Random Walks. IEEE Trans. Pattern Analysis and Machine. Intelligence, vol. 27, no. 7, pages 1100–1111, 2005. (Cité en page 70.)
- [Haralick 1991] R.M. Haralick, C. Lee, K. Ottenberg et M. Nölle. Analysis and solutions of the three point perspective pose estimation problem. pages 592–598, 1991. (Cité en pages 96 et 98.)
- [Horaud 1993] R. Horaud et O. Monga. Vision par ordinateur, outils fondamentaux. 1993. (Cité en page 102.)

- [Lankford 1969] P.M. Lankford. Regionalization : Theory and alternative algorithms. Geographical Analysis, vol. 1, pages 169–212, 1969. (Cité en page 74.)
- [Liu 1990] Y. Liu, T.S Huang et O.D. Faugeras. Determination of camera location from 2D to 3D line and point correspondences. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pages 28–37, 1990. (Cité en pages 102 et 103.)
- [Lowe 2004] G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, vol. 60, no. 2, pages 91–110, 2004. (Cité en page 69.)
- [Matula 1980] D. W. Matula et R. R. Sokal. Properties Of Gabriel Graphs Relevant To Geographic Variation Research And The Clustering Of Points In The Plane. Geographical Analysis, vol. 12, pages 205–222, 1980. (Cité en page 74.)
- [Medasani 2001] S. Medasani et Y. Krishnapuram R. Choi. Graph Matching by Relaxation of Fuzzy Assignments. IEEE Trans. on Fuzzy Systems, vol. 9(1), pages 173–182, 2001. (Cité en page 72.)
- [Moreno-Noguer 2007] F. Moreno-Noguer, V. Lepetit et P. Fua. Accurate non-iterative $o(n)$ solution to the pnp problem. IEEE 11th International Conference on Computer Vision., pages 1–8, 14-21 Oct. 2007. (Cité en page 96.)
- [Nister 2004] D. Nister. An efficient solution to the five-point relative pose problem. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26(6), pages 756–770, 2004. (Cité en page 96.)
- [Quan 1999] L. Quan et Z.D. Lan. Linear n-point camera pose determination. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21(8), pages 774–780, 1999. (Cité en page 70.)
- [Roberto M. 2005] Cesar Jr. Roberto M., E. Bengoetxea, I. Bloch et P. Larrañaga. Inexact graph matching for model-based recognition : Evaluation and comparison of optimization algorithms. Pattern Recognition, vol. 38, no. 11, pages 2099–2113, 2005. (Cité en pages 70 et 73.)
- [Sandhu 2011] Romeil Sandhu, Samuel Dambreville, Anthony Yezzi et Allen Tannenbaum. A Nonrigid Kernel-Based Framework for 2D-3D Pose Estimation and 2D Image Segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence., vol. 33, no. 6, pages 1098–1115, jun 2011. (Cité en pages 69 et 103.)
- [Schmid 1997] C. Schmid et R. Mohr. Local grayvalue invariants for image retrieval. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 5, pages 530–535, 1997. (Cité en page 69.)
- [Shakhnarovich 2005] G. Shakhnarovich, T. Darrell et P. Indyk. In Nearest-Neighbor Methods in Learning And Vision : Theory And Practice, page 252, 2005. (Cité en page 83.)
- [Shi 2000] J. Shi et M. Jitendra. Normalized Cuts and Image Segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22(8), pages 888–905, 2000. (Cité en pages xii, 73, 111, 113 et 114.)

- [Toussaint 1988] G.T. Toussaint. A graph theoretical primal sketch, In G. Toussaint (editor). Computational Morphology - a Computational Geometric Approach to the Analysis of Form, pages 229–260, 1988. (Cité en page 73.)
- [Tsai 1986] R.Y. Tsai. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pages 364–374, 1986. (Cité en page 70.)
- [Wrobel 2001] Bernhard P. Wrobel. Calibration and Orientation of Cameras in Computer Vision. volume 34, pages 7–62. 2001. (Cité en page 92.)
- [Wu 1993] Z. Wu et R. Leahy. An Optimal Graph Theoretic Approach to Data Clustering : Theory and Its Application to Image Segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, no. 11, pages 1101–1113, 1993. (Cité en pages 70 et 114.)
- [Zhang 2000] Z. Zhang. A flexible new technique for camera calibration. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pages 1330–1334, 2000. (Cité en page 70.)
- [Zhang 2006] J. Zhang, A.C. Berg, M. Maire et J. Malik. SVM-KNN Discriminative Nearest Neighbor Classification for Visual Category Recognition. Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2006. (Cité en page 82.)
- [Zhang 2007] J. Zhang, S. Marszalek et C. Schmid. Local Features and Kernels for Classification of Texture and Object Categories : An In-Depth Study. int'l J. Computer Vision and Pattern Recognition, no. 2, pages 213–238, 2007. (Cité en page 82.)
- [Zheng 2006] Y. Zheng et D. Doermann. Robust Point Matching for Nonrigid Shapes by Preserving Local Neighborhood Structures. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 28, no. 4, pages 643–649, 2006. (Cité en pages 69 et 82.)

ANALYSE LOCALE ET PRISE DE DÉCISION

"Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution."
Albert Einstein (1879-1955)

Sommaire

5.1	Introduction et problématique	110
5.2	Méthodes de segmentation par régions testées	111
5.2.1	La segmentation en régions en utilisant les graphes (<i>NCut</i>)	113
5.2.2	Méthodes de segmentation en régions utilisant l'histogramme	115
5.2.3	Détermination de la forme de l'objet : stratégie proposée	118
5.2.3.1	Dilatation	118
5.2.3.2	Érosion	119
5.2.3.3	Ouverture et fermeture <i>Opening and closing</i>	119
5.2.4	Comparaison des méthodes	120
5.3	Le squelette comme outil déterminant la structure topologique	123
5.3.1	Méthodes d'amincissement	123
5.3.2	Méthodes à partir de la carte de distance	123
5.3.3	Méthodes à partir du diagramme de Voronoï	123
5.3.4	Méthodes de morphologie mathématique	124
5.3.5	Régularisation du squelette	124
5.4	Appariement de squelettes et calcul du score	125
5.4.1	Appariement de squelettes à partir de leur graphe de proximité : Méthode proposée	125
5.4.2	Calcul du score d'appariement	125
5.5	Calcul du score d'évaluation des contours et dimension	127
5.6	Les relations entre les paramètres de décision	128
5.6.1	Le squelette <i>VS</i> les contours et la dimension	128
5.6.2	Les contours <i>VS</i> le squelette et la dimension	129
5.6.3	La dimensions <i>VS</i> le squelette et les contours	130
5.7	Système expert flou pour la prise de décision	130
5.7.1	Définition des ensembles flous	131
5.7.1.1	Ensemble triangulaire	131
5.7.1.2	Ensemble trapézoïdale	131
5.7.1.3	Ensemble Gaussienne	132
5.7.2	Les principales opérations floues	133

5.7.3	Fuzzification	134
5.7.3.1	Définition des règles floues	135
5.7.4	Défuzzification et décision	135
5.8	Avons-nous besoin de l'information tridimensionnelle ?	143
5.9	Conclusion	143
	Bibliographie	145

Mots clés : Segmentation en régions, Squelette, Dimension, Évaluation des contours, Système expert flou.

5.1 Introduction et problématique

Ce chapitre précise la méthode permettant de reconnaître les contenus dans les régions d'intérêt déterminées en exploitant la pose, e.g. résultat du chapitre 4. La figure 5.2 présente un exemple du montage Semelle et des différentes régions définies par le recalage du modèle de référence sur l'image du montage réalisé. Afin de connaître la forme de l'objet en question, nous avons besoin d'une image binaire de référence. La figure 5.3 illustre les images binaires des objets de référence créés à partir de la forme de l'objet donnée au modèle CAO en respectant les occultations, objet blanc sur fond noir, pour le cas du montage Semelle.

Pour comparer chaque image binaire théorique à l'image de l'objet réel attendu, nous avons besoin d'extraire correctement sa forme à partir de l'image réelle de luminosité dans sa région. Afin d'améliorer la représentativité de l'image binaire de l'objet nous appliquons des techniques de morphologie¹ mathématique [Forsyth 2002]. Le détail est présenté dans la section 5.2.3.

Une fois la forme de l'objet calculée, nous utilisons trois caractéristiques pour décrire cette dernière : **la structure topologique, la géométrie et la dimension**. Ces caractéristiques sont portées par le **squelette** pour décrire la structure topologique, **les contours** pour décrire la géométrie et la dimension est donnée par **la surface**. Nous montrerons dans la section 5.6 la relation entre ces trois caractéristiques pour décrire la forme de l'objet et leur importance dans la fonction de prise de décision.

Nous attribuons à chaque caractéristique un score de ressemblance. Le score est donné par le rapport entre la mesure de la caractéristique en question calculée sur l'objet réel et la mesure de son correspondant dans l'image binaire de référence. Les sections 5.4 et 5.5 en présentent les détails.

Enfin, l'ensemble des scores calculés à partir de ces trois caractéristiques : **appariement de squelettes, évaluation du contour et dimension** sont analysés dans un système expert flou permettant une décision automatique sur la conformité de la région analysée et par conséquent sur la conformité du montage d'usinage complet. Cette partie est traitée dans la section 5.7. Le chapitre est structuré comme le montre le diagramme de la figure 5.1.

1. Morphologie : Est un mot d'origine grecque "morphē", qui signifie la forme

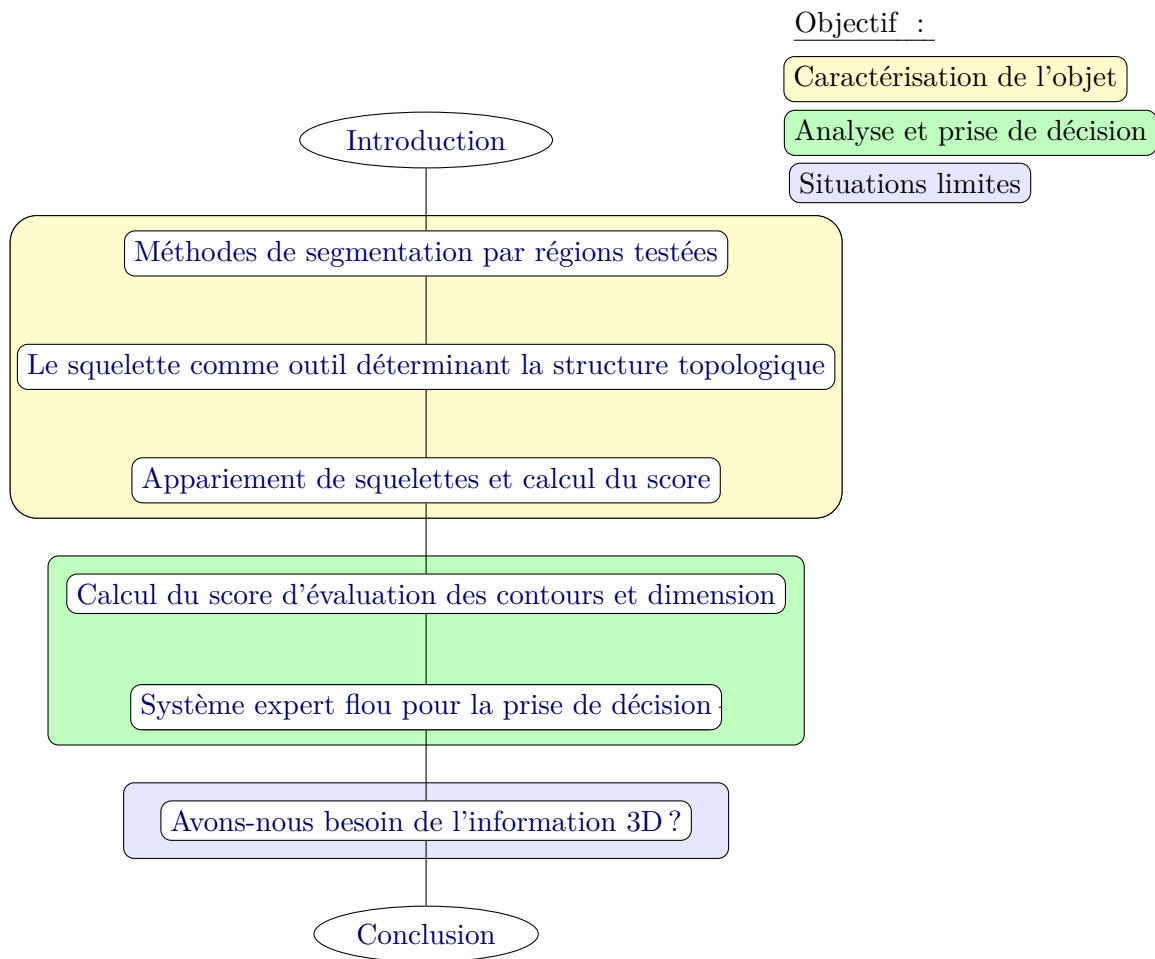


Figure 5.1 – Organigramme du chapitre 5 et objectif de chaque partie.

Avant de procéder au calcul des paramètres descriptifs de l'objet, nous avons besoin d'extraire, dans un premier temps, la forme de ce dernier contenu dans sa région. La section 5.2 qui suit, présente quelques algorithmes de segmentation par régions testées et des éléments de comparaison de performances, donnés par la matrice de confusion.

5.2 Méthodes de segmentation par régions testées

La segmentation de l'image en régions constitue une des étapes importantes dans la prise de décision sur la conformité de l'objet, car le résultat obtenu à l'issue de cette étape conditionne l'interprétation de la forme de l'objet à partir des **paramètres descriptifs** présentés précédemment, e.g. squelette, évaluation des contours (CE) et la surface.

Les méthodes de segmentation par régions peuvent être scindées, selon la technique utilisée, en 3 groupes : méthodes spatiales s'appuyant sur l'information de luminance, i.e. *Neut* [Shi 2000], méthodes utilisant l'information de contours pour remonter à la

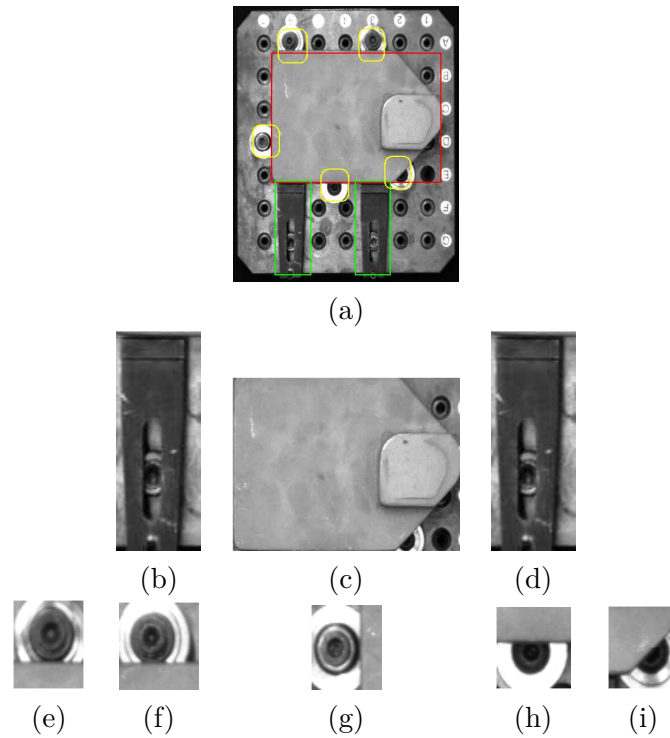


Figure 5.2 – Décomposition du montage d'usinage en régions d'intérêt des éléments de bridage pour la reconnaissance de formes. (a) Image du montage d'usinage, (b), (c), (d), (e), (f), (g), (h) et (i) imagettes des régions d : brut, Bride gauche, bride droite, pion 1, pion 2, pion 3, pion 4 et pion 5, successivement.

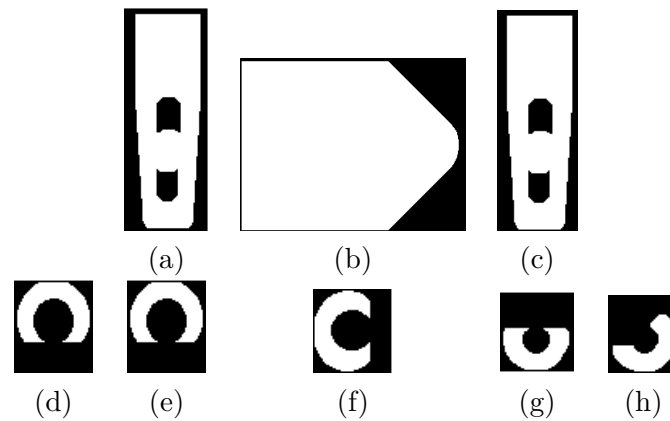


Figure 5.3 – Images binaires de référence des éléments de bridage du montage Semelle. (a) image binaire de la bride1, (b) image binaire du brut (c) image binaire de la bride2, (d) image binaire du pion1, (e) image binaire du pion2, (f) image binaire du pion3, (g) image binaire du pion4 et (h) image binaire du pion5.

segmentation hiérarchique en régions [Arbelaez 2009] et méthodes s'appuyant sur la modélisation de l'histogramme en utilisant un modèle de mélange de Gaussiens *GMM* ou *GGMM* [Nacereddine 2009]. Dans ce dernier groupe, l'estimation des paramètres

du modèle s'appuie sur l'algorithme **EM** (Espérance-Maximisation).

Nous voulions, initialement, utiliser la méthode classique d'Otsu pour trouver automatiquement le seuil de binarisation, suivi par l'application des opérations de morphologie mathématique pour restituer la forme de l'objet originale, le plus fidèle possible. En effet, la méthode d'Otsu est efficace pour des images dont l'histogramme est bimodal. Chaque mode correspond respectivement à la classe fond et à la classe objet. Mais le rapport de surface objet/fond dans la zone d'intérêt conduit à une forme de l'histogramme unimodale, le mode "fond" étant peu marqué. Compte tenu de notre expérience dans la détection des contours, e.g. chapitre 3, pour laquelle le problème étant similaire, mais concerné les classes contours et non-contours, nous avons testés 2 algorithmes réputés performants dans les 2 groupes, méthode de segmentation spatiale, *NCut* et méthode de modélisation de l'histogramme, *GMM*.

5.2.1 La segmentation en régions en utilisant les graphes (*NCut*)

L'idée de l'utilisation des graphes pour la segmentation dans l'algorithme *Normalized Cut* repose sur la construction d'un graphe selon le critère d'homogénéité, i.e. apparence, proximité, etc. à partir des pixels de l'image. Ensuite, on cherche à partitionner le graphe selon le critère d'homogénéité de telle sorte que les parties obtenues représentent les objets d'intérêt. Pour le cas de la segmentation en régions le critère d'homogénéité est donné par l'apparence ou la luminance de l'image, et l'hypothèse employée suppose que chaque objet dans une région de l'image est relativement homogène en distribution de niveau de gris. Le travail introduit par Shi, J et Malik, J. [Shi 2000] fait référence aujourd'hui dans la segmentation d'image par régions en utilisant les graphes.

Formellement, l'algorithme *NCut* proposé par les auteurs s'appuie sur la recherche du graphe couvrant le poids minimal calculé par la différence absolue de niveaux de gris dans un voisinage. Le principe de l'algorithme sur un graphe $G = (V, E)$ est illustré sur la figure 5.4, avec V représentant les nœuds, les pixels, et E les arêtes, le critère d'homogénéité. Les pixels dans le groupe A sont connectés par les arêtes noires, de même pour le groupe B. Les arêtes rouges représentent les relations faibles au sens du critère de similarité choisi, exprimé par la différence absolue de niveau de gris, reliant le groupe A avec le groupe B. Cette relation est donnée par la matrice des poids W . $w(V_x, V_y)$ définit la capacité de l'arête (V_x, V_y) . La matrice des poids W est donnée par l'image des gradients, ce qui signifie que le poids de l'arête (A, B) est différent de zéro pour des gradients non nuls et que les arêtes (A, A) et (B, B) sont nulles ou presque.

La recherche de la partition optimale au sens du critère de similarité, entre le graphe A et le graphe B, se fait en respectant les deux conditions données par l'équation (5.1).

$$\begin{cases} A \cup B = V \\ A \cap B = \emptyset \end{cases} \quad (5.1)$$

Le degré de partitionnement est donné par la fonction coupe d'équation (5.2) suivante :

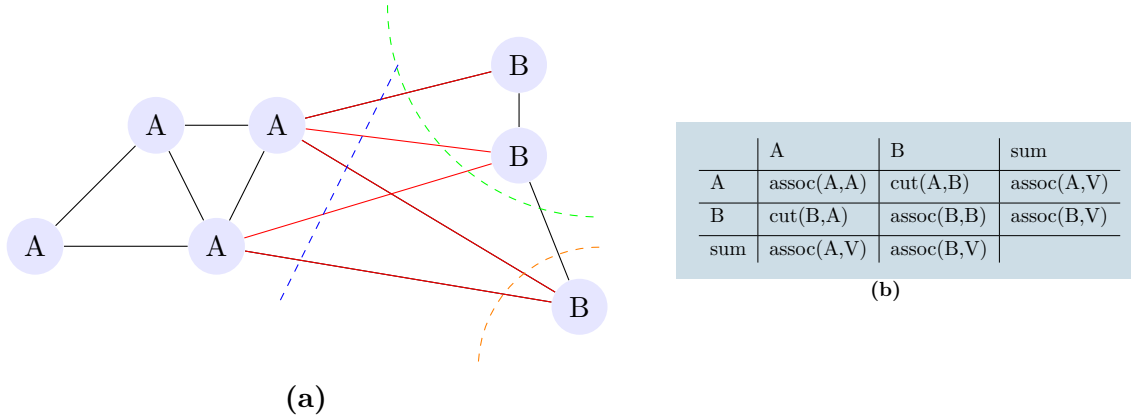


Figure 5.4 – Principe de l’algorithme Ncut [Shi 2000] : (A) un simple graphe et la coupe de poids minimale ; (B) tableau des associations et des coupes pour ce graphe. *assoc* et *cut* sont calculés à partir de la matrice de poids W .

$$cut(A, B) = \sum_{V_x \in A, V_y \in B} w(V_x, V_y) \quad (5.2)$$

Cependant, la recherche de la partition optimale qui minimise cette quantité est un problème relativement long, car la capacité de coupe minimale est fonction du nombre d’arêtes de la partition. Sur l’exemple de la figure 5.4, la coupe bleue donne deux partitions, $P_{b1} = \{A, A, A, A\}$ et $P_{b2} = \{B, B, B\}$; et la coupe verte, les deux partitions $P_{v1} = \{A, A, A, A, B\}$ et $P_{v2} = \{B, B\}$, de même pour la coupe orange, les partitions $P_{o1} = \{A, A, A, A, B, B\}$ et $P_{o2} = \{B\}$. Les coupes verte et orange ont un coût minimal par rapport à la coupe bleue. Cependant, la coupe bleue présente une meilleure délimitation des deux régions A et B .

Les auteurs ont proposé une solution inspirée des travaux de Wu Z. et Leahy R. [Wu 1993] dont l’intérêt est de partitionner le graphe de telle sorte que le moins d’arêtes possible interviennent dans la coupe. Cette solution utilise une nouvelle fonction coût de la fonction coupe, qui permet de supprimer l’influence du nombre d’arêtes dans la coupe, grâce à une étape de normalisation donnée par l’équation (5.3).

$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)} \quad (5.3)$$

avec $assoc(A, V) = \sum_{a \in A, t \in V} w(a, t)$ pour cette nouvelle définition, $assoc(B, V)$ est le critère de similarité. On peut également définir une mesure de l’association normalisée à l’intérieur de chacune des parties de la partition par l’équation (5.4) :

$$Nassoc(A, B) = \frac{assoc(A, B)}{assoc(A, V)} + \frac{assoc(B, B)}{assoc(B, V)} \quad (5.4)$$

A partir de ces deux définitions, le problème de minimisation de la coupe normalisée revient à maximiser l’association correspondante. En effet, supposons x le vecteur indicateur de la partition, avec $x_i = 1$ ssi $i \in A$ et $x_i = -1$ ssi $i \in B$. Soit $D = diag(d_i)$,

avec $d_i = \sum w_{ij}$ le vecteur indicateur de la partition et $W = w_{ij}$ représentant les matrices de pondération, respectivement, symétrique et diagonale. Le problème de minimisation de la coupe pour tout l'ensemble des vecteurs x peut s'écrire sous la forme :

$$\min_y Ncut(x) = \min_y \frac{y^T(D - W)y}{y^T D y} \quad (5.5)$$

Avec $y = ((1+x) - b(1-x))/2$ le vecteur composé de tous les 1^{er} et b^{ieme} éléments, tel que, $\mathbf{y} \cdot \mathbf{d} = \mathbf{0}$. La solution de ce système est donnée par :

$$(D - W)y = \lambda D y \quad (5.6)$$

qui peut être traduit sous la forme du système d'équations suivant :

$$(I - N)z = \lambda z \quad (5.7)$$

Avec $N = D^{-\frac{1}{2}} W D^{-\frac{1}{2}}$ le vecteur propre associé à la seconde plus petite valeur propre de la matrice D .

Cette solution donne de meilleurs résultats pour des images ayant une certaine homogénéité de luminance. La figure 5.5 montre un exemple pour le pion de fixation du montage **Minimanche**. La figure 5.5.a représente la matrice des poids W , la figure 5.5.b le contour de partitionnement et la figure 5.5.c les partitions obtenues de l'image. Cependant, cette méthode demeure relativement longue par rapport à d'autres méthodes classiques de segmentation par régions. De plus elle nécessite un post-traitement pour définir l'objet, car le nombre de régions délimitées sur l'image ne contient pas une information sur leur appartenance à l'objet. Nous proposons ici de ne considérer que la plus grande région dans l'image résultante pour représenter l'objet. Les autres petites régions sont remises au fond.

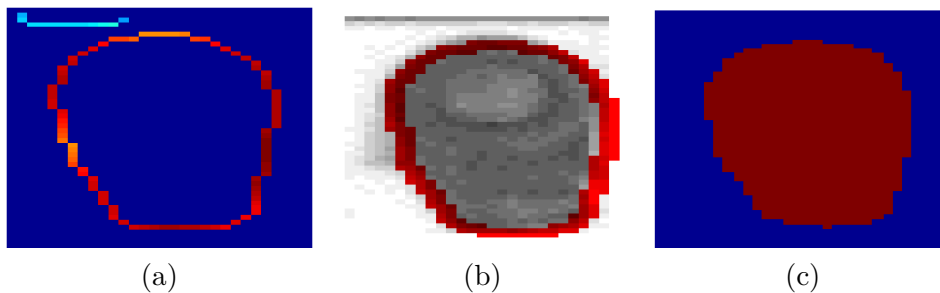


Figure 5.5 – Exemple d'application de l'algorithme *NCut* sur l'image du pion du montage *Minimanche*. (a) le degré de partitionnement calculé sur les contours, (b) les contours de partitionnement et (c) les régions de l'image.

5.2.2 Méthodes de segmentation en régions utilisant l'histogramme

Basé sur le principe de l'algorithme de Maximum de vraisemblance *EM*, les méthodes de segmentation par régions en utilisant l'histogramme s'appuient sur la multi-

modalité (k modes) de la distribution de l'histogramme. Chaque mode représente une zone dans l'image et peut être estimé par un modèle Gaussien en utilisant l'algorithme *EM*. L'histogramme est ainsi modélisé par un mélange de lois Gaussiennes *GMM*, donné par l'équation (5.8).

$$f_k(x_i|\Theta) = \frac{1}{\sqrt{2\pi\sigma_k}} \exp \frac{-(x_i - \mu_k)^2}{2\sigma_k} \quad (5.8)$$

Pour ($k > 1$ mode), le modèle de mélange peut être exprimé par l'équation (5.9) :

$$F(x_i|\Theta) = \sum_{k=1}^K \pi_k f_k(x_i|\theta_k) \quad (5.9)$$

Avec F la densité de probabilité Gaussienne de l'histogramme des niveaux de gris de l'image, $\Theta = (\pi_k, \theta_k)$ le vecteur des paramètres à estimer avec π_k la probabilité *a priori* de la $k^{\text{ième}}$ composante qui satisfait $\pi_k > 0$ et $\sum_{k=1}^K \pi_k = 1$.

Pour k modes, le problème revient à déterminer les k couples $\theta_i(\pi_k, \theta_k)$. En effet, pour un modèle Gaussien, la vraisemblance entre les n observations et le modèle peut être écrite par :

$$L(\Theta) = \prod_{i=1}^n f_k(x_i|\Theta) = \prod_{i=1}^n \sum_{k=1}^K \pi_k f_k(x_i|\theta_k) \quad (5.10)$$

Le principe de l'algorithme *EM* s'appuie sur l'estimation itérative des paramètres du modèle. En introduisant une variable aléatoire $Z = z_i = (z_{i,1} \dots z_{i,n})$ où $z_{i,m} = 1$ si $x_i \in f_k$ et $z_{i,m} = 0$ ailleurs pour définir l'appartenance de l'observation x_i au modèle f_k . Soit $Y = (\chi, Z)$ le nouveau vecteur des observations. La Log-vraisemblance s'écrit :

$$\log L_c(\Theta/Y) = \sum_{n=1}^n \sum_{k=1}^K \pi_k z_{i,m} \log [\pi_k f_k(x_i|\theta_k)] \quad (5.11)$$

- L'étape *E* estime la Log-vraisemblance L_c en utilisant la probabilité *a posteriori* $z_{i,k}^{(t)}$ des paramètres courants du vecteur $\Theta^{(t)}$ donné par $Q = E[L_c\Theta, \Theta^{(t)}]$.
- L'étape *M* maximise cette fonction Q par rapport à Θ pour estimer les nouvelles valeurs des paramètres $\Theta^{(t+1)}$.

Un exemple de résultat est donné par la figure 5.6 pour le cas de la bride du montage Semelle.

Le figure 5.6.a illustre les deux modes estimés. Le mode "rouge" représente la région de la bride dans l'image et le mode "vert", peu marqué, représente le fond. La figure 5.6.b représente le modèle *GMM* complet. Les seuils (S_{k-1}) peuvent être donnés par le minimum local entre deux modes successifs (i) et ($i + 1$), avec $i \leq k$.

Dans cette méthode, le paramètre k est fixé *a priori*. D'autres méthodes [Bouguila 2007] proposent d'utiliser la technique de Monte-Carlo (*MC*) pour estimer les paramètres du modèle. Le paramètre k désignant le nombre de régions dans l'image est calculé en utilisant le facteur de Bayes (*BF*).

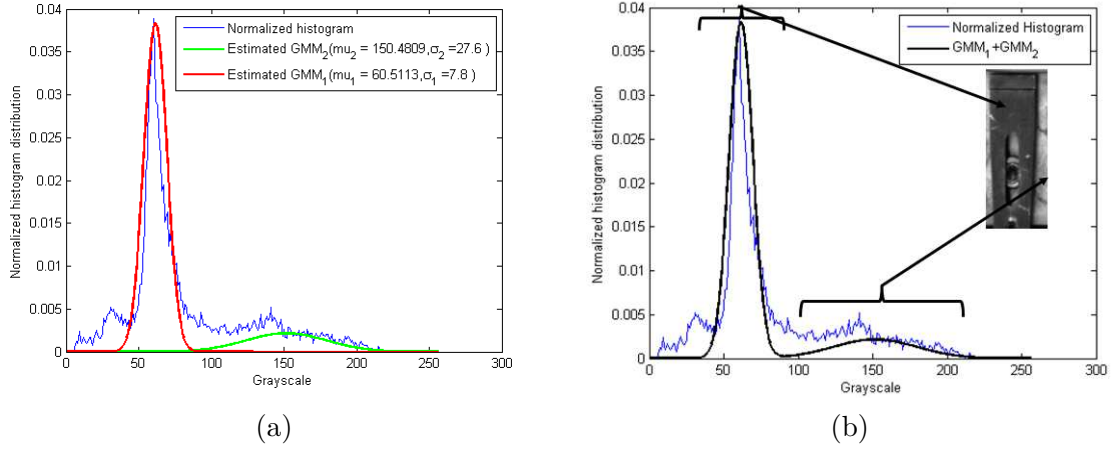


Figure 5.6 – Modélisation de l'histogramme de l'image du pion du montage Semelle par l'algorithme GMM-EM. Application à la segmentation par régions

Pour mieux estimer les modes, Nacereddine *et al.* [Nacereddine 2009] proposent un modèle de mélange de Gaussien généralisé $GGMM$, donné par l'équation (5.12) en introduisant un paramètre d'échelle α_k qui permet d'exprimer la largeur du mode en fonction de l'écart-type.

$$f_k(x/\theta_k) = \frac{\beta_k}{2\alpha_k\Gamma(1/\beta_k)} \exp(-|x - \mu_k|/\alpha_k)^{\beta_k} \quad (5.12)$$

Avec $x \in [0, 255]$, $\theta_k = (\mu_k, \alpha_k, \beta_k)$ représente le vecteur des paramètres à estimer. Le paramètre β_k est donné par l'équation (5.13)

$$\alpha_k = \sigma_k \sqrt{\frac{\Gamma(1/\beta_k)}{\Gamma(3/\beta_k)}} \quad (5.13)$$

Le lecteur intéressé par les approches probabilistes de modélisation de l'histogramme de l'image de luminance peut se référer aux monographies de Forsyth, David. A [Forsyth 2002] et Kevin P. Murphy [Murphy 2012].

Les deux types d'algorithmes présentés peuvent être employés pour le calcul de la forme de l'objet. Néanmoins, par définition l'image de la région ne doit contenir que deux classes Objet/ fond. La littérature est très riche en méthodes de binarisation en régions. La méthode classique d'Otsu, par exemple, peut être utile et suffisante, en la combinant avec des techniques de morphologie mathématique de l'image, ce qui permet de traiter le problème du bruit dû aux pixels "isolés" dans l'image binaire résultante et de combler le vide pour restituer la forme de l'objet.

Nous nous intéressons dans la section suivante 5.2.3 à la méthode retenue. Nous rappelons quelques notions sur les techniques de morphologie mathématique des images. Nous montrons ainsi l'intérêt de l'utilisation de ces dernières pour retrouver la forme "originale" de l'objet, à partir de l'image binaire obtenue par la méthode d'Otsu.

5.2.3 Détermination de la forme de l'objet : stratégie proposée

Pour le calcul de la forme de l'objet de manière la plus "fidèle" possible, nous proposons de binariser l'image de la région de l'objet en utilisant l'algorithme classique d'Otsu. Ensuite, nous appliquons des opérations de la morphologie mathématique de l'image afin de retrouver la forme de l'objet la plus fidèle possible. La méthode proposée s'articule en trois étapes :

- Binarisation de l'image par l'algorithme d'Otsu.
- Application d'opérations de la morphologie mathématique sur l'image binaire permettant de restituer la forme originale de l'objet.
- Dans certains cas, l'étape 2 conduit à intensifier les pixels dûs au bruit en créant des petites régions isolées. Nous proposons de ne considérer que la région la plus grande pour représenter l'objet, les autres régions sont remises au fond.

Cette stratégie de traitement de l'image rend plus facile l'extraction de la forme de l'objet, car la répartition des niveaux de gris dans l'image est supposée homogène, e.g. l'image dans le cas "idéal" ne doit contenir que deux classes : Objet et fond.

Les opérations de morphologie mathématique d'image permettent aussi d'analyser la connectivité des pixels pour en extraire la forme de l'objet. Les applications de morphologie mathématique peuvent être considérées comme un pré ou post traitement, permettant de réaliser différentes opérations : filtrage, amincissement, élagage, squelettisation, calcul de l'enveloppe convexe, etc.

Les deux opérations principales de morphologie sont la **dilatation** et l'**érosion**. Ces opérations utilisent un filtre de taille ($k \times k$), appelé élément structurant, donné dans un voisinage de 4 ou de 8 pixels, comme le montre la figure 5.7.

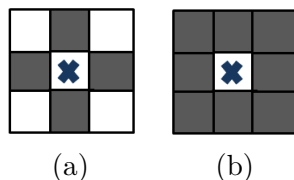


Figure 5.7 – éléments structurants utilisés dans la morphologie d'image. (a) de voisinage 4, (b) de voisinage 8.

5.2.3.1 Dilatation

La dilatation d'une image binaire (A) à partir d'un élément structurant (B) de taille ($k \times k$), notée $A \oplus B$, est une opération mathématique similaire à la convolution qui consiste à étendre la surface des pixels proches, ce qui permet de combler les vides dus au bruit et d'éliminer les pixels "perturbateurs" isolés. Le choix de la taille de l'élément structurant (B) joue un rôle important sur la forme de l'objet final calculé. Il permet d'étendre les pixels connectés et donc de combler plus de vide. Cependant, le risque de fusionner des pixels isolés augmente. Un exemple est donné sur la figure 5.8.

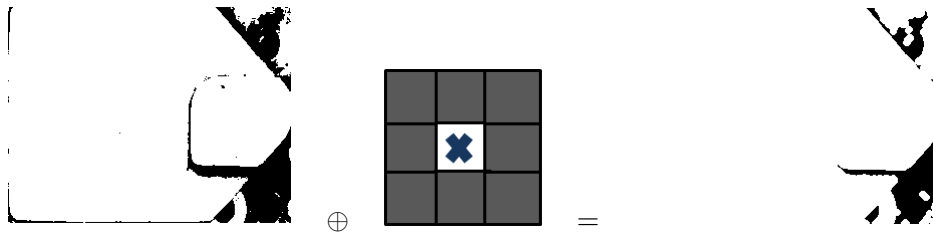


Figure 5.8 – Opération de dilatation en utilisant un filtre de voisinage 8×8 .

5.2.3.2 Érosion

L'érosion est l'opération duale de la dilatation. Elle consiste à dilater les pixels du fond "0". Cette opération a comme objectif de réduire la surface des pixels isolés. Un exemple d'opération de l'érosion en utilisant un élément structurant de voisinage (8×8) est donné sur la figure 5.9.

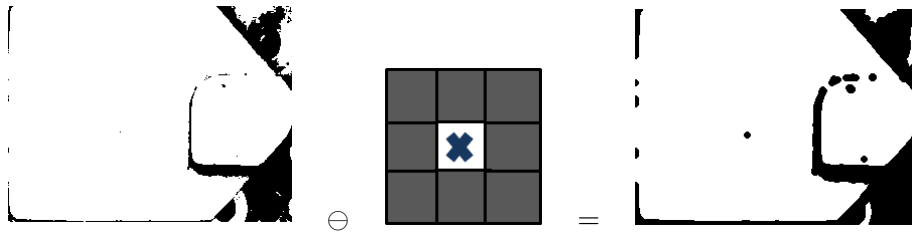


Figure 5.9 – Opération de dilatation en utilisant un filtre de voisinage 8×8 .

Il est à noter sur la figure 5.9 que les zones n'appartenant pas au brut peuvent être réduites avec l'opération d'érosion.

5.2.3.3 Ouverture et fermeture *Opening and closing*

Les deux opérations évoquées précédemment, dilatation et érosion, peuvent être combinées sous forme de deux opérations morphologiques appelée Ouverture, *Opening* ou fermeture, *Closing*. L'ouverture, notée $A \circ B = (A \ominus B) \oplus B$, consiste à appliquer une opération d'érosion suivie d'une dilatation sur l'image binaire. L'objectif de cette opération revient à éliminer les pixels de perturbations "isolés" et à combler le vide des pixels contenant la forme de l'objet. Un exemple de l'opération ouverture en utilisant un élément structurant de connexité (8) pour le brut du montage Semelle, est donné sur la figure 5.10.b. Pour l'opération fermeture, l'exemple est donné sur la figure 5.10.c

L'opération *closing* notée $A \bullet B = (A \oplus B) \ominus B$ est une combinaison de dilatation et d'érosion. Elle consiste à combler le vide et à fermer les trous dûs au bruit.

Certains pixels appartenant à l'objet dans la figure 5.10 passent au fond, faux négatifs, et certains pixels du fond passent à l'objet, faux positifs en référence à l'objet théorique, après application de la fermeture ou l'ouverture. L'ouverture filtre les faux positifs, tandis que la fermeture filtre les faux négatifs.

Comparant les deux résultats (b) et (c) de la figure 5.10, il est facile de remarquer

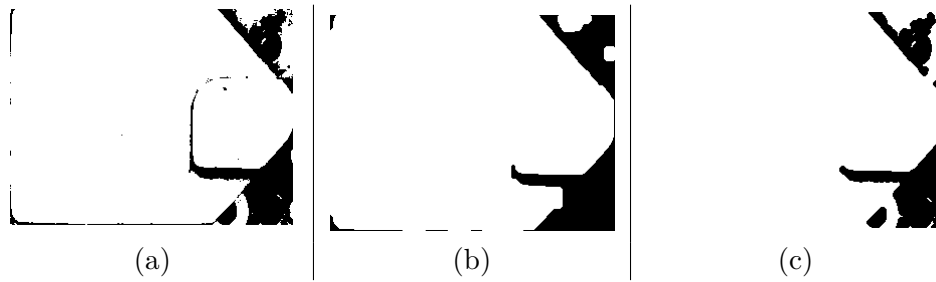


Figure 5.10 – Opérations *Opening* et *Closing* en utilisant un élément structurant de voisinage 4×4 .

que l'ordre de l'opération est important, malgré le fait que les deux opérations utilisent le même élément structurant pour le filtrage.

En général, l'application de la dilatation suivi par l'érosion permet de retrouver les proportions originales de l'objet avec des comblements du vide intérieur.

D'autres opérations morphologiques peuvent être aussi utiles comme par exemple l'extraction de contours, calcul du squelette, amincissement, etc.

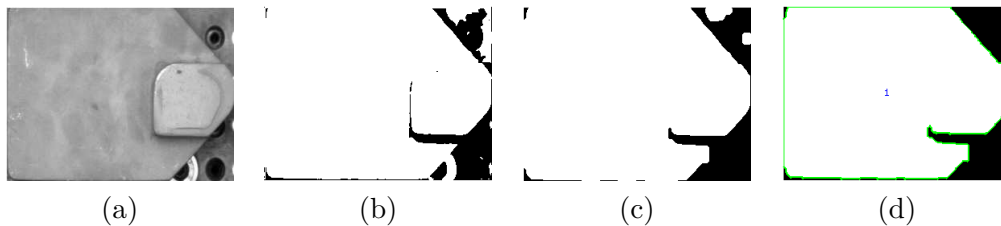


Figure 5.11 – Résultat de différentes étapes de traitement proposées pour le calcul de la forme de l'objet, (a) image de luminance de la région de brut, (b) binarisation par la méthode d'Otsu, (c) application de l'opération de morphologie mathématique, ouverture et (c) restitution de la forme de l'objet par le calcul de la plus grande région.

5.2.4 Comparaison des méthodes

Les trois méthodes donnent de bons résultats. Nous avons évalué ces méthodes par rapport à la matrice de confusion. Rappelons que nous intéressons à déterminer la forme de l'objet la plus correcte possible. L'intérêt de la matrice de confusion repose sur le calcul des taux de classification par rapport à l'image binaire de référence, donnés par les indices : Vrai Positif (**VP**), Vrai Négatif (**VN**), Faux Positif (**FP**) et Faux Négatif (**FN**) permettant d'apporter une information sur la qualité de forme de l'objet calculé.

Nous avons réalisé cette étude sur différents objets de référence pour différents montage d'usinage. La figure 5.12 illustre le résultat pour trois éléments de la scène. Dans la première ligne, le brut du montage sur la plaque Norelem[®], chaque colonne représente le résultat de la méthode testée. La seconde ligne représente les résultats du test pour la bride du montage Semelle et la dernière ligne le pion du montage

Minimanche. Il est à noter que nous n'avons pas appliqué d'opération morphologie pour les deux autres méthodes, car elles sont réputées performantes et sophistiquées.

La précision donnée par l'équation (5.14) calculée à partir de la matrice de confusion ci-dessous, est en moyenne supérieure à 80%, ce qui signifie que les trois méthodes donnent de bons résultats et qu'elles peuvent être utilisées pour notre application. Nous avons adopté la méthode d'Otsu pour sa simplicité et sa rapidité d'exécution.

$$Précision(\%) = \frac{VP(\%) + FN(\%)}{VP(\%) + VN(\%) + FP(\%) + FN(\%)} \quad (5.14)$$

		Référence(%)	
		Objet	Fond
Calculé(%)	Objet	VP	FP
	Fond	VN	FN

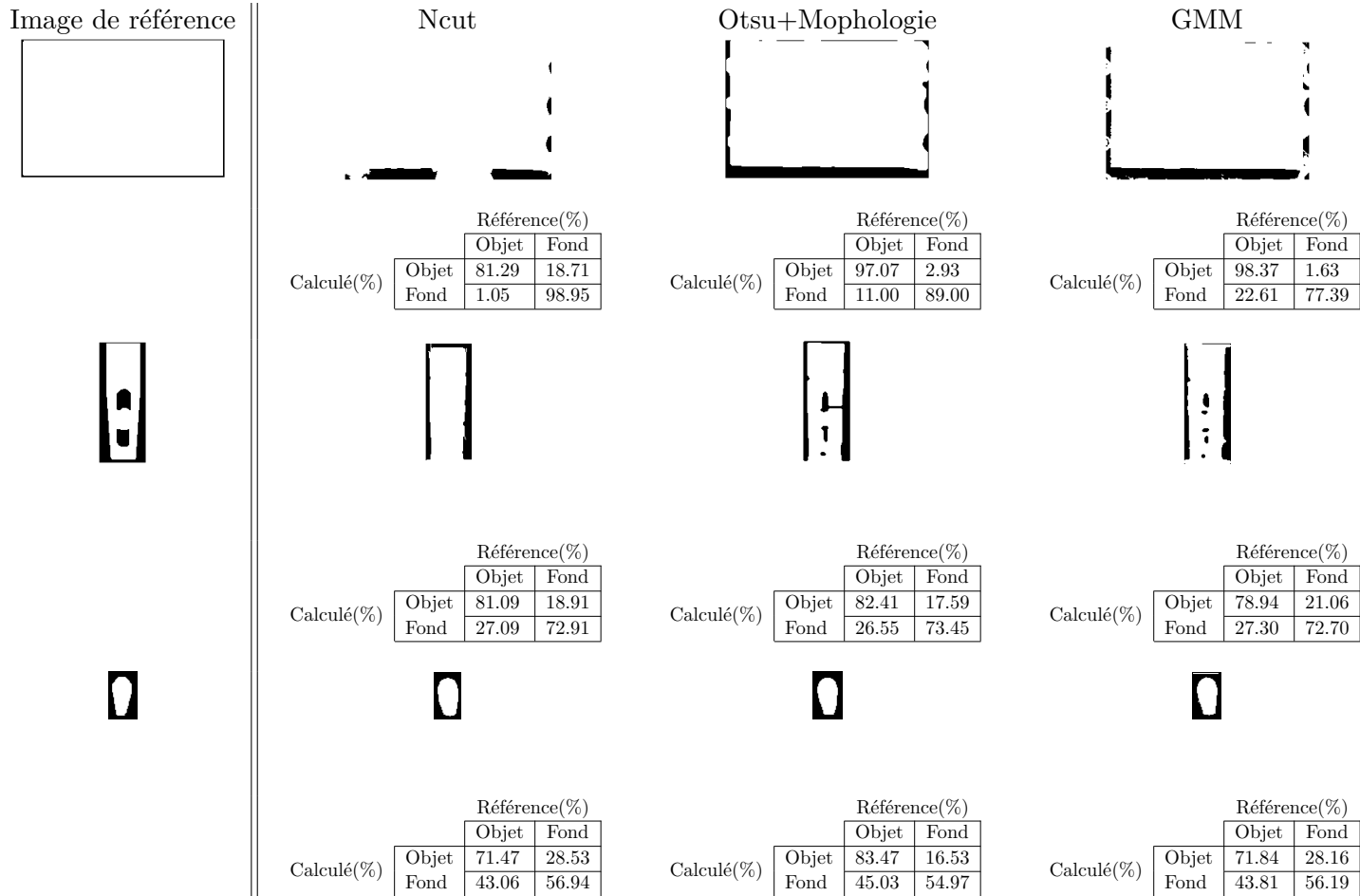


Figure 5.12 – Résultats de comparaison des méthodes de segmentation par régions testées.

5.3 Le squelette comme outil déterminant la structure topologique

Le squelette est un descripteur important de la structure, qui peut être utile pour la reconnaissance de forme tout en gardant un minimum de pixels. La représentation d'une forme par son squelette peut garantir la structure de l'objet original. Les méthodes utilisant le squelette pour la reconnaissance de formes sont nombreuses et peuvent être classées, selon la technique utilisée, en 4 groupes que nous précisons brièvement dans les paragraphes suivants.

5.3.1 Méthodes d'amincissement

Ces méthodes [Arcelli 1985b], [Pudney 1998], [Xie 2003], recherchent à supprimer itérativement les contours, frontières de l'objet tout en conservant sa géométrie. L'inconvénient majeur de cette méthode est qu'elle reste sensible au bruit sur les contours.

5.3.2 Méthodes à partir de la carte de distance

Ces approches basées sur la carte de distance [Arcelli 1985a], [Ge 1996], calculent le squelette à partir de la transformée en distance de l'image binaire. Les maxima locaux de la carte de distance sont les points de l'axe médian du squelette. Les méthodes utilisant la carte de distance garantissent souvent la localisation, mais pas la connectivité, ni la complétude [Arcelli 1985a], [Ge 1996].

La transformée en distance (DT) d'une image binaire consiste à attribuer à chaque pixel la valeur de la distance par rapport au pixel le plus proche. L'algorithme initial du calcul de cette transformation s'appuie sur la distance euclidienne. Il est caractérisé par une complexité en $O(n)$, où n est le nombre des pixels dans l'image [Breu 1995], [Arcelli 1989] impliquant un temps de calcul prohibitif. Pour pallier ce problème du temps de calcul, des nouvelles approches [Meijster 2000] utilisent des définitions approximatives de la distance euclidienne, i.e. distance de Manhattan, distance de Chanfrein, etc.

Ce type d'approche permet de reconstruire la forme de l'objet initial à partir de squelette dans une opération réversible, mais elle ne garantit pas la connectivité.

5.3.3 Méthodes à partir du diagramme de Voronoï

Les méthodes basées sur le diagramme de Voronoï [Ogniewicz 1995], [Brandt 1992], dites continues, cherchent à placer les centres des cercles de diamètres maximaux contenant la forme polygonale de l'objet. Le diagramme de Voronoï est proposé initialement pour des objets de forme polygonale.

Le squelette de Voronoï est un partitionnement de l'espace des points d'intérêt de l'objet en régions convexes appelées les sites de Voronoï. Chaque point appartient au polygone délimité par les contours. Les contours de Voronoï sont les frontières des régions de Voronoï. Supposons C l'ensemble de points caractérisant la forme de l'objet F , le squelette de Voronoï peut être écrit :

$$S_v = \bigcup_{i,j \in V_{sites}} [s_{ij} \cap F] \quad (5.15)$$

Avec s_{ij} le contour de Voronoï séparant les régions $V(i)$ et $V(j)$ définies, respectivement, par les sites v_i et v_j .

Cette méthode possède au moins l'avantage de favoriser la connectivité. Les contours de Voronoï sont fortement connectés. Le squelette obtenu par cette méthode permet de garder la structure de l'objet. Par contre, la sensibilité du calcul des régions à partir des contours rend cette méthode sensible au bruit.

5.3.4 Méthodes de morphologie mathématique

Les méthodes de morphologie mathématique [Vasilevskiy 2002], [Dimitrov 2000], sont un autre type de méthodes d'amincissement basées sur l'érosion successive du contour de l'objet. Ils sont invariants par rotation, mais la connectivité et la localisation ne sont pas garanties [Dimitrov 2003].

5.3.5 Régularisation du squelette

La technique la plus connue pour résoudre le problème de sensibilité du squelette est la méthode d'"élagage", *pruning*. L'élagage peut être considéré comme un post-traitement ou une étape implicite dans le calcul du squelette.

En général, le squelette obtenu à partir de l'image binaire de l'objet est souvent sensible au bruit dû à la segmentation. Cette sensibilité peut créer des branches non significatives de la forme qui peuvent affecter la reconnaissance de l'objet. Les algorithmes d'élagage, *pruning*, sont un type d'algorithme qui cherche à supprimer les branches parasites dans le squelette tout en gardant sa topologie. Xiang B *et al.* [Bai 2007] proposent une nouvelle méthode s'appuyant sur le partitionnement de contours en utilisant la fonction d'évaluation discrète de la courbe, *Discrete Curve Evolution (DCE)*. Le processus se résume à l'élimination successive des nœuds dans le squelette formant des petites zones dans le polygone convexe contenant l'objet. Cette solution peut être exprimée aussi par une approximation polygonale des contours, *boundaries*, en segments. Chaque nœud dans le squelette est, ensuite, relié au segment tangentiel à son cercle de diamètre maximal. Si plusieurs nœuds appartiennent au même segment, on supprime les nœuds dont le diamètre du cercle retenu est inférieur aux autres afin de garder un seul nœud pour chaque segment. Ils montrent dans un autre travail [Bai 2008] son utilité dans le calcul de la mise en correspondance de squelettes. Shaked D. et Bruckstein A.M. ont proposé dans [Shaked 1998] une analyse comparative de ce type de méthodes basées sur plusieurs paramètres : épaisseur maximale de contour, disque maximal contenant le point, la longueur des branches, etc.

Dans notre travail nous nous n'intéressons pas à la régularisation de squelette, car d'une part, la forme de l'objet obtenue est souvent stable et continue et d'autre part, les branches parasites ne sont pas prises en compte dans le calcul du score d'appariement entre le squelette réel et son correspondant de référence. Le détail est donné dans la section 5.4. En plus, contrairement à d'autres travaux s'appuyant seulement sur

l'appariement de squelettes pour la reconnaissance de formes, nous utilisons d'autres paramètres descriptifs de la forme, à savoir les contours et la dimension donnée par la surface.

5.4 Appariement de squelettes et calcul du score

5.4.1 Appariement de squelettes à partir de leur graphe de proximité : Méthode proposée

La méthode proposée pour la mise en correspondance des squelettes s'appuie sur le même principe que la méthode de mise en correspondance des graphes présentée dans le chapitre 4.

Nous calculons, dans un premier temps, le graphe de proximité à partir du squelette de l'objet. Nous proposons de reprendre la définition du type du nœud pour caractériser les différentes parties du graphe. Cette fois-ci, nous attribuons un vecteur d'étiquette, labelisation, $P = (x_i, y_i, 1)$ pour les nœuds terminaux et les nœuds de jonction sont représentés par le vecteur d'étiquette $Q = (x_j, y_j, 0)$. Les arêtes entre les nœuds terminaux et les nœuds de jonction représentent les arêtes de coupures. Le problème peut être modélisé par la recherche du graphe de référence (H) dans le graphe réel (G).

En plus, dans cette étape d'analyse locale, nous n'avons pas de problème de mise en échelle (\mathcal{Y}), ni de rotation (Φ), ni de translation (T), car ces informations ont été calculées à partir de l'estimation automatique de la pose. Autrement dit, la région de l'objet en question déterminée par la pose tient compte de l'orientation de l'objet attendu, la translation est supposée nulle et la dimension est donnée par le graphe des attributs (ARG).

Basé sur l'appariement des nœuds terminaux et de jonction, l'algorithme se résume à la recherche du voisin proche donné par la distance minimale pour chaque nœud de référence dans le graphe réel, selon l'équation (5.16). Si un nœud réel dans le graphe (G) ne trouve pas de voisin dans le graphe de référence (H), celui-ci n'est pas pris en compte dans le calcul du score d'appariement. Ceci explique que les nœuds parasites ne peuvent pas affecter l'appariement et par conséquent le calcul du score. Un exemple de résultat pour le brut du cas conforme du montage de la plaque Norelem[©] est donné sur la figure 5.13.

$$S(G|H) = \underbrace{\min_{i,j \in \mathcal{I}, \mathcal{J}} (Dist(v_i, w_j | \mathcal{Y}, \Phi, T))}_{\text{nœuds appariés}} \quad (5.16)$$

5.4.2 Calcul du score d'appariement

Pour qualifier la reconnaissance de formes à partir de squelettes, certains travaux [Bai 2008] s'appuient sur la mesure de ressemblance portée par les nœuds terminaux. Ils estiment que l'organisation topologique des nœuds terminaux est suffisante pour caractériser la forme de l'objet à reconnaître.

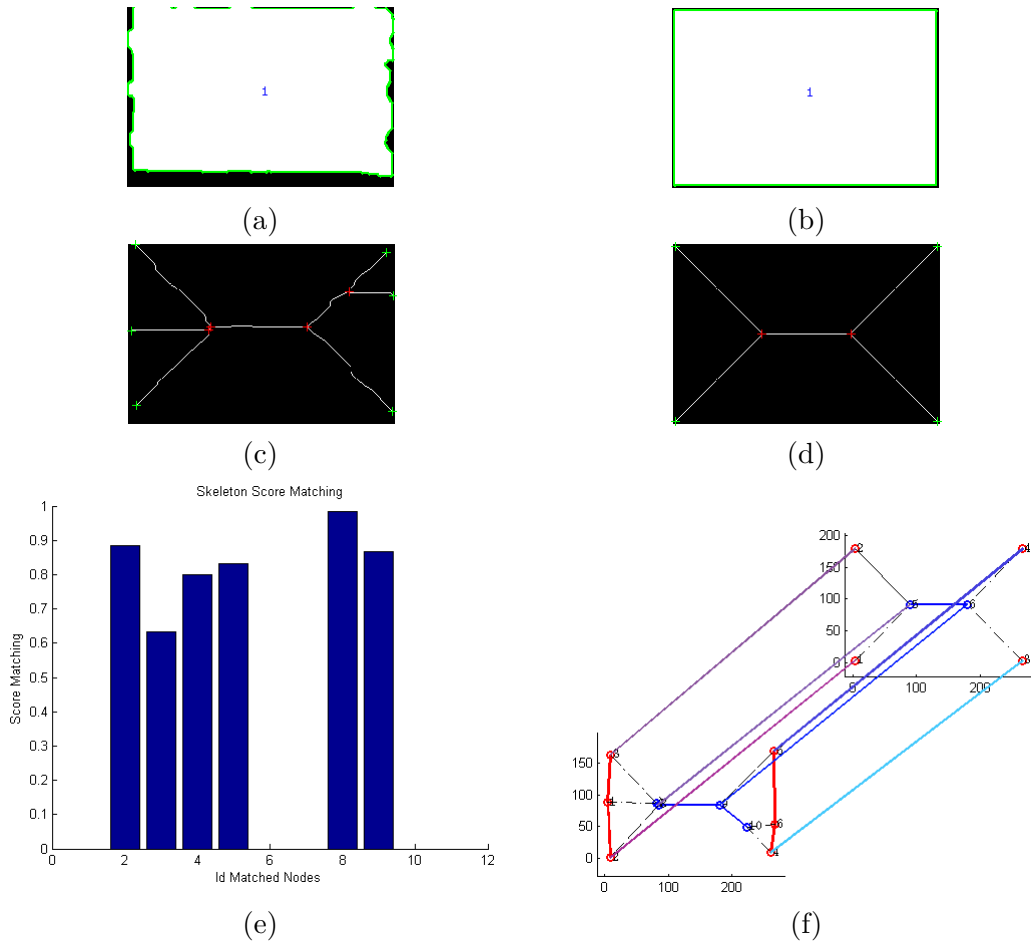


Figure 5.13 – Mise en correspondance de squelettes du brut. (a) image du brut réel calculé, (b) image de référence du brut théorique, (c) squelette correspondant à (a), (d) squelette correspondant à (b), (e) score d'appariement pour chaque nœud du graphe réel, (f) appariement des graphes

Nous proposons dans notre travail d'exploiter à la fois les nœuds terminaux et les nœuds de jonctions dans le calcul du score d'appariement. Cette mesure se traduit dans l'équation (5.17) par une normalisation de la distance de chaque nœud de référence par rapport à la distance maximale entre les paires de nœuds appariés, sans prendre en compte leur nature.

$$Score(v_i, w_j) = 1 - \frac{Dist(v_i, w_j)}{\max(Dist(P, Q))} \quad (5.17)$$

Avec P et Q les vecteurs d'étiquetages contenant les coordonnées des nœuds du graphe (G) et (H) respectivement.

Un résultat est donné sur la figure 5.13.f. L'ensemble des nœuds (de 1 à 6) du graphe de référence (H) ont été appariés, représentés par 6 barres dans le diagramme de barre de la figure 5.13.e. L'axe des abscisses dans ce dernier représente le numéro des nœuds donné par le graphe réel (G). Les nœuds 1, 6 et 7 par exemple dans le

graphe (G) n'ont pas été appariés, donc ils ne sont pas pris en compte dans le calcul du score d'appariement.

Le **score final d'appariement** est donné par la **valeur moyenne** des scores d'appariement de l'ensemble des nœuds.

5.5 Calcul du score d'évaluation des contours et dimension

Le score de l'évaluation des contours (CE) entre l'image des contours de référence et l'image des contours réels repose sur la précision de pose. Pour pallier le problème de l'erreur de la pose, évaluée en moyenne à moins de 3 pixels, nous proposons de faire un balayage avec un masque de taille (5×5) sur les pixels de contours de références par rapport aux pixels des contours réels, pour compter les pixels appartenant aux contours dans les deux images. En langage de matrice de confusion ce rapport peut être exprimé par le ratio entre les vrais positifs et les faux positifs.

Il est à noter que ce paramètre (CE) nécessite à la fois la **présence du contour** et la **bonne localisation** de ce dernier. Les deux "sous-paramètres" sont indépendants l'un de l'autre. Le premier "sous-paramètre" dépend des performances de l'algorithme de calcul des contours et de l'homogénéité en luminance de l'objet. Dans [Liu 2010] par exemple, les auteurs proposent une méthode de localisation d'un objet de référence à partir de ces contours dans l'image en utilisant la méthode de chanfrein, *Chamfer Matching*. Dans cette approche, le processus de calcul de la distance suppose que les contours de l'objet sont convexes, ce qui n'est pas toujours le cas dans nos scènes. L'appariement entre les deux formes revient à minimiser la somme des distances des points de l'objet dans l'image, par rapport à l'objet de référence. Le second "sous-paramètre", donné par la localisation, dépend essentiellement de la précision de la pose. Cela justifie le choix de la taille du masque, e.g. l'erreur moyenne de la pose est évaluée à 3 pixels, et explique aussi que le score calculé sur certains objets conformes demeure relativement moyen. La figure 5.14 en présente un exemple. Il faut donc prendre en compte ce cas de figure dans la méthode de décision.

Enfin, le score de dimension donnée par la surface est calculé par le rapport de la superficie de l'objet réel sur celle de l'objet de référence, selon l'équation (5.18)

$$S_{Dim} = \begin{cases} \frac{Aire(Obj_{Real})}{Aire(Obj_{Ref})} & \text{si } Aire(Obj_{Real}) \leq Aire(Obj_{Ref}) \\ \frac{Aire(Obj_{Ref})}{Aire(Obj_{Real})} & \text{ailleurs} \end{cases} \quad (5.18)$$

La section 5.6 présente les relations entre ces paramètres et nous montrerons ainsi la nécessité de leur combinaison pour la reconnaissance de formes et la prise de décision.

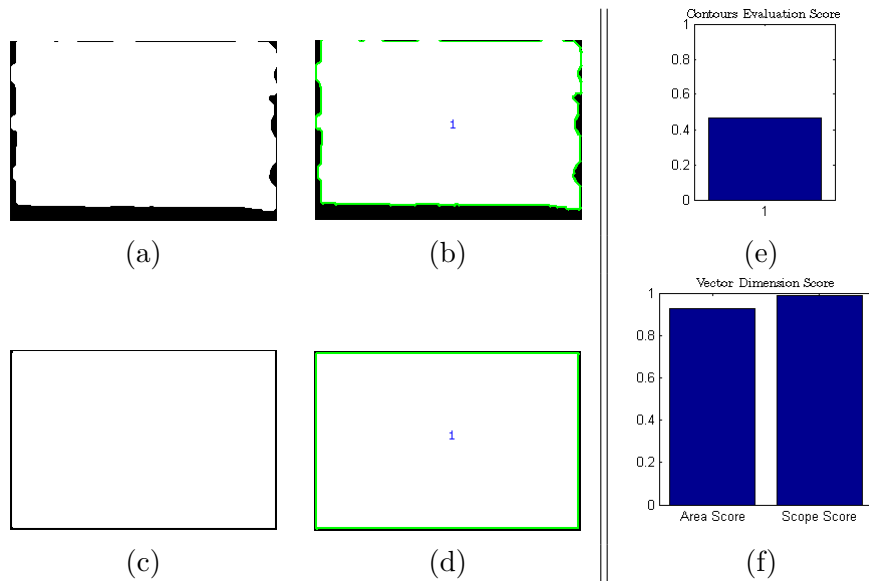


Figure 5.14 – Calcul du score d'évaluation de contours et le score de dimension. (a) image binaire du brut réel calculé, (b) image des contours du brut réel, (c) image binaire du brut de référence, (d) image des contours du brut de référence, (e) score d'évaluation de contours, (f) score de dimension donné par la surface

5.6 Les relations entres les paramètres de décision

Avant de développer un système de vision capable de qualifier l'incertitude et de décider automatiquement de la conformité des montages d'usinage, il faut mesurer les relations entre les paramètres indispensables pour la prise de décision, soit en proposant un modèle mathématique, souvent probabiliste, soit en utilisant des mesures possibilistes, dans le cas où la modélisation du comportement du système est insuffisante et nécessite l'intégration d'un raisonnement logique. En effet, l'étude de la relation entre les variables descriptives de la forme est une étape importante qui pourrait contribuer à ce raisonnement logique.

Dans notre travail, nous proposons d'utiliser trois caractéristiques pour décrire la forme, la géométrie et la dimension afin de reconnaître l'objet. L'utilisation de ces trois paramètres permet de renforcer la décision pour gérer les problèmes de redondance ou de confusion. Les paragraphes suivants décrivent les relations entre les trois caractéristiques et montrent que l'utilisation d'une seule ne peut pas suffire pour prendre une décision sur la conformité de l'objet et par conséquent la conformité du montage réalisé.

5.6.1 Le squelette VS les contours et la dimension

Nous nous intéressons dans cette partie à montrer la relation entre la forme donnée par le squelette et les deux autres caractéristiques. En effet, le squelette d'une forme n'est pas unique, surtout pour des objets polyédriques, sphérique et cylindriques, que nous retrouvons principalement dans nos scènes. On peut trouver par exemple deux

formes différentes ayant le même squelette. La figure 5.15 en montre un exemple. Sur cet exemple, les deux formes ont le même squelette, mais elles n'ont pas les mêmes contours, ni la même dimension, e.g. superficie.

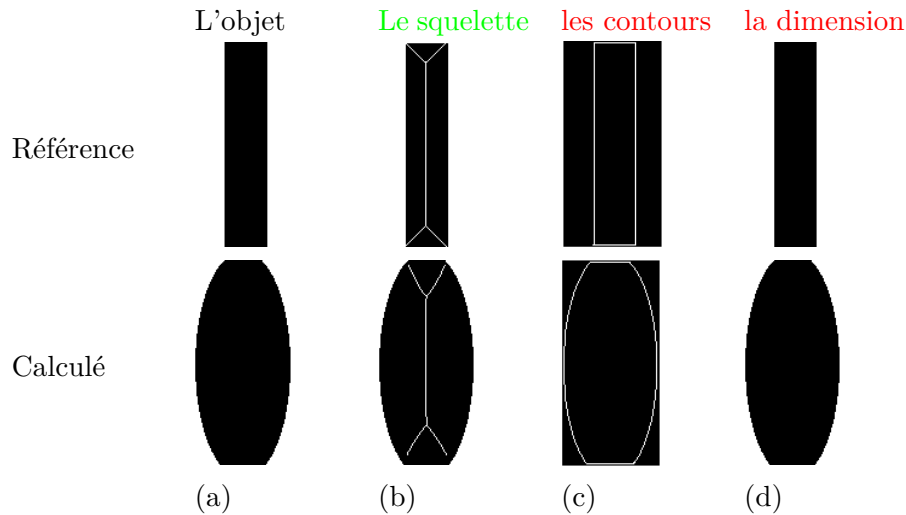


Figure 5.15 – Deux formes ayant le même squelette, mais elles n'ont pas la même dimension ni les mêmes contours.

Il est évident que le squelette seul n'est pas suffisant pour la reconnaissance de l'objet permettant la prise de la bonne décision.

5.6.2 Les contours VS le squelette et la dimension

Pour renforcer la description de la forme de l'objet donnée par le squelette, nous exploitons les contours pour décrire la ressemblance géométrique entre l'objet de référence et l'objet réel.

La figure 5.16 illustre un exemple permettant de comprendre la relation entre les contours et les deux autres paramètres : le squelette et la superficie.

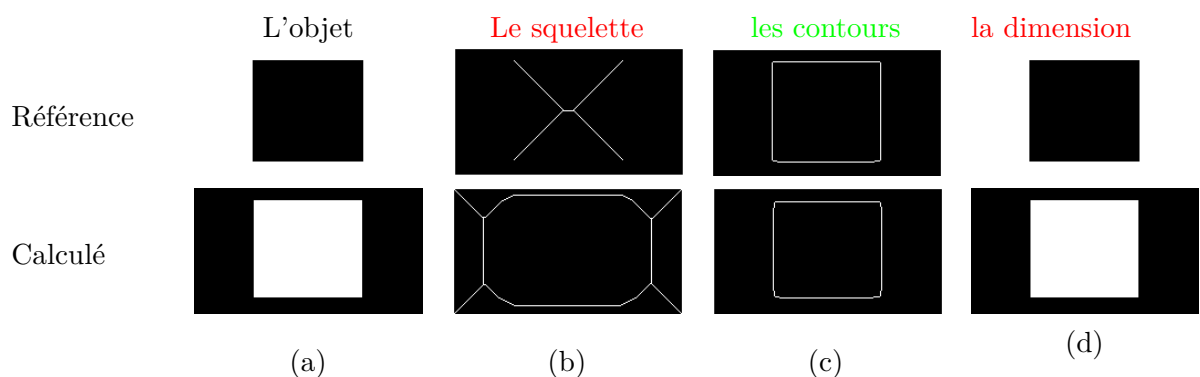


Figure 5.16 – Deux formes ayant les mêmes contours, mais elles n'ont pas la même dimension ni le même squelette.

Il est également évident que les contours seuls ne sont pas suffisants pour la reconnaissance de l'objet. Néanmoins leur combinaison avec le squelette peut assurer la prise de la bonne décision.

5.6.3 La dimensions VS le squelette et les contours

Enfin, la dimension quant à elle vient renforcer le vecteur descriptif de l'objet par la mesure de sa superficie. La figure 5.17 montre la relation entre la dimension et les deux autres paramètres descriptifs de la forme. Il est compréhensible que l'utilisation unique de la dimension, donnée par la superficie, pour décrire l'objet n'est pas suffisante. Cependant, la combinaison avec le squelette et les contours peut renforcer la prise de la bonne décision.

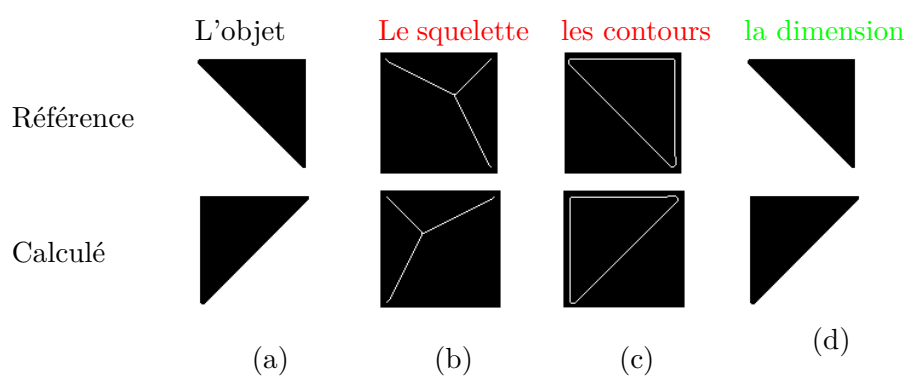


Figure 5.17 – Deux formes ayant la même dimension, mais elles n'ont pas le même squelette, ni les mêmes contours.

5.7 Système expert flou pour la prise de décision

Pour la prise de décision sur la conformité du montage d'usinage, nous proposons d'utiliser un système expert flou pour deux raisons principales.

- La première est que les caractéristiques descriptives de l'objet et les scores mesurés à partir de ces derniers peuvent être entachés par du bruit qui conduit à des scores relativement imprécis.
- Deuxièmement, compte tenu des imprécisions des scores, il nous est apparu impossible de hiérarchiser les caractéristiques pour en sortir un arbre de décision. La section 5.6 a montré les relations entre les caractéristiques de la forme.

Considérant ces deux points et l'**incertitude** ou l'imprécision de mesure portée sur les différentes étapes du traitement, partant de calcul des contours, chapitre 3, jusqu'à la localisation des régions d'intérêts, chapitre 4 et calcul de la forme de l'objet, chapitre 5, la nécessité d'une étape de quantification dans laquelle le raisonnement des systèmes intelligents ou experts pourrait remplacer l'intervention humaine, est une étape importante.

Dans des situations similaires d'incertitude, le problème est traité souvent par des méthodes **probabilistes** ou **possibilistes**. Souvent, les méthodes probabilistes cherchent à modéliser le comportement interne du système en utilisant des relations mathématiques entre ces variables aléatoires. Cependant, dans certains cas, la source de l'incertitude ne vient pas uniquement des variables internes du système, mais aussi d'autres paramètres extérieurs affectant la décision. L'intégration de la connaissance humaine dans la quantification des scores est une étape importante pour l'automatisation du processus [Keller 1999],[Li 1989] et [Pao 1989]. Nous nous intéressons dans la suite à la définition de quelques notions associées à la logique floue et nous montrons la méthode proposée pour la prise de décision.

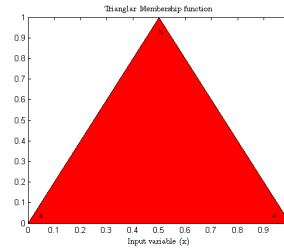
5.7.1 Définition des ensembles flous

La logique floue, fondée par Lotfi ZEDEH [Zedeh 1965], est une technique possibiliste caractérisée par un ensemble de règles, *Membership Functions (MF)* définies par un expert. Chaque règle floue est caractérisée par sa forme qui traduit un syllogisme exprimé par un expert, mais en relativisant la valeur de décision en la remplaçant par un intervalle pondéré par une valeur de la certitude définissant un ensemble flou. Il existe plusieurs formes géométriques modélisant les règles floues, triangulaires, trapézoïdales, Gaussiennes, sinusoïdales et sigmoïdes, etc. En général 3 formes sont très utilisées. Nous les présentons ici : Triangulaire, Trapézoïdale et Gaussienne.

5.7.1.1 Ensemble triangulaire

Un triangle peut être défini par trois paramètres (a, b, c) , avec $(a < b < c)$ comme suit :

$$Triangle(x; a, b, c) = \begin{cases} 0, & \text{si } x \leq a \\ \frac{x - a}{b - a}, & \text{si } a \leq x \leq b \\ \frac{c - x}{c - b}, & \text{si } b \leq x \leq c \\ 0, & \text{si } c \leq x \end{cases} \quad (5.19)$$



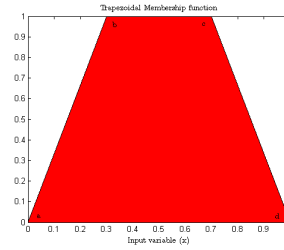
En utilisant les fonctions min et max, nous obtenons une expression de la variable x dans le triangle (a, b, c) :

$$Triangle(x; a, b, c) = \max \left(\min \left(\frac{x - a}{b - a}, \frac{c - x}{c - b} \right), 0 \right) \quad (5.20)$$

5.7.1.2 Ensemble trapézoïdale

Un trapézoïde peut être défini par quatre paramètres (a, b, c, d) , avec $(a < b < c, d)$ comme suit :

$$\text{Trapézoïde}(x; a, b, c, d) = \begin{cases} 0, & \text{si } x \leq a \\ \frac{x-a}{b-a}, & \text{si } a \leq x \leq b \\ 1, & \text{si } b \leq x \leq c \\ \frac{d-x}{d-c}, & \text{si } c \leq x \leq d \\ 0, & \text{si } d \leq x \end{cases} \quad (5.21)$$



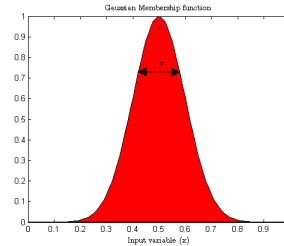
En utilisant les fonctions min et max, nous obtenons une expression de la variable x dans le trapézoïde (a, b, c, d) :

$$\text{Trapézoïde}(x; a, b, c, d) = \max \left(\min \left(\frac{x-a}{b-a}, \frac{d-x}{d-c} \right), 0 \right) \quad (5.22)$$

5.7.1.3 Ensemble Gaussienne

La règle Gaussienne peut être modélisée par deux paramètres (μ, σ) .

$$\text{Gaussien}(x; \mu, \sigma) = \exp \frac{-1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \quad (5.23)$$



Les valeurs de μ et σ , représentent la moyenne et l'écart-type de la fonction Gaussienne.

Les choix d'une forme de l'ensemble flou dépendent de la complexité du problème et des plages de variation des variables. Les deux formes, triangulaire et trapézoïdale, ont une caractéristique discontinue par rapport à la forme Gaussienne. La forme triangulaire a tendance à trancher par son maximum marqué par le point (b) , ce qui donne deux parties (a, b) et (b, c) . La forme trapézoïdale a tendance à répartir son maximum sur l'intervalle (b, c) , ce qui donne trois parties (a, b) , (b, c) et (c, d) . Cela peut être interprété par un étalement de l'incertitude pour des valeurs proches de "1". L'ensemble Gaussien caractérisé par sa forme continue permet de regrouper les propriétés des deux ensembles, triangulaire et trapézoïdale et ainsi de donner une forme continue aux petites valeurs d'incertitude. Cette caractéristique est conditionnée par le choix des valeurs μ et σ . Nous adoptons pour notre problématique la forme Gaussienne pour définir les variables d'entrée du système flou, et la forme triangulaire pour définir la variable de sortie du système flou.

5.7.2 Les principales opérations floues

Contrairement à la logique classique, Booléenne, les opérations sur les ensembles flous permettent de décrire l'interaction entre ces ensembles en présence d'une variable x . Cette interaction dépend de la forme des règles floues choisie auparavant dans l'étape de la définition des ensembles flous. Il existe plusieurs types d'opérations. En général, les plus souvent utilisées sont : **le complément**, **l'intersection** et **l'union**.

De plus, dans la logique floue, on adopte le mode de raisonnement approximatif plutôt que exact. Le raisonnement humain est approximatif et le système expert flou tend à le modéliser. Dans certains cas la variable x peut appartenir à deux intervalles différents avec des taux d'appartenance différents, pour 2 parties du discours de l'expert.

A fuzzy set is a class of objects with a continuum of grades of membership. Such a set is characterized by a membership (characteristic) function which assigns to each object a grade of membership ranging between zero and one. Lotfi Zedeh[Zedeh 1965]

Cette approximation du langage linguistique est exprimée en langage flou par le **le degré d'appartenance** ou **le degré de vérité**. En effet, la logique floue associe à chaque valeur de l'entrée x un degré d'appartenance ou de vérité entre 0 et 1 dans chaque classe définissant l'ensemble flou de la variable en question. Un exemple est donné sur la figure 5.18. Dans cet exemple, la variable floue d'entrée x appartient à l'ensemble **Moyen** avec un degré d'appartenance égal à 0.2 et à l'ensemble **Grand**, avec un degré d'appartenance égal à 0.6. Il est à noter que le degré d'appartenance dépend aussi de la courbure de l'ensemble flou. Cette caractéristique est l'une des propriétés de la logique floue qui la distingue à la logique "classique", binaire. Cette propriété permet de donner plus de précision à des définitions ambiguës, comme par exemple, *faible, peu, chaud, petit, etc.*

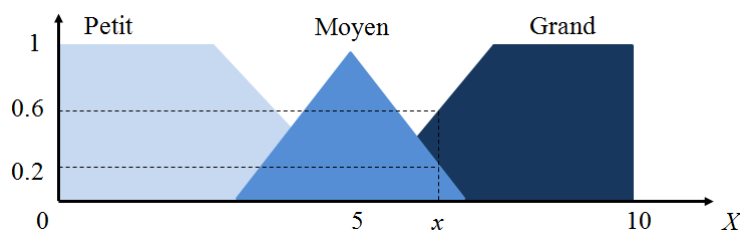


Figure 5.18 – Définition du degré d'appartenance.

- **L'intersection floue**, notée \cap entre deux ensembles flous consiste à déterminer la valeur minimale d'appartenance de la variable x aux deux ensembles (A) et (B).

$$(A \cap B)(x) = \min(A(x), B(x)) \quad (5.24)$$

- **Le complément flou** permet de vérifier le degré de non appartenance d'un variable x à l'ensemble (A).

$$cA(x) = 1 - A(x) \quad (5.25)$$

- **L'union floue**, notée \cup est l'opération inverse de l'opération intersection. L'union floue cherche à déterminer le degré d'appartenance de la variable x à l'un des ensembles (A) et (B) en utilisant la valeur maximum d'appartenance.

$$(A \cap B)(x) = \max(A(x), B(x)) \quad (5.26)$$

Géométriquement, la figure 5.19 illustre les trois opérations.

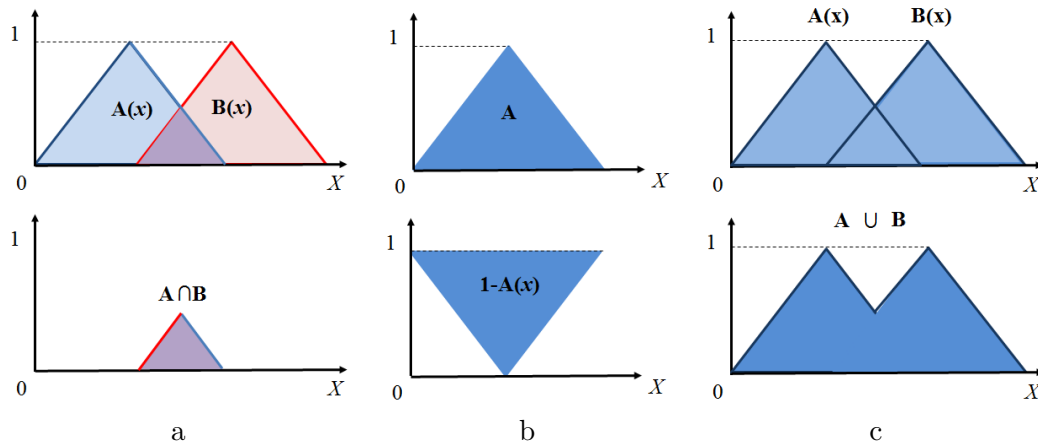


Figure 5.19 – Exemple d'opérations sur les ensembles flous : (a) intersection floue, (b) complément flou et (c) union floue

5.7.3 Fuzzification

La fuzzification est l'opération qui consiste à transformer l'entrée numérique x en une donnée linguistique. Elle s'appuie sur le concept :

Si condition Alors décision.

Les variables linguistiques sont données par les ensembles flous issus du discours de l'expert. Chaque ensemble possède généralement trois classes, i.e. petit, moyen, grand ; lent, moyen, rapide ; froid, tiède, chaud, etc. Il faut noter qu'il n'y a pas une règle pour choisir le nombre de classes, ni la forme des ensembles flous, ni l'intervalle de définition des classes pour chaque règle floue. Tous ces paramètres s'appuient sur l'expérience humaine.

La section 5.7.3.1, présente la définition de nos règles floues et de l'intervalle de ces ensembles.

5.7.3.1 Définition des règles floues

Chaque ensemble flou d'entrée représente un des trois paramètres descriptifs de l'objet : score d'appariement de squelettes, score d'évaluation de contours (CE) et score de dimension.

Nous proposons de définir trois ensembles pour chaque règle floue d'entrée, et deux ensembles pour la règle floue de sortie, e.g. "Décision". Nous rappelons que pour l'instant nous ne répondons pas à la situation *Orange*, e.g. "montage réalisé non conforme, mais ne provoquant pas de collision". Des éléments de solution pour cette situation sont traités dans le chapitre 6, par l'estimation de la carte de profondeur. L'ensemble des règles floues définies sont illustrées sur la figure 5.20. Les figures 5.20.b, 5.20.c et 5.20.d représentent les règles d'entrées : score d'appariement de squelettes, score d'évaluation de contours et score de dimension successivement. Nous proposons de prendre la forme Gaussienne pour définir les règles floues d'entrées. L'intervalle de chaque variable est divisé en trois parties, la première partie appelée "Petite" est caractérisée par une moyenne centrée à zéro et un écart-type égale à "0.12", ce qui permet d'étaler la gaussienne sur des petites valeurs proche de zéro et permet aussi d'étaler les valeurs proches de "1" sur l'intervalle de l'écart-type, i.e. 0.12. L'ensemble appelée "Moyen" est centrée sur 0.5 avec le même écart-type, 0.12, et enfin l'ensemble "Grand" centré sur la valeur "1" avec le même écart-type, 0.12. Cet écart-type a été choisi de façon à obtenir une équi-répartition des ensembles flous.

La figure 5.20.e présente la règle de sortie donnée par la "Décision". La forme choisie pour définir les ensembles flous est triangulaire, ce qui va permettre de trancher la décision. Nous proposons aussi de donner plus d'importance à l'ensemble "Non conforme" défini sur un intervalle $[0, 0.70]$ et par conséquence l'ensemble "Conforme" est défini sur le reste de l'intervalle $]0.70, 1]$. Cette définition intègre en quelque sorte une mesure de "prudence" permettant de garantir la bonne décision, car **"il vaut mieux décider que ce n'est pas bon quand c'est bon, plutôt que de décider que c'est bon quand ce n'est pas bon"**. Autrement dit, pour la sécurité de la machine, il vaut mieux être prudent dans des décisions "Vrai négatifs" que dans des décisions "Faux positifs".

Les relations entre les règles floues sont illustrées sur la figure 5.21

En trois dimensions la variable linguistique "Décision" peut être représentée comme le montre la figure 5.22.

5.7.4 Défuzzification et décision

Le processus de défuzzification consiste à transformer la variable linguistique de sortie en variable numérique. Dans une première étape, il faut fusionner les variables linguistiques à l'aide des opérateurs de la logique floue et des règles définies auparavant. Pour nous, l'opération logique choisie est le "ET". Par exemple, appelons S_{Skel} le score d'appariement du squelettes, S_{CE} le score de l'évaluation des contours et S_{Dim} le score de dimension, dans ce cas, une des règles floues que nous avons définies est donnée par :

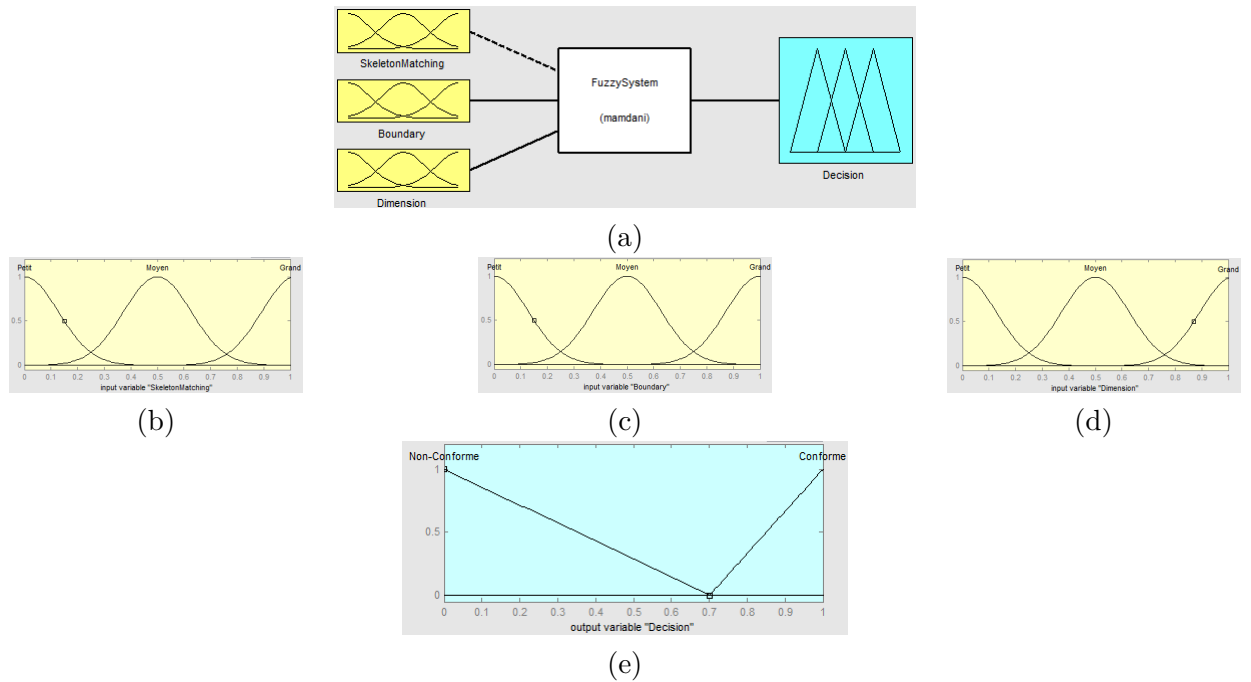


Figure 5.20 – Définition des règles floues, (a) architecture général du système expert flou, (b) règle de score d'appariement de squelettes, (c) règle de score CE, (d) règle de score de dimension et (e) règle de la décision

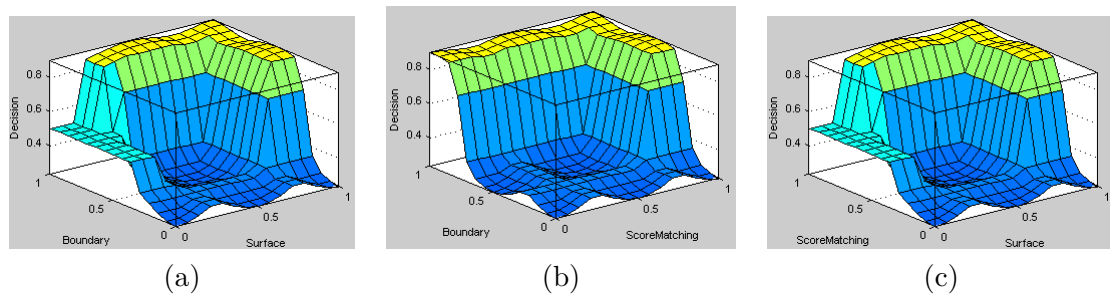


Figure 5.21 – Relation 2D entre les règles floues, (a) relation 2D entre l'ensemble flou **CE** et l'ensemble flou **Dimension**, (b) relation 2D entre l'ensemble flou **CE** et l'ensemble flou **Score d'appariement de squelettes**, (c) relation 2D entre l'ensemble flou **Dimension** et l'ensemble flou **Score d'appariement de squelettes**

Si S_{Skel} est Grand ET Si S_{CE} est Grand ET Si S_{Dim} est Grand Alors L'objet est conforme.

La seconde étape, consiste à trouver la valeur quantitative en fonction du degré d'appartenance des variables linguistiques. Il existe plusieurs techniques de défuzzification, parmi lesquelles le début de la zone de la valeur maximale (**LOM**), la fin de la zone de la valeur maximale (**SOM**), la moyenne de la zone de la valeur maximale (**MOM**), le centre de gravité (**Centroid**) du polygone et la méthode des moyennes

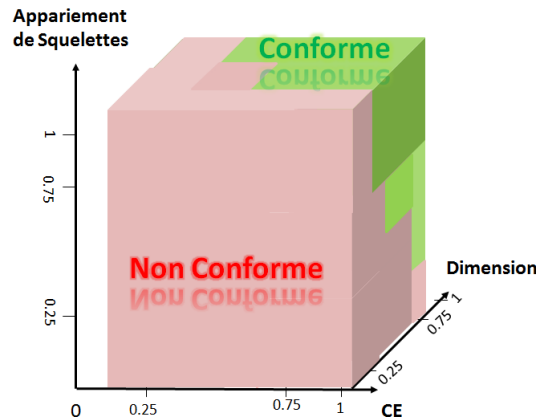


Figure 5.22 – Relation 3D entre les variables d'entrées et la variable de sortie floues

pondérées, etc. La figure 5.23 montre le positionnement de chacune des valeurs selon la méthode choisie.

Il existe deux méthodes de défuzzification connues en logique floue. **La défuzzification de Mamdani** et **la défuzzification de Sugeno**. Les deux méthodes se distinguent dans le calcul de l'**inférence** et l'agrégation des sorties de règles.

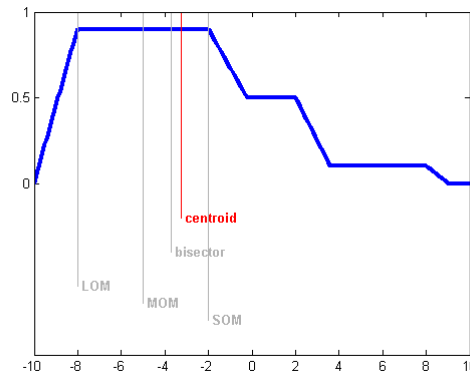


Figure 5.23 – Relation entre les méthodes de fuzzification et la forme des règles floues. Exemple de www.mathworks.com

- La méthode d'inférence proposée par Mamdani, consiste à calculer le polygone englobant les valeurs d'appartenance de la variable d'entrée. La hauteur du polygone est définie par la plus grande valeur d'appartenance. L'agrégation de cette variable est calculée par l'une des techniques de défuzzification présentées ultérieurement, i.e. MOM, LOM, SOM, Centre de gravité et la méthode des moyennes pondérées. On utilise ici la méthode du **centre de gravité** donnée par l'équation (5.27).

$$CG_{Mamdani} = \frac{\sum_{x=a}^b \mu_A(x) \cdot x}{\sum_{x=a}^b \mu_A(x)} \quad (5.27)$$

- La méthode d'inférence proposée par Sugeno s'appuie sur la modélisation de l'ensemble flou en un **pic**. L'amplitude du pic est donnée par la valeur d'appartenance de la variable x à l'ensemble flou en question. Le résultat de l'agrégation revient à utiliser la méthode des moyennes pondérées, donnée par l'équation (5.28).

$$CG_{Sugeno} = \frac{\sum_i \mu_A(k_i) \cdot k_i}{\sum_i \mu_A(k_i)} \quad (5.28)$$

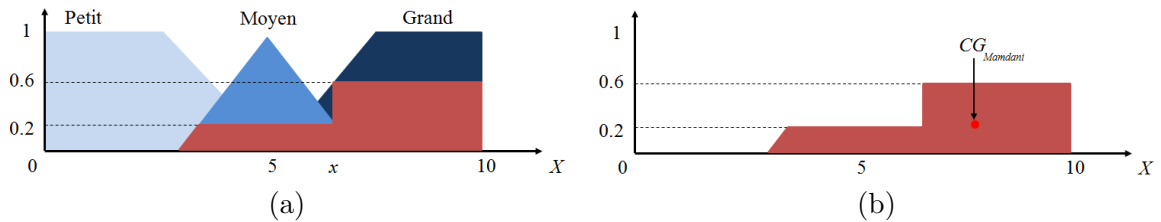


Figure 5.24 – Principe de l'inférence Mamdani et agrégation des règles de sortie, (a) calcul de taux d'appartenance, (c) inférence et calcul de la valeur de la sortie.

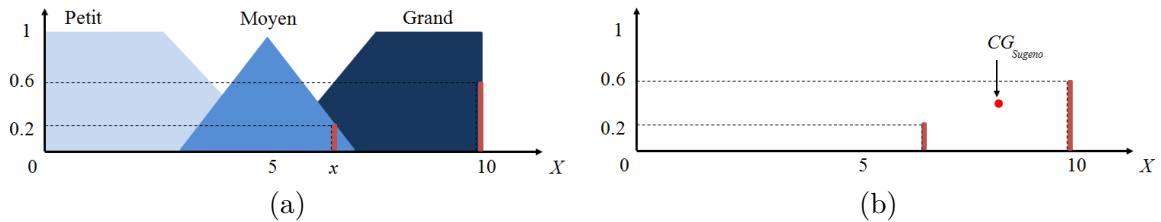


Figure 5.25 – Inférence Sugeno et agrégation des règles de sortie, (a) calcul de taux d'appartenance, (c) inférence et calcul de la valeur de la sortie.

Nous avons testé les deux méthodes de défuzzification, données par Mamdani et Sugeno. Nous n'avons pas remarqué de différence dans la prise de décision, car la forme des ensembles flous définissant la variable de sortie "Décision" sont séparés et ne présentent pas une zone d'intersection, zone de confusion. En plus elles sont centrées chacune, sur "0" pour l'ensemble "non conforme" et "1" pour l'ensemble "Conforme", ce qui implique que le résultat obtenu de la décision est souvent supérieur à 0.8 pour le cas conforme et inférieur à 0.2 pour le cas "non conforme".

Les différentes étapes de l'algorithme et la prise de décision pour le cas conforme du montage Minimanche et le cas non conforme du montage sur la plaque Norelem, e.g. bride en plus placée sur le brut, sont illustrés sur les figures suivantes². D'autres exemples sont présentés dans l'annexe A. La première colonne représente l'image de luminance de la région de l'objet à analyser, la seconde colonne l'image de l'objet calculée

2. Les pages des figures sont en format A3

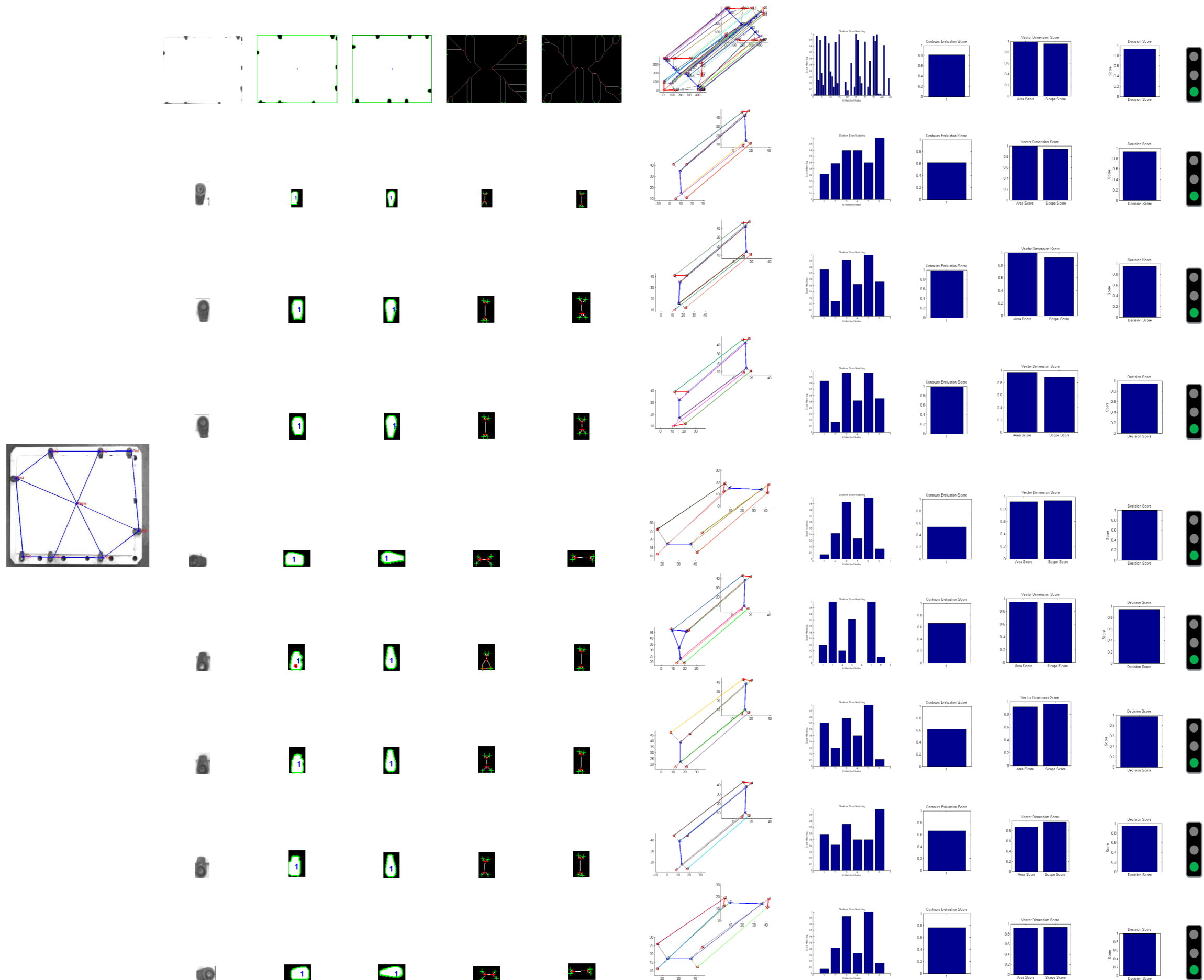
par notre méthode, e.g. Otsu et opérations morphologies pour l'objet réel, la troisième colonne l'image théorique de l'objet de référence, la quatrième et la cinquième colonne les squelettes (G) et (H) calculés sur l'image binaire réel et théorique, successivement, la sixième colonne l'appariement de squelettes (G) et (H), la septième colonne le score d'appariement de chaque nœud dans le squelette de référence, la huitième colonne de score (CE), la neuvième colonne le score de la dimension, la dixième colonne le score de la Décision et enfin la décision finale.

La troisième figure illustre un cas où l'élément de bridage n'a pas la bonne dimension, e.g. bride de fixation. Le système a pu détecter l'anomalie, car le paramètre caractérisant le score d'évaluation des contours est mesuré à une valeur "petite" dans la règle floue (CE). Par conséquent, la variable "Décision" est évaluée à une valeur autour de 0.6, e.g. "non-conforme", et inférieure au seuil minimal pour l'ensemble "conforme", donné par 0.7. Cet exemple montre, entre autres, l'importance de la combinaison des paramètres descriptifs pour la prise de décision. Il présente aussi l'intérêt de l'utilisation du système expert flou et la robustesse des règles définies. D'autres exemples de résultats pour différentes configurations des montages d'usinage, e.g. Norelem, Semelle et Minimanche, sont illustrés sur l'annexe A.

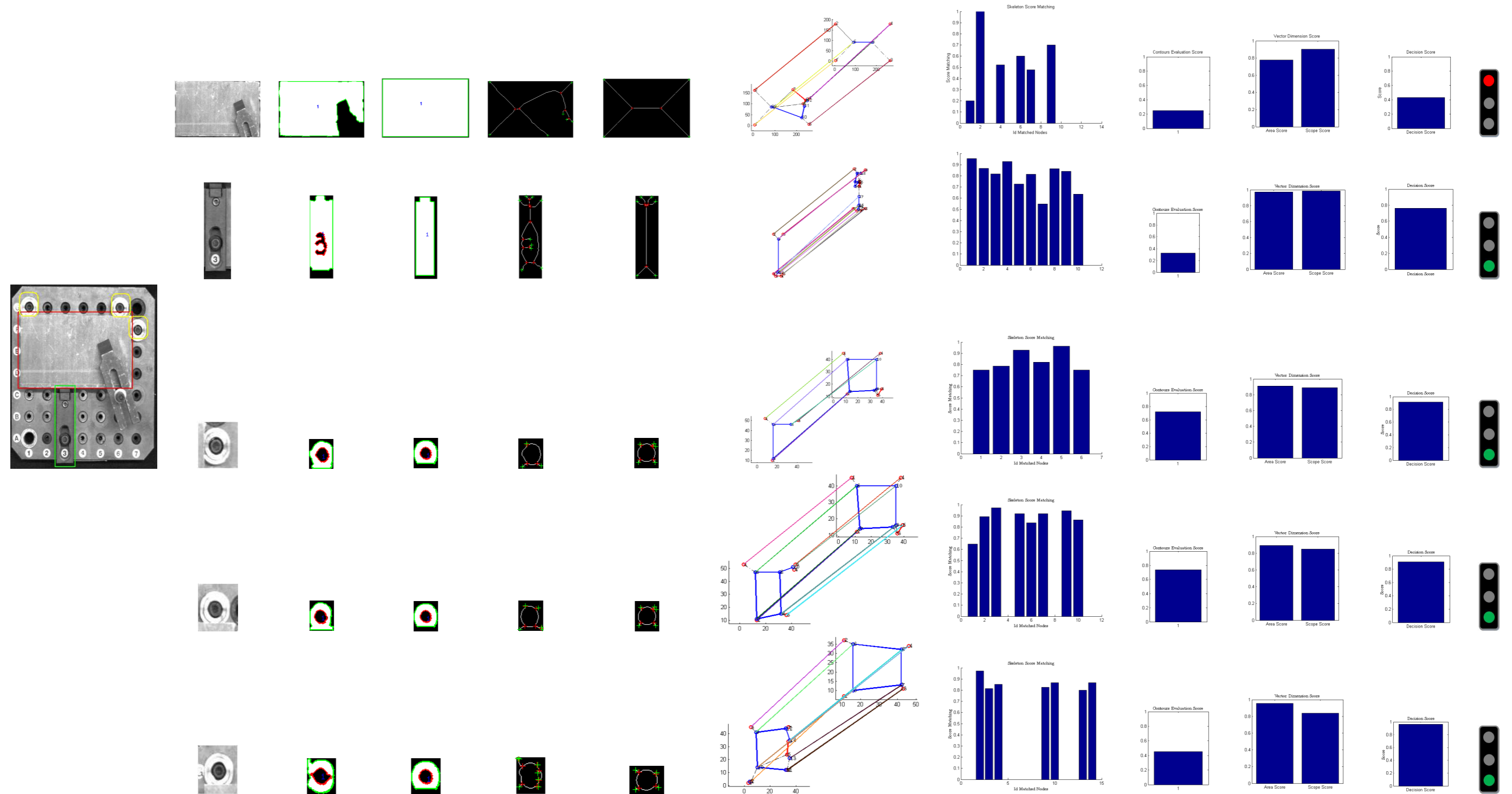
Cependant, le pion décalé n'est pas localisé! Car la définition CAO ne donne pas un pion à cet endroit. Nous nous sommes interrogés sur l'analyse de la plaque pour répondre à ce cas de figure. Si un objet est ajouté sur la plaque en dehors des régions d'intérêt, comment l'extraire à partir du reste³ de l'image 2D de luminance? Avons-nous besoin de l'information 3D? La section suivante 5.8 présente les résultats obtenus et le chapitre 6 apporte des éléments de solution pour répondre à ces questions.

3. les régions des éléments localisés sont remis au fond.

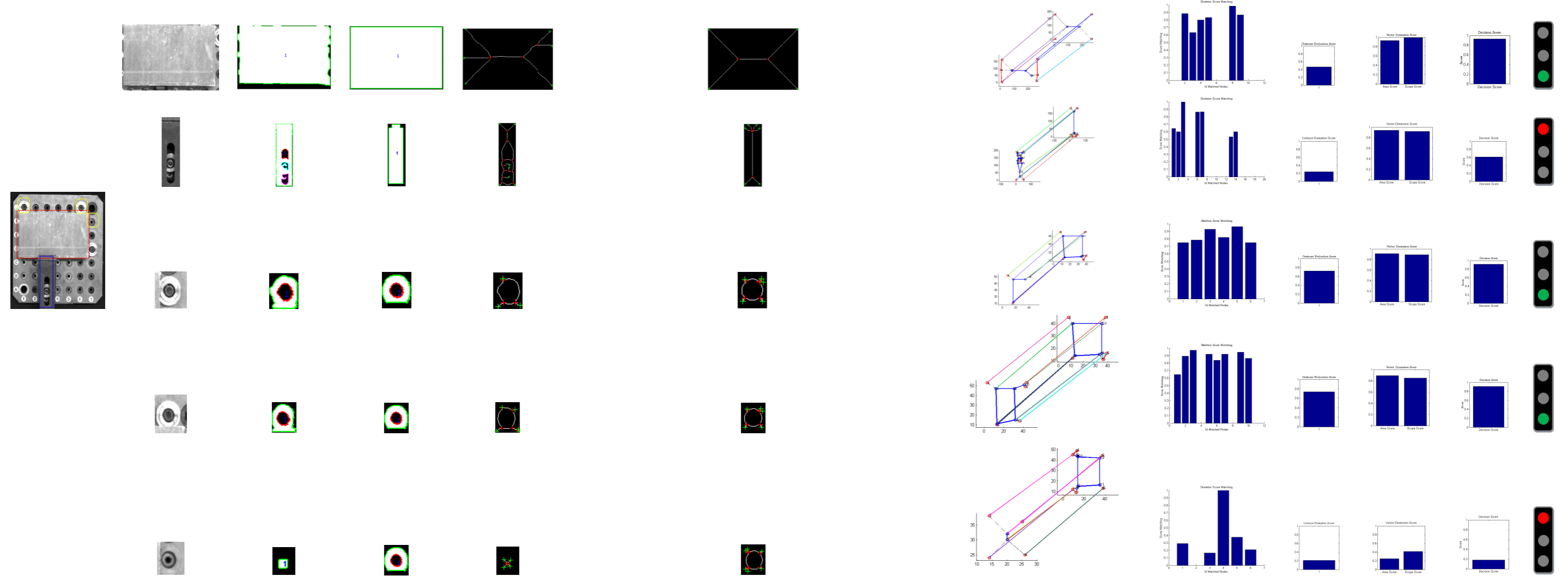
Montage Minimanche : Configuration conforme



Montage Norelem : Configuration Bride en plus



Montage Norelem : Configuration Bride qui n'a pas la bonne dimension et Pion en moins



5.8 Avons-nous besoin de l'information tridimensionnelle ?

Nous avons jusqu'à maintenant pu répondre à la problématique posée sur la vérification automatique des montages d'usinage par vision. Les différentes formes des pièces et les différentes configurations de non conformité du montage d'usinage que nous avons pu imaginer, nous ont conduit à reposer la question de la nécessité de l'information tridimensionnelle notamment pour répondre au cas "Orange" de la chaîne numérique de traitement de l'usinage. Cette réponse doit inclure l'information 3D issue de la scène pour la comparer à la trajectoire de l'outil afin de répondre à la situation d'un montage d'usinage qui n'est pas conforme mais qui ne provoque pas de collision avec l'outil. Cette information 3D peut informer notre système du cas où l'élément de bridage plus épais et qui donne presque la même empreinte dans la résolution de l'image.

La figure 5.26 montre un autre exemple, dans lequel une bride a été ajoutée dans un endroit en dehors des régions d'intérêt. Les paramètres descriptifs que nous employons, en utilisant seulement la luminance, ne sont pas toujours suffisants pour distinguer une telle situation en 2D. Sur cet exemple, la solution 3D peut être observée aussi comme un élément confortant la segmentation.

Les éléments en dehors des régions définies par le modèle CAO ne sont pas vus et par conséquent, ne sont pas analysés. Il est important alors d'analyser la plaque qui fait partie des éléments constituant le montage d'usinage. Cependant la région de la plaque est caractérisée par une diversité de formes circulaires représentées par les trous et qui peuvent affecter la segmentation. En effet, nous avons testé les 3 algorithmes présentés précédemment. Le résultat est illustré sur la figure 5.26. Les régions des éléments de bridage définis par le modèle CAO sont remis au fond "noir". La figure 5.26.a présente les régions et la figure 5.26.b l'image de la plaque seule après remise des régions analysées au fond. Notre objectif est de séparer la bride de la plaque en utilisant une des méthodes de segmentation par région, présentées précédemment. La figure 5.26.c représente le résultat pour l'algorithme *Ncut*, la figure 5.26.d le résultat de l'algorithme Otsu plus l'application de la morphologie mathématique, la figure 5.26.e le résultat de l'algorithme *GMM*, et enfin, l'image de profondeur calculée à partir d'un nuage de points donnés par un système de numérisation 3D du chapitre 3. La séparation de deux régions par une segmentation de luminance n'est pas facile voire impossible. Le calcul de la carte de profondeur va faciliter la segmentation en marquant les différents plans, niveaux des éléments de la scène, et la séparation des éléments par couplage de l'information 3D avec l'information 2D rendra plus facile la segmentation par région [Richtsfeld 2012].

5.9 Conclusion

Nous avons vu dans ce chapitre les différentes étapes de traitement permettant une prise de décision automatique sur la conformité de l'élément de bridage en cours

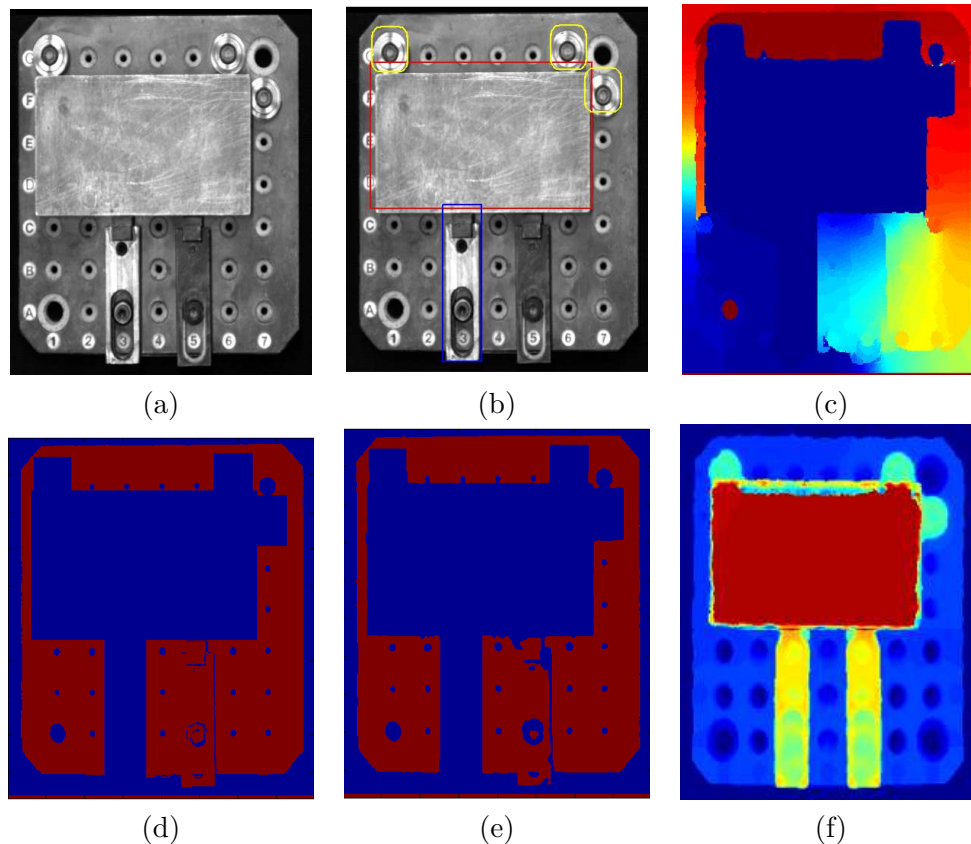


Figure 5.26 – Nécessité de l'information 3D, (a) Image du montage Norelem sans les éléments de bridage analysés, (b) localisation des régions d'intérêt, (c) Segmentation par la méthode *Ncut*, (d) Segmentation par la méthode proposée, (e) segmentation par la méthode *GMM* et (f) carte de profondeur de la scène.

d'analyse et par conséquent le montage d'usinage réalisé.

L'utilisation de l'image théorique binaire de l'objet permet de réduire la complexité du problème par une comparaison directe des deux images, au lieu de chercher l'image correspondante dans une base de donnée des formes d'objets.

La détermination de la forme de l'objet est un élément clé dans le calcul des paramètres descriptifs. Nous avons montré que les trois paramètres sont indispensables pour une prise de décision automatique en utilisant un système expert flou.

Le système expert flou s'appuie sur le raisonnement approximatif humain. Il intervient de façon prépondérante pour qualifier la reconnaissance de l'objet à partir de sa région et permet une prise de décision.

Cependant, le système expert flou ne peut réagir que selon les règles programmées. Le développement d'une mesure adaptative de l'incertitude pour la décision nécessite, parfois, le couplage de ce système avec des systèmes d'apprentissage automatique, e.g. réseaux de neurones, algorithmes génétiques, etc. , ce qui permet de mettre à niveau la base des règles floues afin d'adapter la décision à la situation courante par l'ajustement des intervalles. Cependant, pour notre application, cette situation est peu envisageable

compte tenu de la variété des formes et la géométrie des pièces produites.

Dans ce chapitre, nous n'avons pas encore répondu à la situation sur le risque de collision, cas **Orange** dans la chaîne du traitement proposée, car l'évaluation du risque de collision nécessite le couplage de l'information 3D pour la comparaison avec la trajectoire de l'outil. Le chapitre 6 présente le début d'une solution permettant d'estimer la carte de profondeur de la scène à partir d'une seule acquisition d'image couleur **RGB**.

Bibliographie

- [Arbelaez 2009] P. Arbelaez, M. Maire, C. C. Fowlkes et J. Malik. From Contours to Regions : An Empirical Evaluation. In Conf. Computer Vision and Pattern Recognition, 2009., 2009. (Cité en page 112.)
- [Arcelli 1985a] C. Arcelli, K.-M. Lam et W.-C. Siu. Extraction of the Euclidean skeleton based on a connectivity criterion. Pattern Recognition, vol. 36, pages 721–729, 1985. (Cité en page 123.)
- [Arcelli 1985b] C. Arcelli, G. Sanniti di Baja et R. Perucchio. A Width Independent Fast Thinning Algorithm. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 7, pages 463–474, 1985. (Cité en page 123.)
- [Arcelli 1989] C. Arcelli et G. S. Di Baja. A one-pass two-operation process to detect the skeletal pixels on the 4-distance transform. Mathematical Morphology and its Applications to Image and Signal Processing, vol. 11, no. 4, pages 411–414, 1989. (Cité en page 123.)
- [Bai 2007] X. Bai, L.J. Latecki et W.-Y. Liu. Skeleton Pruning by Contour Partitioning with Discrete Curve Evolution. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 29(3), pages 449–462, 2007. (Cité en page 124.)
- [Bai 2008] X. Bai et L.J. Latecki. Path Similarity Skeleton Graph Matching. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30(7), pages 1282–1292, 2008. (Cité en pages 124 et 125.)
- [Bouguila 2007] N. Bouguila, D. Ziou et R. I-Hammoud. A Bayesian Non-Gaussian Mixture Analysis : Application to Eye Modeling. Conf. Computer Vision and Pattern Recognition, 2007., pages 1–8, June 2007. (Cité en page 116.)
- [Brandt 1992] J. W. Brandt et V. R. Algazi. Continuous skeleton computation by Voronoi diagram. Computer Vision, Graphics, Image Processing, vol. 55, pages 329–338, 1992. (Cité en page 123.)
- [Breu 1995] H. Breu, J. Gil, D. Kirkpatrick et M. Werman. Linear time euclidean distance transform algorithms. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 17, pages 529–533, 1995. (Cité en page 123.)
- [Dimitrov 2000] P. Dimitrov, C. Phillips et K. Siddiqi. Robust and Efficient Skeletal Graphs. Conf. Computer Vision and Pattern Recognition, 2000. (Cité en page 124.)

- [Dimitrov 2003] P. Dimitrov, J. N. Damon et K. Siddiqi. Flux Invariants for Shape. Conf. Computer Vision and Pattern Recognition, 2003. (Cité en page 124.)
- [Forsyth 2002] D.A. Forsyth et J. Ponce. Computer Vision : A Modern Approach. 2002. (Cité en pages 110 et 117.)
- [Ge 1996] Y. Ge et J.M. Fitzpatrick. On the Generation of Skeletons from Discrete Euclidean Distance Maps. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18(11), pages 1055–1066, 1996. (Cité en page 123.)
- [Keller 1999] P. Keller J.M.and Matsakis. Aspects of high level computer vision using fuzzy sets. In Conf. Fuzzy Systems,IEEE, volume 2, pages 847 – 852, 1999. (Cité en page 131.)
- [Li 1989] H. Li et H.S. Yang. Fast and reliable image enhancement using fuzzy relaxation technique. IEEE Trans. Systems, Man and Cybernetics, vol. 19(5), pages 1276–1281, 1989. (Cité en page 131.)
- [Liu 2010] M-Y. Liu, O. Tuzel, A. Veeraraghavan et R. Chellappa. Fast directional chamfer matching. In Conf. Computer Vision and Pattern Recognition. IEEE Computer Society, 2010. (Cité en page 127.)
- [Meijster 2000] A. Meijster, J. Roerdink et H. Hesselink. A general algorithm for computing distance transforms in linear time. Mathematical Morphology and its Applications to Image and Signal Processing, pages 331–340, 2000. (Cité en page 123.)
- [Murphy 2012] K.-P. Murphy. Machine Learning : A Probabilistic Perspective. 2012. (Cité en page 117.)
- [Nacereddine 2009] N. Nacereddine, S. Tabbone, D. Ziou et L. Hamami. L’algorithme EM et le Modèle de Mélanges de Gaussiennes Généralisées pour la Segmentation d’images. Application au contrôle des joints soudés par radiographie. Conf. Traitement et Analyse de l’Information : Méthodes et Applications, pages 217–222, 2009. (Cité en pages 112 et 117.)
- [Ogniewicz 1995] R. L. Ogniewicz et O. Kùbler. Hierarchic Voronoi skeletons. Pattern Recognition, vol. 28(3), pages 343–359, 1995. (Cité en page 123.)
- [Pao 1989] Y.-H. Pao. Vague features and Vagues Decision Rules : The Fuzzy-Set Approach. 1989. (Cité en page 131.)
- [Pudney 1998] C Pudney. Distance-Ordered Homotopic Thinning : A Skeletonization Algorithm for 3D Digital Images. Computer Vision and Image Understanding, vol. 72, no. 3, pages 404–413, 1998. (Cité en page 123.)
- [Richtsfeld 2012] T. Richtsfeld A.and Morwald, J. Prankl, M. Zillich et M. Vincze. Segmentation of Unknown Objects in Indoor Environments. In IEEE RSJ International Conf. Intelligent Robots and Systems (IROS), 2012. (Cité en pages 143 et 169.)
- [Shaked 1998] D. Shaked et A. M. Bruckstein. Pruning Medial Axes. Computer Vision and Image Understanding, vol. 69(2), pages 156–169, 1998. (Cité en page 124.)

- [Shi 2000] J. Shi et M. Jitendra. Normalized Cuts and Image Segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22(8), pages 888–905, 2000. (Cité en pages [xii](#), [73](#), [111](#), [113](#) et [114](#).)
- [Vasilevskiy 2002] A. Vasilevskiy et K. Siddiqi. Flux Maximizing Geometric Flows. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24(12), pages 1565–1578, 2002. (Cité en page [124](#).)
- [Wu 1993] Z. Wu et R. Leahy. An Optimal Graph Theoretic Approach to Data Clustering : Theory and Its Application to Image Segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, no. 11, pages 1101–1113, 1993. (Cité en pages [70](#) et [114](#).)
- [Xie 2003] W. Xie, R.P. Thompson et R. Perucchio. A topology-preserving parallel 3D thinning algorithm for extracting the curve skeleton. Pattern Recognition, vol. 36, pages 1529–1544, 2003. (Cité en page [123](#).)
- [Zedeh 1965] L.A. Zedeh. Fuzzy Sets. Information and Control, vol. 8, pages 338–353, 1965. (Cité en pages [131](#) et [133](#).)

COMMENT OBTENIR L'INFORMATION 3D À PARTIR D'UNE SEULE CAMÉRA ?

"It is absolutely certain that there is a definite dynamical theory for waves and light, to be enriched not abolished by electromagnetic theory". William Thomson¹, (1824-1907).

Sommaire

6.1	Introduction et problématique	149
6.2	Formation de l'image	152
6.3	Relation entre le flou et l'éloignement	154
6.4	Méthodes de perception de la profondeur (DFD)	156
6.4.1	Méthode exploitant le flou sur les contours [Zhuo 2011]	157
6.4.1.1	Recherche de l'indice de flou	157
6.4.1.2	Interpolation de la carte des flous	158
6.4.2	Méthode adoptée [Deschenes 2003]	160
6.4.2.1	Utilisation des dérivées appliquée à la transformée S	160
6.4.2.2	Flou couleur : notre proposition	162
6.5	Expérimentation et résultats	163
6.6	La réflectance de l'image pour l'estimation de la carte de profondeur : <i>Shape From Shading</i>	167
6.7	Création de l'image RGBD	169
6.8	Conclusion	170
	Bibliographie	172

Mots clés : Depth From Defocus, Image couleur, Longueur d'onde.

6.1 Introduction et problématique

Le chapitre précédent a montré que, dans certains cas, une information 3D est nécessaire pour décider de la conformité du montage d'usinage. Le cas de l'analyse de la

1. Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light, 1904

plaque, e.g. support du montage d'usinage, sans l'exploitation de l'information tridimensionnelle est un vrai challenge en exploitant seulement la luminance, car la plaque est constituée de plusieurs trous de fixations qui pourraient perturber la segmentation et l'extraction de la forme d'un objet ajouté.

Les premiers essais présentés dans le chapitre 2 sur les performances des systèmes de numérisation 3D permettaient de conclure qu'il n'existe pas un dispositif répondant aux critères de rapidité d'acquisition, de coût et de faible encombrement.

Nous avons donc étudié la possibilité de créer un système original d'acquisition 3D capable de fournir cette information rapidement, à la cadence vidéo. Nous avons privilégié une solution monoculaire capable de fournir une **carte des profondeurs denses**, dans la **résolution de l'image**, pour éviter tout problème de fusion de données. La précision en profondeur n'est pas obligatoirement fine, car les déplacements de la broche en hauteur ne sont jamais programmés au ras des éléments du montage d'usinage. Une précision de l'ordre du millimètre semble suffisante selon les spécialistes. Le coût devra piloter les solutions matérielles que nous sélectionnerons.

Une solution monoculaire n'est possible que par l'acquisition de plusieurs images, d'un même point de vue, mais avec des variations des paramètres d'acquisition. Dans cette gamme de techniques, nous trouvons, celles proposant de faire varier la focale, celles avec un déplacement du plan de focalisation, *depth from focus (DFF)* [Trevor 1988], celles exploitant le flou optique par modification de l'ouverture [Zhou 2009] ou par un léger changement de mise au point [Deschenes 2003], méthodes appelées *depth from defocus (DFD)*, et celles utilisant une variation de l'éclairage [Daniel 2000], *shape from shading (SFS)*.

Nous avons éliminé les méthodes utilisant une variation de la focale, car les algorithmes d'appariement sont coûteux en temps de calcul. Nous avons aussi éliminé les méthodes de *DFF*, car elles nécessitent une multitude d'acquisitions, selon la précision souhaitée en profondeur. Une image par position du plan focal, image dans laquelle les zones de netteté sont recherchées.

Certaines méthodes basées sur l'ouverture, proposent de placer un code barre, e.g. filtre, devant l'iris pour contrôler la quantité de lumière traversant la caméra. Cette solution peut être observée comme élément d'ajustement de la quantité de lumière et peut être un indice pour l'estimation du flou optique.

Les travaux proposant un changement d'ouverture entre deux images acquises, nette et floue, s'appuient sur le principe du changement du diamètre de la tâche floue formée sur les pixels. D'autres travaux [Simon 2004], [Zhuo 2011], s'intéressent aux contours pour estimer le diamètre de la tâche donnant la profondeur. En effet, les contours dans l'image peuvent représenter une discontinuité de profondeur et toujours une discontinuité de luminance. L'intérêt de l'utilisation des contours pour l'estimation de l'indice de profondeur, rend ces méthodes plus générales. Cependant, elles ne donnent qu'une information de profondeur sur les contours. Dans [Zhuo 2011] les auteurs proposent une estimation d'une carte d'éparse en utilisant un filtrage bilatéral joint. La section 6.4.1 en présente le détail.

Nous avons concentré nos efforts sur l'évaluation de méthodes *DFD*, en proposant une contribution originale d'amélioration de méthodes existantes. Nous avons testé

notre solution avec la méthode de F. Deschenes *et al.*[Deschenes 2003]. Ce travail que nous évaluons dans les sections 6.4.2 et 6.5, a été réalisé dans le cadre d'une collaboration avec le centre de recherche MOIVRE² de l'université de Sherbrooke, Québec.

Notre contribution s'articule sur la proposition d'un nouveau principe pour le calcul de la *carte de profondeur* à partir d'une *seule acquisition d'image couleurs* et par conséquent, sans changement des paramètres d'acquisition.

Enfin nous avons confié un montage d'usinage à des spécialistes de SFS pour une évaluation de performances. Un exemple de résultat est présenté dans la section 6.6.

Nous nous intéressons dans la section 6.2 à rappeler le principe de la formation d'image et le phénomène optique du flou.

Objectif :

État de l'art

Notre contribution

Autres techniques

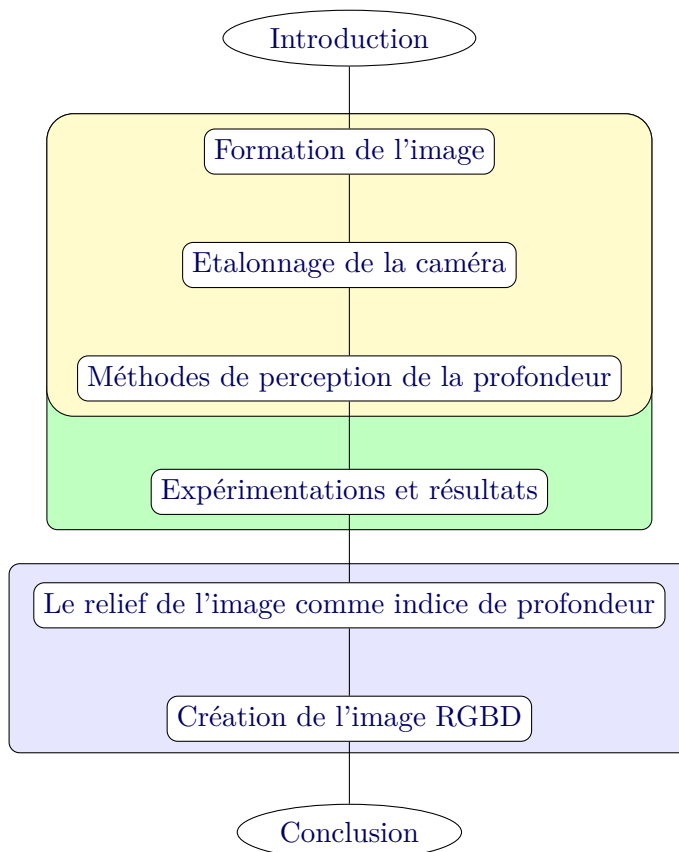


Figure 6.1 – Organigramme du chapitre 6.

6.2 Formation de l'image

Depuis les premiers travaux de Pentland P.A [Pentland 1987] et Subbarao M. et Gurumorthy N. [Subbarao 1988], la modélisation de la formation des images floues est formulée selon l'équation (6.1).

$$I_b(x, y, s_0) = I_s(x, y) \otimes_{2D} h(x, y, s_0) \quad (6.1)$$

Dans cette équation, la formation de l'image floue $I_b(x, y, s_0)$ est exprimée comme la convolution bidimensionnelle d'une image théorique $I_s(x, y)$ considérée nette en toute distance s_0 , par la réponse impulsionnelle du système optique $h(x, y, s_0)$, qui dépend de la distance s_0 entre le plan de la lentille et le plan objet. Depuis 1988 et jusqu'aux travaux de Favaro P. et Saotto S. [Paolo 2007] et plus récemment [Zhuo 2011], la réponse impulsionnelle est modélisée par une Gaussienne bidimensionnelle donnée par l'équation (6.2).

$$h(x, y, s_0) = g_{\sigma_0}(x, y) = \frac{1}{2\pi\sigma_0^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_0^2}\right) \quad (6.2)$$

Le principe de la formation des images nettes et floues est présenté respectivement, dans sa forme géométrique, sur la figure 6.2 avec les réponses impulsionnelles correspondantes.

Pour former une image nette, deux cas peuvent être retenus. Celui d'un plan objet parfaitement focalisé, figure 6.2.a, et celui des rayons paraxiaux, figure 6.2.b. e.g. hypothèse conduisant au modèle *sténopé*. Dans ce dernier cas la quantité de lumière arrivant sur le capteur est faible, conduisant à un rapport signal à bruit faible, le bruit de l'électronique dominant la conversion photon \curvearrowright électron. Une ouverture trop petite induira aussi des phénomènes de diffraction non négligeables compte tenu de la nature ondulatoire de la lumière. Pour former donc une image nette, il faut utiliser une ouverture plus grande pour de bonnes conditions d'acquisition. Cette condition détermine une approximation de netteté liée à la taille du pixel. En effet pour les plans objet non focalisés une tâche image se forme. Si cette tâche s'inscrit dans un pixel, l'image paraîtra nette, définissant une profondeur de champ, sinon le diamètre de la tâche dépassera la taille du pixel et l'image paraîtra floue. La figure 6.3 montre la taille de la zone de profondeur du champ en fonction de la taille du pixel et le l'ouverture. Plus l'ouverture est petite et plus la taille du pixel est grande, plus la zone de profondeur de champ est grande. Cette zone constitue pour nous une imprécision, puisque le flou n'y est pas perceptible.

Sur la figure 6.2, ξ représente le décalage de focalisation par rapport à la distance s_i focalisée, distance entre le plan lentille et le plan image. Le flou est aussi observé, si pour un même réglage de mise au point on déplace le plan objet de la quantité d .

La répartition de l'intensité lumineuse dans la tâche image n'est pas uniforme, même si certains auteurs utilisent cette approximation [Schneider 1994]. Une approximation plus proche de la réalité est donnée par le modèle Gaussien de la réponse

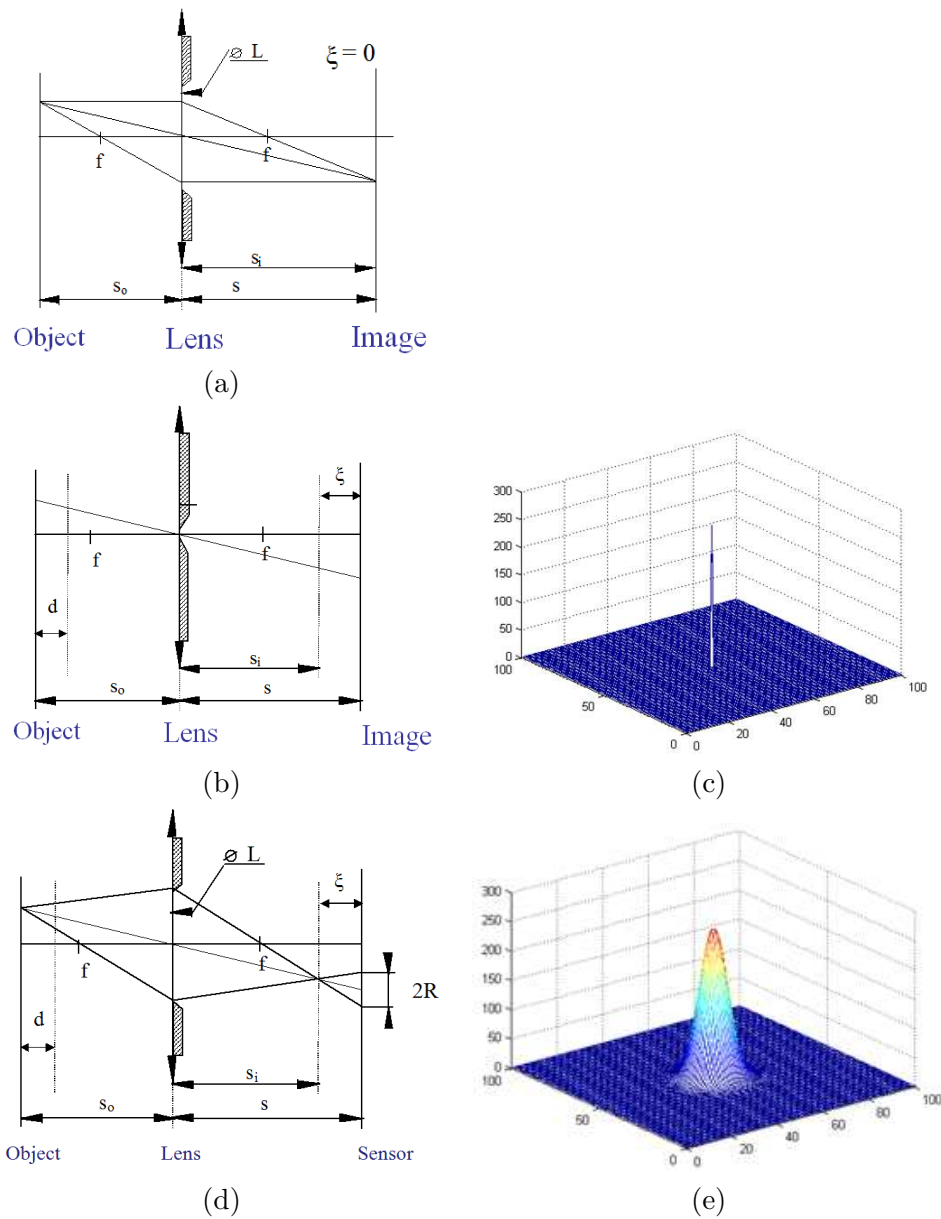


Figure 6.2 – Formation d'image : (a) image nette formée par les rayons Paraxiaux, (b) image nette focalisée, (c) réponse impulsionnelle correspondant au système donné en (b), (d) image formée par un ajustement de l'ouverture (L) et (e) réponse impulsionnelle du système donné en (d).

impulsionnelle. La forme théorique a été approximée par Stokseth [Stokseth 1969] à partir du modèle de Hopkins, elle est donnée par l'équation (6.3)³. La figure 6.4 illustre la forme de la réponse impulsionnelle correspondant au modèle Stokseth.

3. Day N. *et al.* ont utilisé le modèle Stokseth pour modéliser le flou dans [Dey 2002].

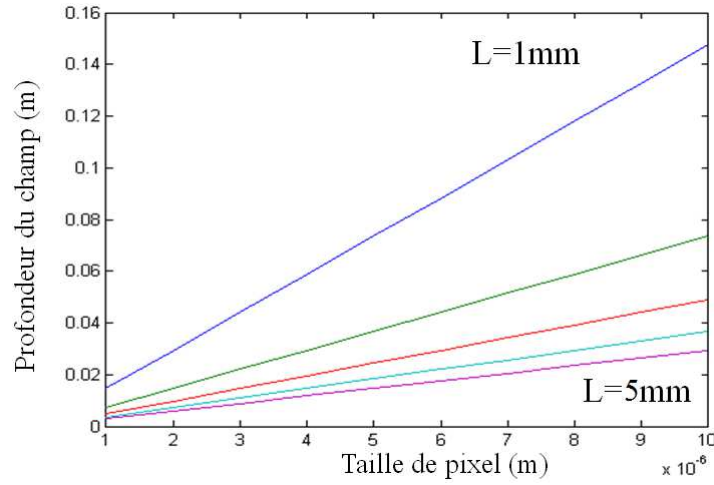


Figure 6.3 – Relation la taille du pixel et la profondeur du champs par rapport au flou optique.

$$h(x, y) = TF^{-1}\left\{\left(1 - 1.38\left(\frac{\mu}{\mu_c}\right) + 0.0304\left(\frac{\mu}{\mu_c}\right)^2 + 0.344\left(\frac{\mu}{\mu_c}\right)^3\right) Jinc\left(4kw\left(1 - \frac{\mu}{\mu_c}\right)\frac{\mu}{\mu_c}\right)\right\} \quad (6.3)$$

Où λ est la longueur d'onde dominante, $\mu = \sqrt{u^2 + v^2}$ la fréquence spatiale, $\mu_c = \frac{2L}{\lambda s_0}$, $k = \frac{2\pi}{\lambda}$, L l'ouverture, s_0 distance Objet/Caméra et $Jinc(x) = 2\frac{J_1(x)}{x}$ avec, J_1 la fonction de Bessel de première espèce et d'ordre 1.

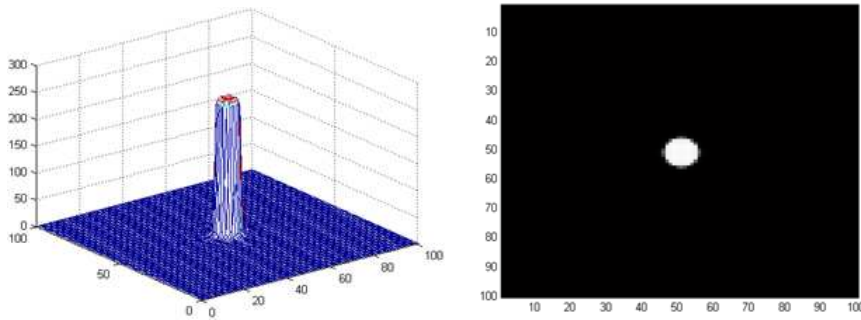


Figure 6.4 – Réponse impulsionnelle du système optique donnée par [Stokseth 1969], (a) le modèle Stokseth, (b) la tâche floue créée sur l'image.

6.3 Relation entre le flou et l'éloignement

Le flou peut être observé physiquement en changeant la mise au point du plan net de l'image au plan proche, e.g. $\xi < 0$ ou au plan loin, e.g. $\xi \geq 0$. Le flou est marqué sur les deux acquisitions d'image. La figure 6.5 en montre l'exemple.

La relation entre le flou, représenté par le rayon de la tâche R et la distance entre le plan objet et la caméra, s_0 , issue de la loi de conjugaison des lentilles minces, e.g.

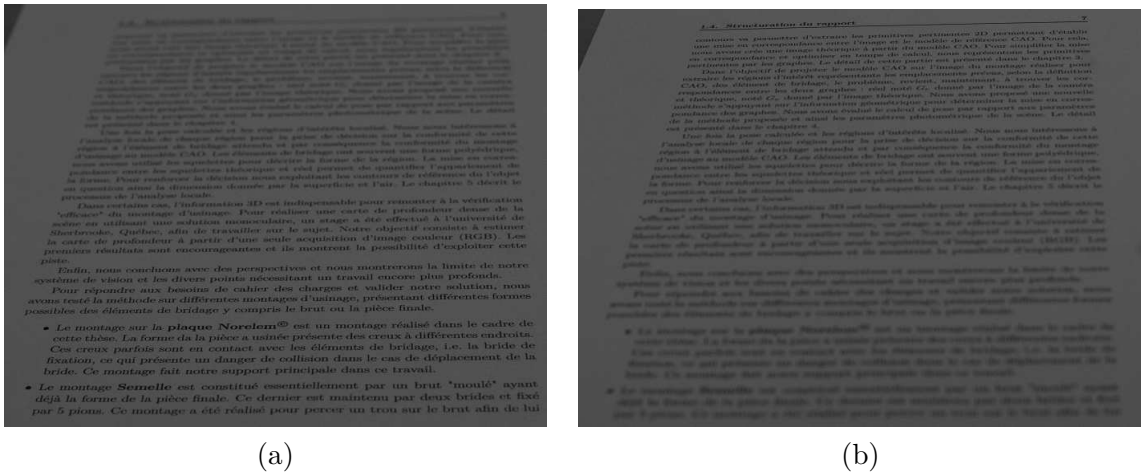


Figure 6.5 – Relation flou-éloignement, (a) cas de flou proche au plan image, e.g. $\xi < 0$, (b) cas de flou loin au plan image, e.g. $\xi \geq 0$

équation (6.4), est donnée par l'équation (6.5).

$$\frac{1}{f} = \frac{1}{s_0} + \frac{1}{s_i} \tag{6.4}$$

$$2R = Ls_i \left(\frac{1}{f} - \frac{1}{s_0} - \frac{1}{s_i} \right) \tag{6.5}$$

Plus exactement la relation dépend du signe de ξ . Le système d'équations suivant en précise la forme :

$$\begin{cases} s_0 = \frac{fs_i}{s_i - f - 2R\frac{f}{L}} & \text{Pour } \xi < 0 \\ s_0 = \frac{fs_i}{s_i - f + 2R\frac{f}{L}} & \text{Pour } \xi \geq 0 \end{cases} \tag{6.6}$$

La figure 6.6 présente la loi pour différentes valeurs de l'ouverture L .

Une estimation du diamètre équivalent à une distribution uniforme de la luminance pour une distribution Gaussienne permet, par la relation de proportionnalité établie avec l'écart type de la Gaussienne, de mettre en évidence deux paramètres globaux m , c de l'équation (6.7).

$$s_0^{-1} = m\sigma + c \tag{6.7}$$

Le coefficient de proportionnalité est combiné avec les paramètres du système optique, f focale, d_i distance plan image/plan lentille et L diamètre de l'ouverture. Ces paramètres m et c sont déterminés dans une phase d'étalonnage. Cette phase consiste à présenter dans le champ de vision de la caméra un plan noir et blanc pour marquer la discontinuité lumineuse, placés à différentes distances, permettant une mesure du flou

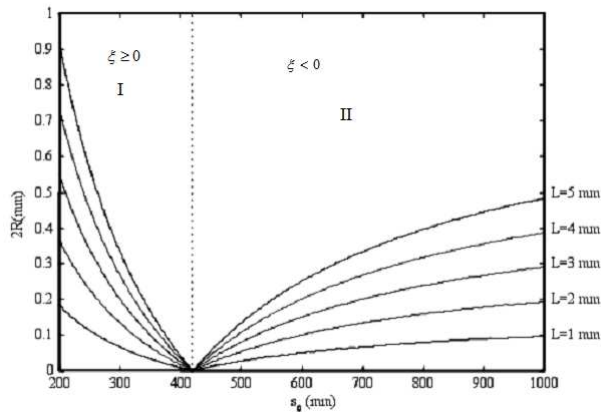


Figure 6.6 – Relation flou-éloignement par rapport à la caméra.

sur les contours par un ajustement de Gaussienne. Une courbe est construite passant par les points de coordonnées (flou, distance), puis les paramètres sont estimés par régression linéaire.

6.4 Méthodes de perception de la profondeur (DFD)

Le flou optique n'est perceptible que sur les discontinuités de luminance. Les figures 6.7.a et 6.7.b montrent respectivement un exemple sur des contours et sur des textures. La figure 6.7.a présente un plan incliné s'éloignant de la caméra portant une page d'écriture, l'évolution continue du flou y est perceptible. La figure 6.7.b est composée de trois plans, à différentes profondeurs et avec différentes textures.

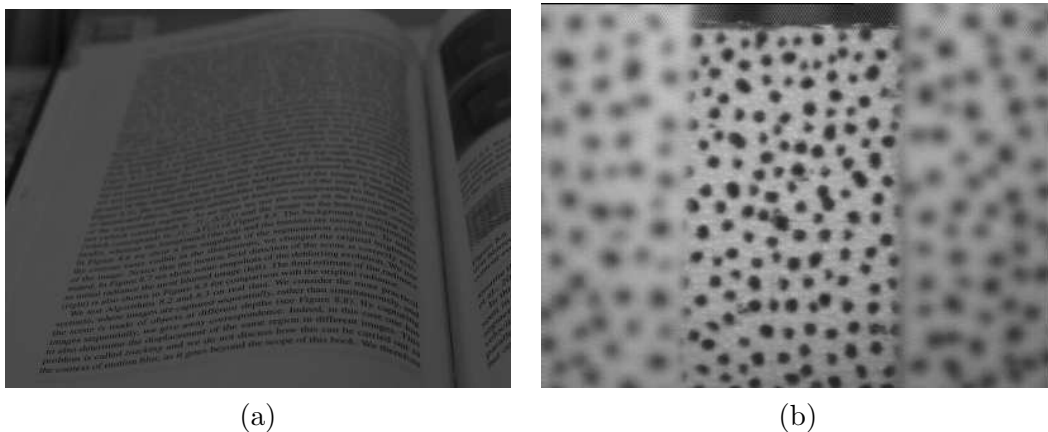


Figure 6.7 – Perception du flou optique sur les images, (a) contours flous sur le texte, (b) texture floue.

Nous proposons de préciser deux méthodes récentes utilisant les deux types d'information, puis nous présenterons la méthode que nous avons adaptée.

6.4.1 Méthode exploitant le flou sur les contours [Zhuo 2011]

La méthode exploitant le flou sur les contours est composée de deux grandes étapes. La première est la recherche d'un indice de flou, la seconde est une densification de la carte éparse des profondeurs obtenue à l'issue de la première étape. Cette méthode a été proposée par Zhuo, S. et Sim, T. [Zhuo 2011].

6.4.1.1 Recherche de l'indice de flou

L'indice de flou à rechercher est relié à la profondeur par l'équation (6.8). Il est pour les auteurs proportionnel à l'écart-type de la gaussienne modélisant la réponse impulsionnelle du système optique et évaluable uniquement sur les contours. Les contours $c(x)$ sont modélisés à partir d'une fonction de Heaviside $u(x)$ localisée en x_0 d'amplitude a et d'offset b dans un voisinage réduit permettant une définition du profil local de luminance. La définition monodimensionnelle est à étendre dans le voisinage selon la direction y .

$$c(x) = a.u(x - x_0) + b \quad (6.8)$$

La convolution 2D de l'image nette théorique par la réponse impulsionnelle du système optique peut être résumée à une transformation du profil du contour par l'équation (6.9) suivante :

$$i_b(x) = c(x) \otimes h(x, \dots, s_0, \sigma) \quad (6.9)$$

Les auteurs proposent de convoluer l'image floue dont les contours flous sont modélisés par l'équation (6.9), ci-dessus, par un noyau gaussien d'écart type s_0 . La transformation sur les contours peut être exprimée par l'expression (6.10).

$$i_{b_f}(x) = i_b(x) \otimes g(x, \sigma) \quad (6.10)$$

Cette convolution va permettre de filtrer le bruit et de mieux localiser les contours.

Le calcul du gradient des images floue i_b et floue filtrée i_{b_f} va conduire à l'indice de flou recherché. Le module du gradient de l'image i_b est donné par l'équation (6.11) et le module du gradient de l'image i_{b_f} est donné par l'équation (6.12).

$$|\nabla(i_b(x))| = |\nabla(c(x) \otimes h(x, \dots, s_0, \sigma))| = |\nabla(c(x)) \otimes h(x, \dots, s_0, \sigma)| \quad (6.11)$$

$$|\nabla(i_{b_f}(x))| = |\nabla(i_b(x) \otimes h(x, \dots, s_0, \sigma))| = |\nabla(c(x)) \otimes h(x, \dots, s_0, \sigma) \otimes g(x, \sigma)| \quad (6.12)$$

Le rapport des modules des gradients vaut alors :

$$R(x) = \frac{|\nabla(i_b(x))|}{|\nabla(i_{b_f}(x))|} = \sqrt{\frac{\sigma^2 + \sigma_0^2}{\sigma^2}} \exp \left[- \left(\frac{(x - x_0)^2}{2\sigma^2} \right) - \left(\frac{(x - x_0)^2}{2(\sigma^2 + \sigma_0^2)} \right) \right] \quad (6.13)$$

Sa valeur sur la discontinuité x_0 peut être exprimée par l'équation (6.14).

$$R(x_0) = \frac{|\nabla(i_b(x))|}{|\nabla(i_{b_1}(x))|} = \sqrt{\frac{\sigma^2 + \sigma_0^2}{\sigma^2}} \quad (6.14)$$

A cet endroit, l'indice de flou σ est obtenu par le maximum du rapport, selon l'équation (6.15) :

$$\sigma = \frac{1}{\sqrt{R(x_0)^2 - 1}} \sigma_0 \quad (6.15)$$

La figure 6.8 suivante illustre les résultats des opérations successives de la méthode.

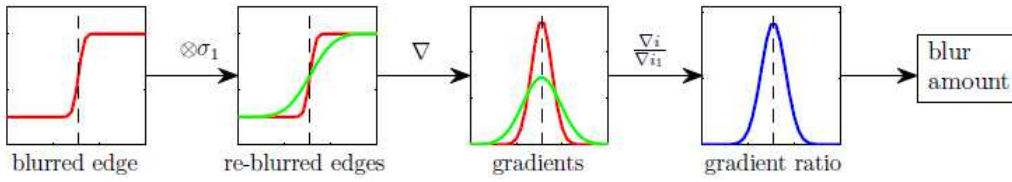


Figure 6.8 – Estimation de l'indice de profondeur sur les contours [Zhuo 2011].

Les modules des gradients et la position des contours peuvent être déterminées par le détecteur de Canny. Le résultat est une valeur de flou uniquement sur les contours entachée d'erreurs dues au bruit et à la configuration d'acquisition pour laquelle des ombres portées pourraient donner du flou lié à la diffraction. L'image obtenue est une carte éparsée des profondeurs. Pour améliorer la carte, un filtrage bilatéral joint (**JBF**) est appliqué selon l'expression (6.16) suivante.

$$BF(d(\hat{x})) = \frac{1}{W(x)} \sum_{y \in V_x} G_{\sigma_s}(\|x - y\|) G_{\sigma_r}(\|i_b(x) - i_b(y)\|) d(\hat{y}) \quad (6.16)$$

Avec $W(x)$ un facteur de normalisation, V_x le voisinage défini par le filtre gaussien G_{σ_s} sachant que σ_s , écart type de la gaussienne, et σ_r , écart type de la gaussienne G_{σ_r} . L'influence des gradients est fixées empiriquement et respectivement à 10% de la taille de l'image et 10% de l'échelle d'intensité et de la carte des flous.

6.4.1.2 Interpolation de la carte des flous

Une fois la carte des flous ou des profondeurs obtenue, la densification est réalisée par une technique de nattage par un opérateur laplacien. La technique peut ainsi être reformulée par la minimisation d'une fonction coût de la forme donnée par l'équation (6.17).

$$E(d) = d^T L d + \lambda (d - \hat{d})^T D (d - \hat{d}) \quad (6.17)$$

Avec \hat{d} et d respectivement des vecteurs formés à partir des matrices éparées $\hat{d}(x)$ et dense $d(x)$ que l'on cherche à obtenir. L est la matrice de nattage laplacien et D est une matrice diagonale dont les éléments sont à 1 quand le pixel est situé en un point de contour et 0 sinon.

L'élément (i, j) de la matrice L est défini par l'équation (6.18) :

$$L(i, j) = \sum_{k|(i, j) \in \omega_k} \left[\delta_{ij} - \frac{1}{|\omega_k|} \left(1 + (I_i - \mu_k)^T \left(\Sigma_k + \frac{\varepsilon}{|\omega_k|} U_k \right)^{-1} (I_j - \mu_k) \right) \right] \quad (6.18)$$

Avec δ_{ij} symbole de Kronecker, U_3 la matrice identité de taille 3, μ_k et Σ_k la moyenne et la matrice de covariance des chrominances sur la fenêtre ω_k , I_i les chrominances de l'image acquise au point i , ε un paramètre de régularisation pour éviter l'inversion d'une matrice nulle et $|\omega_k|$ la taille de la fenêtre ω_k .

La solution est obtenue par la résolution du système d'équations exprimé par l'expression (6.19).

$$\hat{d}(x)(L + \lambda D) = \lambda D \hat{d} \quad (6.19)$$

La figure 6.9 suivante illustre les résultats de l'application de la méthode dans laquelle plus le niveau de gris est élevé plus la distance à la caméra est grande.

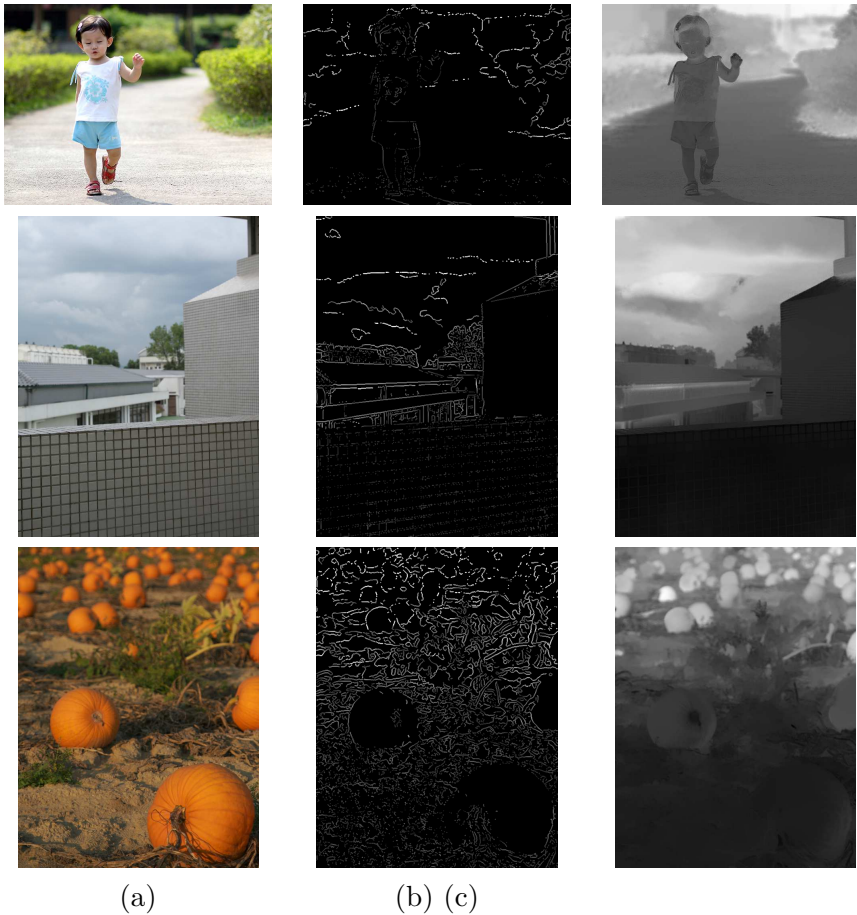


Figure 6.9 – Résultat de l'estimation de la carte de profondeur sur des images réelles donnée par la méthode [Zhuo 2011], (a) images couleurs, (b), matrice éparsse, (c) carte de profondeur

Les premiers essais que nous avons réalisés avec cette méthode montrent que l'ajustement des différents paramètres joue sur l'obtention d'un résultat et de sa validité. Un exemple est donné sur la figure 6.10. La carte de profondeur obtenue n'est pas exploitable. L'ajustement des paramètres dépend fortement des conditions photométriques de la scène et du résultat d'obtention des contours.

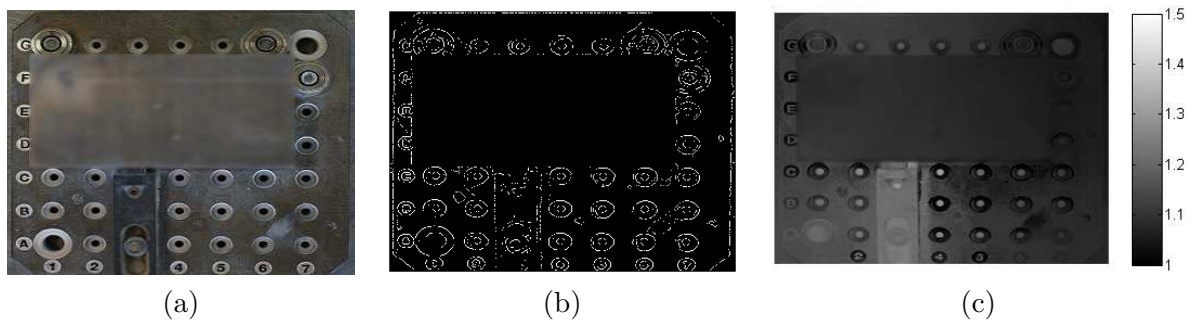


Figure 6.10 – Résultat de l'estimation de la carte de profondeur sur des images du montage d'usinage par la méthode [Zhuo 2011], (a) images du montage d'usinage, (b), matrice sparse, (c) carte de profondeur estimée.

6.4.2 Méthode adoptée [Deschenes 2003]

Nous nous intéressons dans cette section à une méthode n'exploitant pas directement les points de contours pour estimer la profondeur. La méthode que nous avons adoptée pour intégrer notre solution s'appuie sur l'estimation d'une carte dense de profondeurs à partir des dérivées des images, e.g. floue et nette, en utilisant les dérivées appliqués à la transformée S sur des textures [Subbarao 1994], plus difficilement sur des motifs.

Dans un premier temps, nous présentons les bases théoriques permettant la compréhension du principe de l'utilisation des dérivées appliquées à la transformée S . Nous expliquons aussi les détails liés au comportement des différentes équations.

6.4.2.1 Utilisation des dérivées appliquée à la transformée S

Basée sur le principe des moments, *Generalized Expansion Moment (GEM)*, les auteurs proposent d'exprimer l'équation (6.1) du principe de flou⁴ par une approximation en utilisant les dérivées partielles, e.g. $I^{(i)(j)}(x, y) = \frac{\partial^{i+j} I(x, y)}{\partial x^i \partial y^j}$, sur un domaine (P, Q) de l'image I . Cette relation est donnée par l'équation (6.20)

4. Le principe de flou pour les méthodes (*DFD*) s'appuie sur l'expression d'une image floue par une convolution 2D d'une image focalisée avec une gaussienne.

$$\begin{aligned}
I_b^{(n+p)(m+q)}(x_b, y_b) &= I_s^{(n)(m)}(x_b, y_b) + \\
\sum_{p=0}^P \sum_{q=0, p+q \geq 1}^Q & [s_p(\frac{\delta_x}{2}, \beta) s_q(\frac{\delta_y}{2}, \beta) I_s^{(n+p)(m+q)}(x_b, y_b) - \\
& s_p(-\frac{\delta_x}{2}, 0) s_q(-\frac{\delta_y}{2}, 0) I_b^{(n+p)(m+q)}(x_b, y_b)]
\end{aligned} \tag{6.20}$$

Où les coefficients $s_p(\sigma_x, \beta)$ ($s_q(\sigma_y, \beta)$) donnés par l'équation (6.21) sont proportionnels au moment d'ordre $p^{\text{ième}}$ de la réponse impulsionnelle, donnée par une gaussienne h_g centrée sur δ_x et δ_y respectivement. L'indice de flou est donné par la variable β . δ_x et δ_y représentent les disparités en x et y respectivement. Cette dernière provient du changement de la mise au point, ou à la modification de l'ouverture, entre les deux acquisitions d'images, I_b et I_s .

$$s_p(\delta_x, \beta) = \begin{cases} \frac{1}{p!} \sum_{i=0}^r C_{2i}^r \delta_x^{2i} \mu_{2r-2i}(\beta) & \text{si } \mathbf{p=2r} \\ \frac{1}{p!} \sum_{i=0}^r C_{2i+1}^r \delta_x^{2i+1} \mu_{2r-2i}(\beta) & \text{si } \mathbf{p=2r+1} \\ 1 & \text{si } \mathbf{p=0} \end{cases} \tag{6.21}$$

Avec,

$$u_b(\beta) = \begin{cases} (p-1)(p-3) \dots 3\beta^p & \text{si } \mathbf{p>1} \\ 1 & \text{si } \mathbf{p=0} \\ 0 & \text{ailleurs} \end{cases} \tag{6.22}$$

C_n^r le nombre des combinaisons possibles.

Le problème maintenant revient à déterminer les valeurs du triplet $(\beta, \alpha_x$ et $\alpha_y)$ en résolvant un système d'équations polynomiales multi-variables d'ordre P et Q . L'utilisation des dérivées des images consiste à supposer que ces dernières peuvent également être approchées par un polynôme cubique, e.g. de la même manière que la transformée S.

En posant $\beta_{(n)(m)}^2 = \sigma_b^2 - \sigma_s^2$, différence entre les écart-types des gaussiennes estimées sur le domaine (P, Q) à partir des dérivées $^{(n)(m)}$. β peut être exprimé par l'équation (6.23).

$$\beta^2(x, y) = 2 \frac{\sum_p \sum_q (I_b - I_s)^2}{\sum_p \sum_q (\nabla^2 I_f)} \tag{6.23}$$

Avec I_f donnée par la relation (6.24)

$$\nabla^2 I_f = \frac{\nabla^2 I_b + \nabla^2 I_s}{2} \tag{6.24}$$

Pour les trois premiers couples de dérivées $(n, m) = (0, 0), (0, 1), (1, 0)$ on obtient trois formes de l'équation (6.20). Il est désormais nécessaire de déterminer laquelle des équations est la plus pertinente à retenir en chaque voisinage (P, Q) centré sur (x, y) .

Les auteurs proposent de calculer l'erreur quadratique, donnée par la relation (6.25), et de conserver le $\beta_{(n)(m)}$ fournissant l'erreur la plus faible.

$$e_{(n),(m)}^2(x, y) = \sum_P \sum_Q (I_b(x + p, y + q) - (I_s \otimes \beta_{(n)(m)})(x + p, y + q))^2 \quad (6.25)$$

La recherche de la disparité $\delta_x = x + p$ et $\delta_y = y + q$ pour le calcul de l'erreur quadratique et la fusion des données peut alourdir l'algorithme. Notre contribution proposée dans la section suivante permet aussi de garantir un temps de calcul plus rapide en éliminant la disparité entre les deux images couleurs, car nous n'avons pas de réglage préalable à réaliser, e.g. les deux images ont été acquises avec les mêmes paramètres de la caméra.

6.4.2.2 Flou couleur : notre proposition

Plusieurs méthodes utilisant le flou optique comme source de l'information de profondeur à partir d'un seul point de vue existent. Les techniques appelées *Depth From Defocus*, exploitant le degré de flou perceptible sur les zones non homogènes de l'image correspondant à des contours ou à des textures, utilisent au moins deux images d'une même scène acquises avec des réglages différents des paramètres du système optique. Pour les méthodes que nous exploitons il s'agit de l'ouverture.

Le modèle de Stokseth et la relation (6.26) modélisant la réponse impulsionnelle du système optique met très clairement en évidence la dépendance de la forme de la réponse à la longueur d'onde. Pour un capteur photométrique couleur, il faudrait considérer la relation de l'équation (6.26) suivante pour modéliser la réponse impulsionnelle.

$$H_g(\mu) = \int_{\lambda_1}^{\lambda_2} S(\lambda)H(\mu)d\lambda \quad (6.26)$$

$S(\lambda)$ est la sensibilité spectrale du capteur et λ_1 et λ_2 les longueurs d'ondes limites de conversion photon-électron. Cette nouvelle relation conforte l'approximation gaussienne, la sommation assurant un lissage pondéré du model de Stokseth. Dans le cas où nous disposons d'un système d'acquisition d'image couleurs, les bornes des longueurs d'ondes dépendent du plan couleur choisi. La figure 6.11.a montre le principe physique de notre proposition de solution, et la figure 6.11.b. illustre l'efficacité quantique (QE) de notre caméra, permettant de caractériser la sensibilité électronique de la caméra par rapport à la longueur d'onde. La sensibilité du capteur aux trois longueurs d'ondes principales, *bleu*, *vert* et *rouge* est une gaussienne.

Pour un seul réglage, e.g. une seule acquisition d'image couleurs, nous avons trois modèles de la réponse impulsionnelle, un par canal de couleur, ce qui implique, trois images floues, une image par canal couleur.

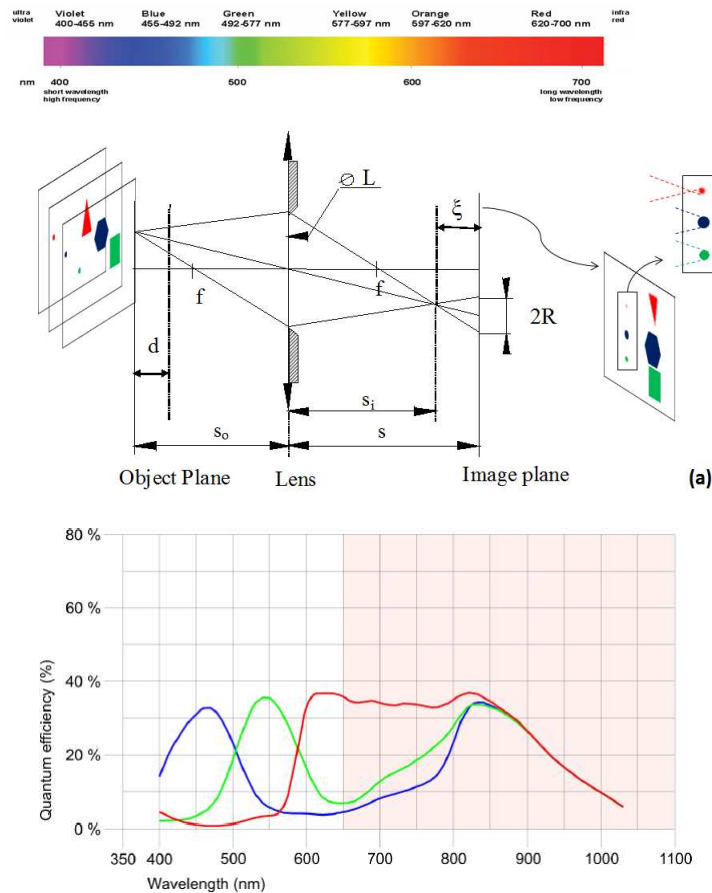


Figure 6.11 – Relation flou-éloignement par rapport à la caméra.

6.5 Expérimentation et résultats

Afin d'établir la relation entre les trois modèles de la réponse impulsionnelle et l'éloignement, e.g. profondeur, nous avons réalisé une expérimentation pour l'étalonnage de la caméra.

Nous proposons de placer un objet, noir et blanc, pour marquer les contours à différentes profondeurs du champ. La figure 6.12 présente les images de l'objet à différentes profondeurs.

L'objectif revient à estimer les écarts-types, σ_i , avec $i \in R, G, B$ de la gaussienne, sur les trois images couleur. L'estimation de paramètre σ_i est réalisée sur une fenêtre de taille (200×200) centrée sur le contour. Le résultat est donné sur la figure 6.13 pour l'image du canal "Rouge", l'image du canal "Vert" et l'image du canal "Bleu".

L'allure des trois gaussiennes est presque constante. La figure 6.14 montre le rapport presque constant des écarts-types α_C . La constance du rapport implique que nous pouvons utiliser chaque plan couleur comme une image avec un flou différent dans les méthodes déjà existantes.

Nous adoptons le modèle proposé par F. Deschênes *et al.* [Deschenes 2003], qui

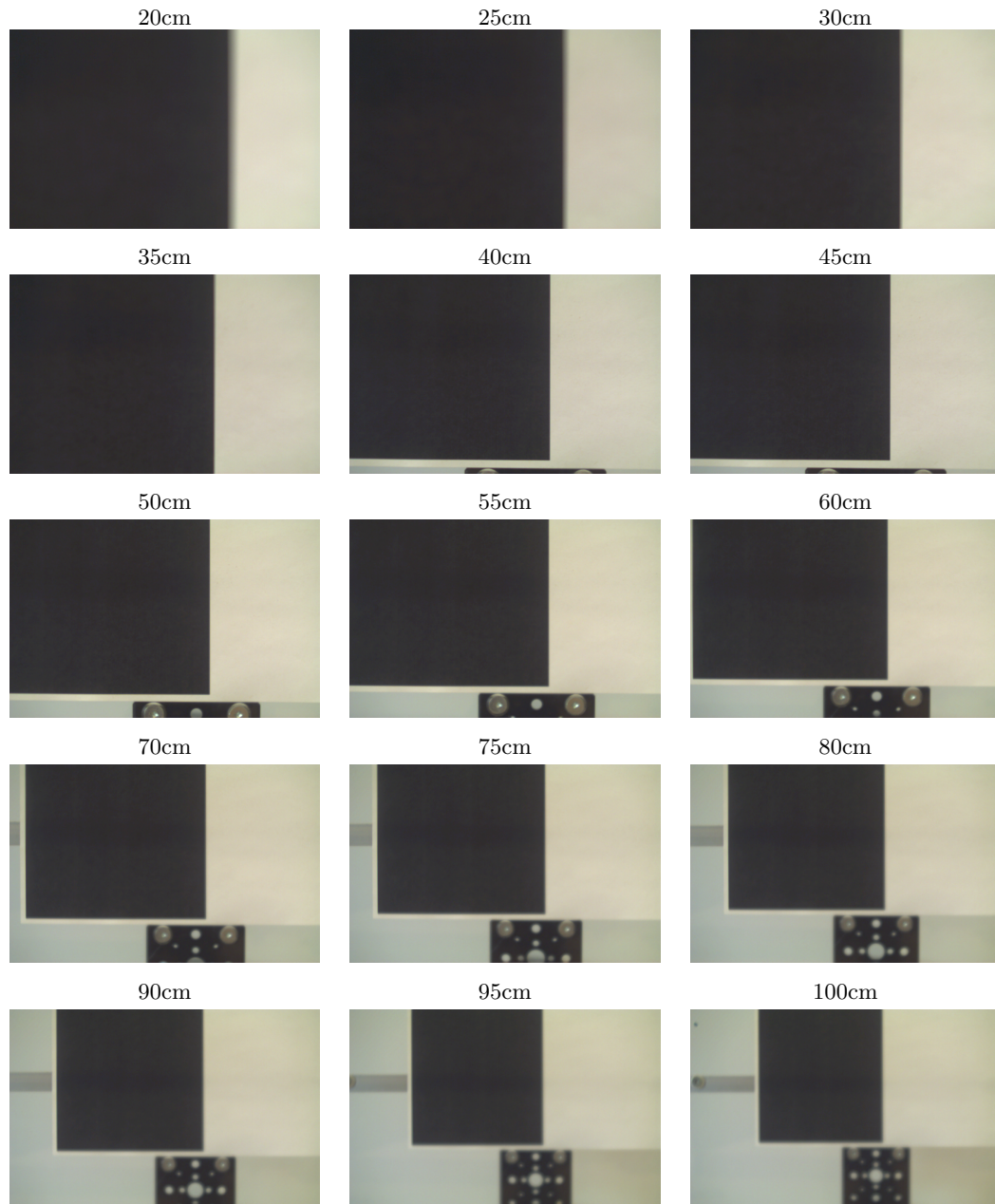


Figure 6.12 – Calibration de la caméra pour la détermination de la relation entre le flou et l'éloignement.

permet d'exprimer la relation entre une image nette I_s et une image floue I_b à partir de l'équation principale de convolution, e.g. équation (6.1).

Pour nous la valeur de β n'est que la différence entre le flou de l'image rouge, estimé par σ_R et le flou de l'image bleue, estimé par σ_B , selon l'équation (6.27) suivante.

$$\beta^2 = \sigma_R^2 - \sigma_B^2, \text{ pour } \sigma_R^2 \geq \sigma_B^2 \quad (6.27)$$

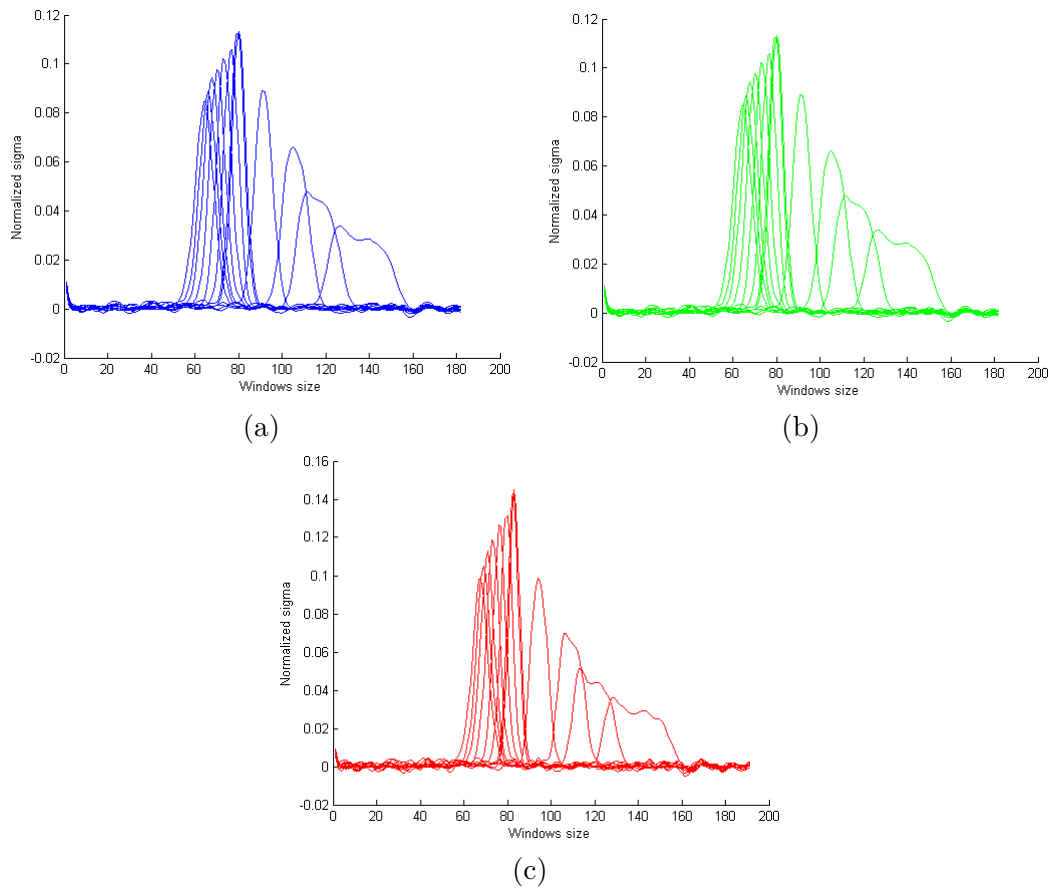


Figure 6.13 – Détermination des valeurs de σ_i , avec $i \in R, G, B$ (a) σ_R , (b) σ_G , (c) σ_B

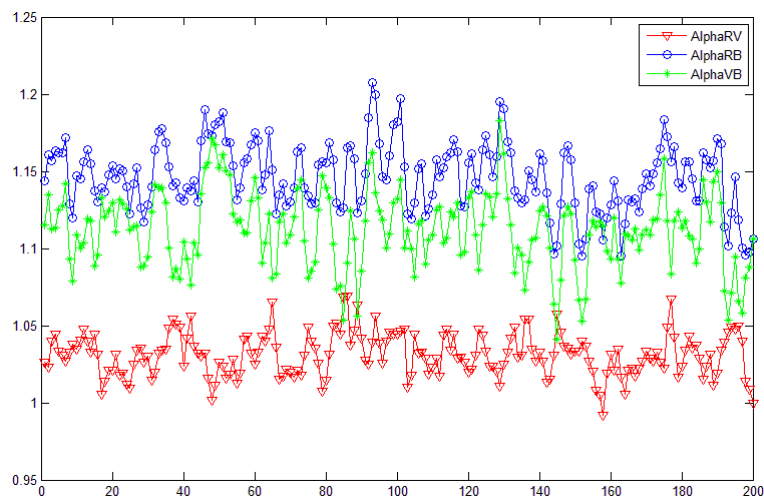


Figure 6.14 – Détermination des valeurs du rapport des α_{RG} , α_{RB} et α_{GB}

La figure 6.15 donne le résultat de la modélisation de la relation flou-éloignement selon le canal couleur, la position du plan net y est précisée. Cette étude montre ainsi

la forme quasi constante des réponses impulsionnelles selon le canal.

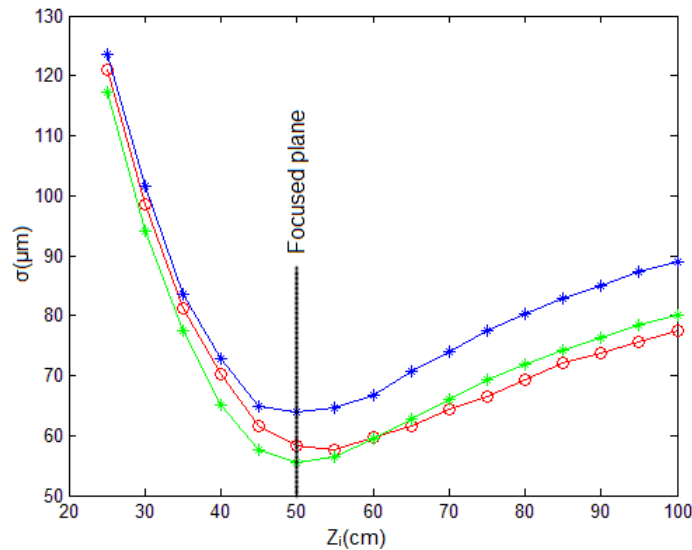


Figure 6.15 – Relation entre le flou et l'éloignement pour les 3 canaux couleurs.

Les résultats obtenus sur des images réelles sont illustrés sur la figure 6.16. Les premiers résultats que nous obtenons sont encourageants, mais restent prohibitifs en temps de calcul, car dans cette méthode on cherche à estimer les disparités données en x et y par ε_x et ε_y . Cependant, avec notre proposition, les disparités calculées dans l'étape de fusion des données et le calcul de l'erreur quadratique peuvent être remis à zéro, ce qui peut conduire à une réduction du temps de calcul non négligeable.

La figure 6.16.a est une image de 3 plans fournie par le centre MOIVRE. Le résultat de l'application de l'algorithme [Deschenes 2003] est montré sur la figure 6.16.b. Sur la figure 6.16.c, nous avons aussi augmenté artificiellement⁵ le flou sur la zone du brut, pour les trois images couleurs, tout en respectant le rapport entre les flous, donné par les α_C , e.g. figure 6.14, afin d'intensifier la région du brut. Nous remarquons aussi que la région de la bride est aussi marquée. Le résultat est un peu bruité, notamment à la frontière des régions présentant une discontinuité de luminance, car notre proposition est sensible à la longueur d'onde. Dans le cas idéal, il faut utiliser une image contenant des plans de la même texture placés à différentes profondeurs. Cas de l'image fournie par le centre MOIVRE. De plus dans l'expérimentation de la figure 6.16.b, le plan net est positionné au niveau de la plaque. Autrement dit, tous les plans proches de la caméra sont flous. Sur la figure de l'étalonnage de notre caméra, e.g. la figure 6.15, il est montré que le rapport entre l'éloignement et la longueur d'onde est petit.

5. Par l'application d'un filtre gaussien.

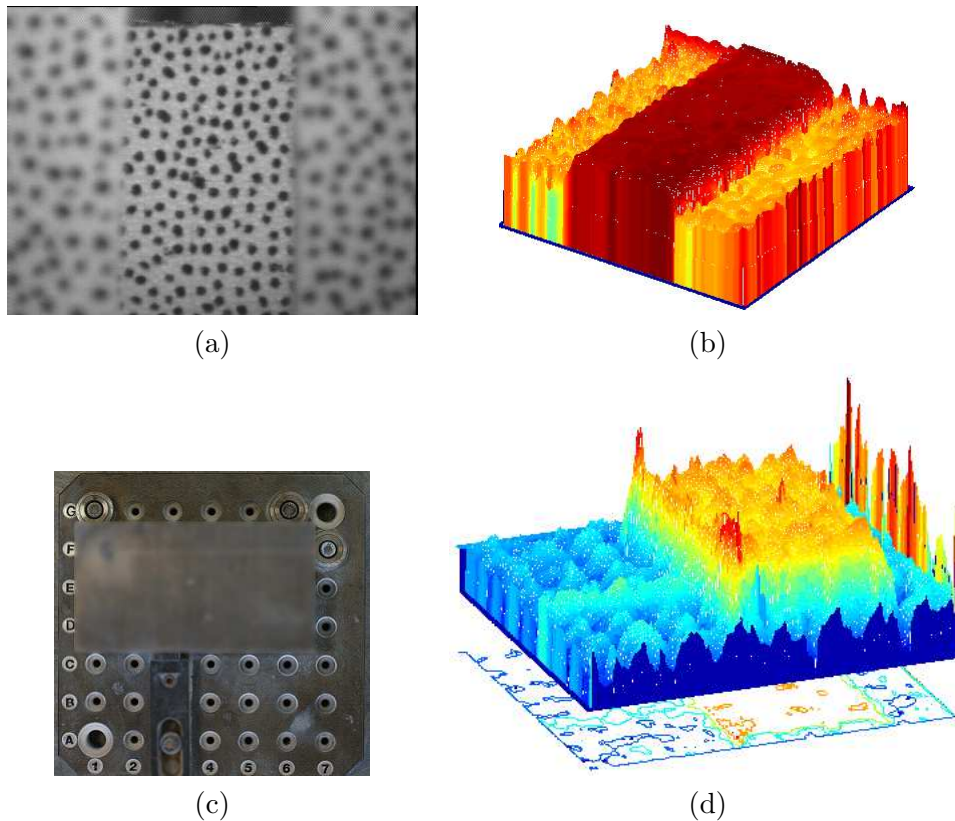


Figure 6.16 – Résultat de l'estimation de la carte de profondeur. (a) image de 3 plans, (b) carte de profondeur estimée, (c) image du montage d'usinage sur la plaque Norelem[®], (d) carte de profondeur estimée.

6.6 La réflectance de l'image pour l'estimation de la carte de profondeur : *Shape From Shading*

Dans l'objectif d'estimer la carte de profondeur à partir d'une seule caméra, nous avons testé la méthode *Shape from Shading* qui est réputée rapide. Nous présentons brièvement dans cette section le principe de la méthode et les résultats obtenus à l'issue de l'application de la méthode proposée dans la thèse de P. DANIEL [Daniel 2000].

La technique de *shape from shading*, exploite la fonction de réflectance d'une scène pour en déduire le relief, la configuration 3D. En général, la méthode demande l'acquisition de plusieurs images d'un même point de vue mais avec plusieurs orientations de l'éclairage. Les images en niveaux de gris acquises sont d'autant plus valides pour représenter le relief que les hypothèses classiques de la théorie du *shape from shading* sont respectées. Les hypothèses usuelles sont résumées dans la liste suivante issue du travail de thèse de P. Daniel [Daniel 2000] :

- L'image est supposée nette en tout point.
- L'image est considérée comme la projection orthogonale de la scène 3D, e.g. modèle *sténopé*

- Le récepteur est linéaire dans sa conversion photon \rightarrow électron.
- Il n'y a qu'une seule source ponctuelle de lumière placée suffisamment loin de la scène pour que l'on puisse considérer qu'elle est à *l'infini* et qu'elle a de très faibles dimensions angulaires. L'éclairage de la scène se ramène donc à un faisceau lumineux "parallèle" et "uniforme", e.g. flux lumineux uniforme, qui peut être décrit par un vecteur \vec{S} .
- Les surfaces observées sont lisses, la fonction $Z(x, y)$ représentant chaque surface est continue et dérivable.
- Les surfaces sont opaques, homogènes en réflexion.
- Le faisceau lumineux est parallèle à l'axe optique \vec{OZ}
- La surface est diffusante, lambertienne.

Avec ces hypothèses le niveau de gris $I(x, y)$ d'une image est lié à la surface par l'expression (6.28) :

$$I(x, y) = R(\vec{N}) \quad (6.28)$$

Où \vec{N} est la normale à la surface au point P éclairé et visible, R est la fonction de réflectance dépendante du vecteur, des caractéristiques de l'observateur et des caractéristiques photométriques de la surface.

On trouve alors la relation (6.29) :

$$I(x, y) = \frac{I_{max}}{\sqrt{1 + p^2 + q^2}} \quad (6.29)$$

où I_{max} est le niveau de gris maximal atteint lorsque \vec{N} est parallèle à \vec{S} , p et q sont les dérivées partielles de $Z(x, y)$.

La solution revient à trouver $Z(x, y)$ par l'intermédiaire de p et q , connaissant $I(x, y)$ et I_{max} pour différents \vec{S} .

Avec,

$$\vec{S} = S \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \quad (6.30) \quad \text{et} \quad \vec{N} = \frac{1}{\sqrt{1 + p^2 + q^2}} \begin{pmatrix} -p \\ -q \\ 1 \end{pmatrix} \quad (6.31)$$

Plusieurs méthodes permettent de résoudre les équations différentielles non linéaires obtenues par l'équation (6.29) et les différentes acquisitions.

Nous avons demandé à des spécialistes⁶ du domaine *shape from shading* de tester leur démonstrateur sur nos scènes de montage d'usinage. La figure 6.17 suivante montre un exemple de résultats. Le démonstrateur est composé d'une caméra munie d'un éclairage annulaire à "led" commandable selon 8 sections de l'anneau pour obtenir 8 éclairages différents. La figure 6.17.a montre l'image du montage d'usinage sur la plaque Norelem[©] après application de ressuyage et création d'une texture. La figure 6.17.b montre la carte de profondeur estimée à partir du démonstrateur.

6. IRIT- Université Paul Sabatier, 118 Route de Narbonne, F-31062 TOULOUSE CEDEX 9.

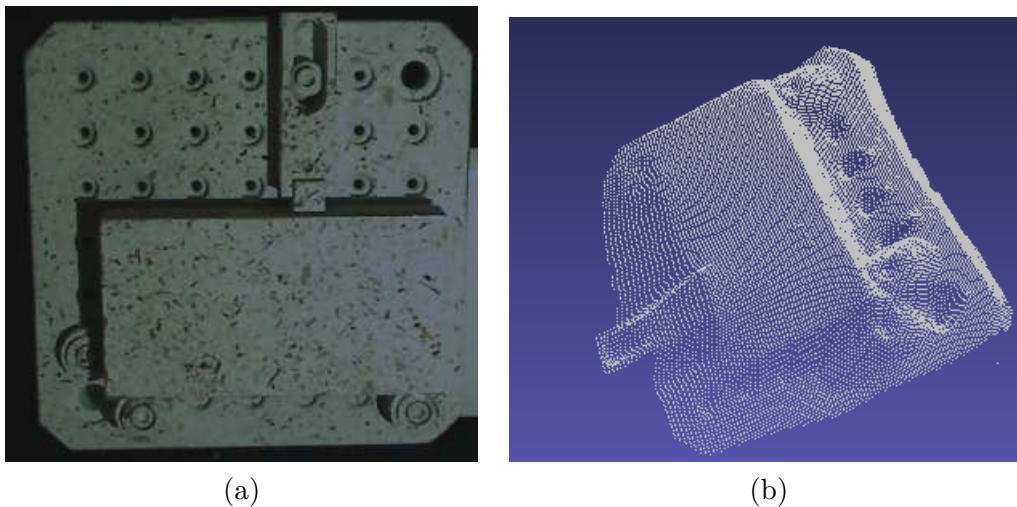


Figure 6.17 – Résultat de l'estimation de la carte de profondeur à partir de méthode *shape from shading*. (a) image du montage d'usinage sur la plaque Norelem[©] après application de ressuage et création de texture, (b) carte de profondeur estimée.

Les résultats obtenus ne sont pas bons car l'hypothèse de surfaces lisses n'est pas respectée, les éléments du montage d'usinage présentent de nombreuses ruptures des normales.

Afin de tester la possibilité d'une solution 3D, nous avons numérisé le montage d'usinage sur la plaque Norelem[©] en utilisant les résultats du chapitre 2. Ces résultats sont donnés sous forme d'un nuage de points 3D. Afin de pouvoir exploiter cette information 3D, nous proposons de créer une image de profondeur à partir de ce nuage de points. Le détail est présenté dans la section 6.7 suivante.

6.7 Création de l'image RGBD

Le couplage de l'information couleur RGB et l'information de profondeur, *Depth*, synthétisée dans une image RGBD, peut jouer un rôle important dans la segmentation d'image [Holz 2012], [Richtsfeld 2012], [Filliat 2012]. Pour remédier aux cas nécessitant l'information 3D pour la prise de décision, i.e. cas d'une bride en plus, en dehors des régions d'intérêt. Nous avons créé une image de profondeur à partir du nuage de points donné par un système optique de numérisation tridimensionnel sans contact.

Notre objectif est de créer une image de profondeur dans la même résolution que l'image RGB. Effectivement, cela est possible avec une connaissance *a priori* des paramètres intrinsèques et extrinsèques du système de numérisation 3D. Une fois ces paramètres fournis, nous projetons le nuage de points 3D calculés sur une image dont la résolution est la même que l'image RGB.

Selon la résolution et la qualité du système de numérisation utilisé, la création de l'image de profondeur à partir du nuage de points peut faire apparaître du bruit dûs aux trous ou aux surfaces réfléchissantes. Pour remédier au problème de bruit qui pourrait affecter la segmentation et le calcul des contours 3D dans l'image de profondeur, nous

appliquons un filtre médian d'une fenêtre (5×5). Un exemple de résultat est donné sur la figure 6.18.

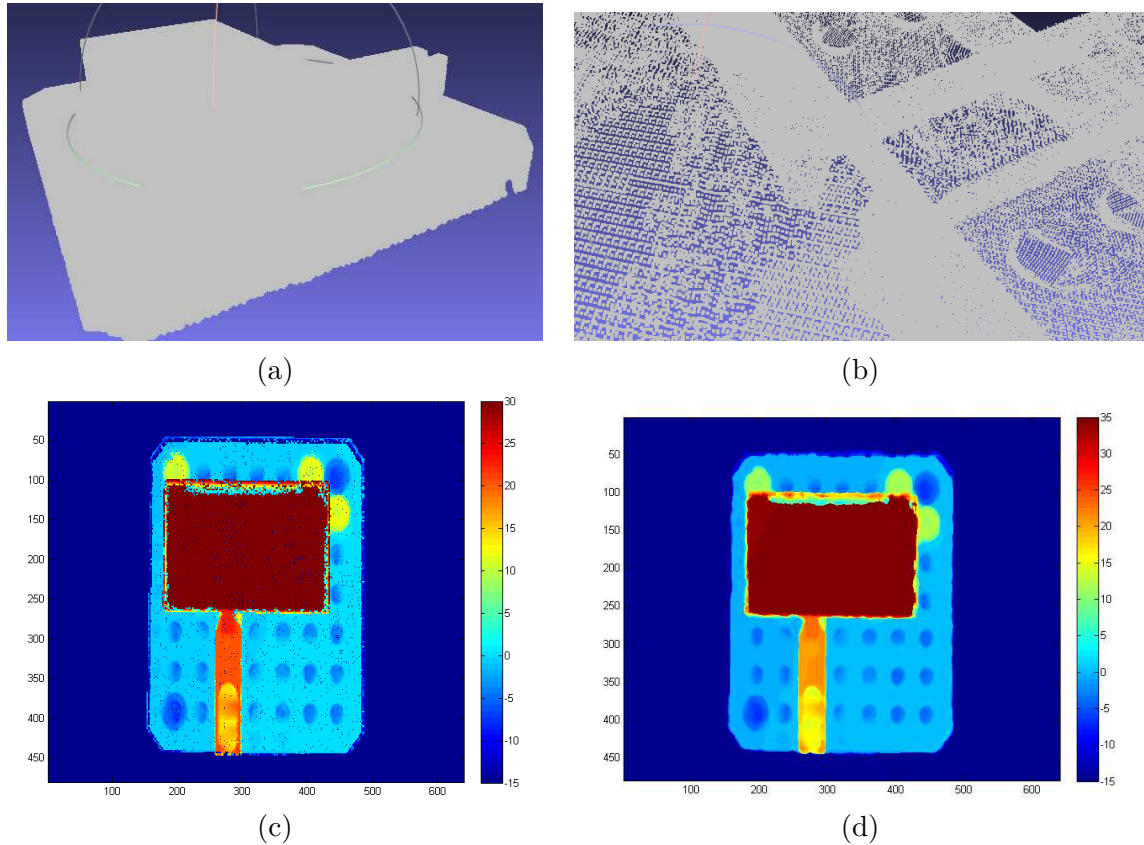


Figure 6.18 – Création de l'image RGBD : (a) nuage de points 3D, (b) Zoom avant sur le nuage du point, (c) image de profondeur (d) image de profondeur après application du filtre médian.

Cette carte de profondeur, e.g. carte d'élévation, peut aussi permettre de faciliter la segmentation des régions d'intérêt. La segmentation peut être donnée par une comparaison directe entre la carte de profondeur calculée et la surface de référence donnée par le modèle CAO. Un exemple est donné sur la figure 6.19.

Dans le cas de présence d'un objet en dehors des régions d'intérêt données par le modèle de référence CAO, l'utilisation de l'information 3D permet la détection de la présence cette l'anomalie, contrairement à l'utilisation unique de la luminance 2D. Un résultat de comparaison entre une représentation de la plaque avec un élément en dehors de la zone d'intérêt à la fois en utilisant la carte de profondeur et la luminance, est illustré sur la figure 6.20.

6.8 Conclusion

Les premiers résultats de notre proposition d'amélioration de la méthode de l'université de Sherbrooke par l'exploitation des flous couleurs montre la possibilité d'ob-

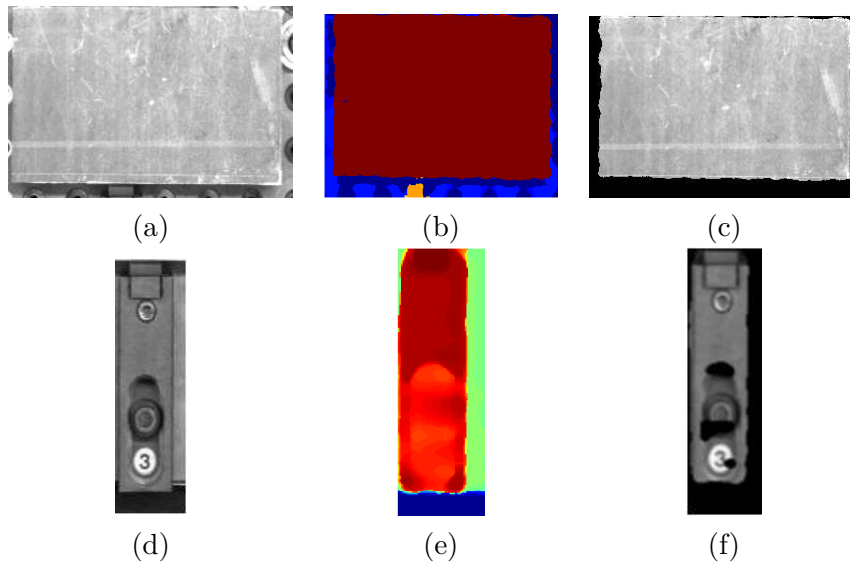


Figure 6.19 – Segmentation de l'image RGBD : (a) image de luminance de la région du brut, (b) image de profondeur de la région du brut, (c) image RGBD de la région du brut (d) image de luminance de la région de la bride de fixation, (e) image de profondeur de la région de la bride, (f) image RGBD de la région de la bride

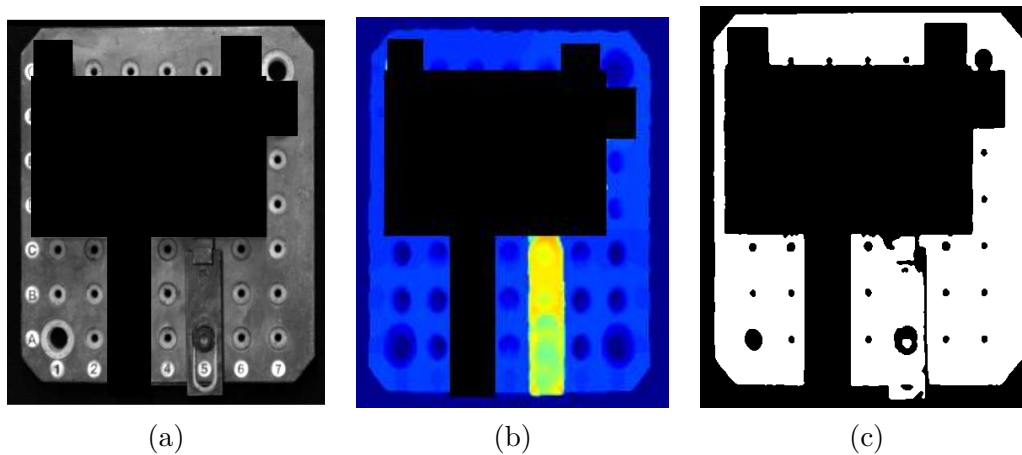


Figure 6.20 – Représentation de l'intérêt de l'information 3D pour la segmentation : (a) image de luminance du montage d'usinage, dont les régions d'intérêt sont remises au fond, "0", (b) image de profondeur donnant l'apparition de l'anomalie, (c) image segmentée par la méthode d'Otsu.

tenir une carte des profondeurs denses d'un montage d'usinage. Ce travail conséquent n'a pu être mené en totalité dans le temps de ce travail doctoral, mais il ouvre des perspectives.

L'information 3D reste une donnée importante à obtenir pour confronter le montage d'usinage réalisé, observé avec la trajectoire de l'outil. Dans une extension de ce travail, nous imaginons la possibilité de proposer à l'opérateur dans la phase d'optimisation de

la trajectoire, de préparer informatiquement les transformations qu'il a réalisées pour une récupération par le système de FAO pour le calcul de la nouvelle trajectoire qui permettrait à l'opérateur de ne pas ressaisir ses modifications dans la nouvelle FAO.

L'information 3D permettra de lever toutes les ambiguïtés liées à l'ajout d'éléments hors des zones prédéfinies par la CAO.

Bibliographie

- [Daniel 2000] P. Daniel. Peut-on extraire le relief d'une seule image ? PhD thesis, Institut de Recherche en Informatique de Toulouse, 28 Janvier 2000. (Cité en pages 150 et 167.)
- [Deschenes 2003] F. Deschenes, D. Ziou et P. Fuchs. Improved Estimation of defocus blur and spatial shifts in spatial domain : A homotopy-based approach. Pattern Recognition, vol. 36, pages 2105–2125, 2003. (Cité en pages v, 28, 149, 150, 151, 160, 163, 166 et 177.)
- [Dey 2002] N. Dey, A. Boucher et M. Thonnat. Image formation model of a 3D translucent object observed in light microscopy. volume 2, pages 469–472, 2002. (Cité en page 153.)
- [Filliat 2012] D. Filliat. RGBD object recognition and visual texture classification for indoor semantic mapping. pages 127–132, 2012. (Cité en page 169.)
- [Holz 2012] D. Holz, R. Holzer S. Bogdan.-Rusu et S. Behnke. Real-Time Plane Segmentation using RGB-D Cameras. pages 306–317, 2012. (Cité en page 169.)
- [Paolo 2007] Favaro Paolo et Soatto Stefano. 3-D Shape Estimation and Image Restoration - Exploiting Defocus and Motion Blur. pages 1–249, 2007. (Cité en pages 28 et 152.)
- [Pentland 1987] P-A. Pentland. A New Sense for Depth of Field. volume 9, pages 523–531, 1987. (Cité en page 152.)
- [Richtsfeld 2012] T. Richtsfeld A.and Morwald, J. Prankl, M. Zillich et M. Vincze. Segmentation of Unknown Objects in Indoor Environments. In IEEE RSJ International Conf. Intelligent Robots and Systems (IROS), 2012. (Cité en pages 143 et 169.)
- [Schneider 1994] G. Schneider, B. Heit, J. Honig et J. Bremont. Monocular depth perception by evaluation of the blur in defocused images. volume 2, pages 116–119, 1994. (Cité en page 152.)
- [Simon 2004] C. Simon, F. Bicking et T. Simon. Depth Estimation Based On Thick Oriented Edges in Images. volume 1, pages 135–140, 2004. (Cité en page 150.)
- [Stokseth 1969] P.A. Stokseth. Properties of a defocused optical system. volume 59(10), pages 1314–1321, 1969. (Cité en pages xiii, 153 et 154.)
- [Subbarao 1988] M. Subbarao et G. Surya. Depth Recovery from Blurred Edges. pages 498–503, 1988. (Cité en page 152.)

-
- [Subbarao 1994] M. Subbarao et G. Surya. Depth From Defocus : A spatial Domain Approach. volume 13, pages 271–297, 1994. (Cité en page 160.)
- [Trevor 1988] D. Trevor et K. Wohn. Calculating Depth From Focus Using A Pyramid Architecture. volume 57, pages 271–297, 1988. (Cité en page 150.)
- [Zhou 2009] C. Zhou, S. Lin et S. K. Nayar. Coded Aperture Pairs for Depth from Defocus. In IEEE International Conference on Computer Vision (ICCV), Oct 2009. (Cité en page 150.)
- [Zhuo 2011] S. Zhuo et T. Sim. Defocus Map Estimation from a Single Defocused Image. Pattern Recognition, vol. 44, no. 9, pages 1852–1858, 2011. (Cité en pages v, xiii, 28, 149, 150, 152, 157, 158, 159 et 160.)

CONCLUSION ET PERSPECTIVES

Sommaire

7.1 Conclusion	175
7.2 Perspectives	177

7.1 Conclusion

L'objectif de cette thèse concerne le développement d'un système de vision pour la vérification automatique des montages d'usinage. Il s'agit d'augmenter la productivité des PME spécialisées dans l'usinage, grâce à la sécurisation des machines contre les risques de collision de l'outil avec les éléments de bridage.

Le cahier des charges impose de prendre en compte trois paramètres principaux, à savoir le temps de calcul, le coût du système et sa simplicité, et l'encombrement pour pouvoir installer le système de contrôle au sein de la machine. Nous avons proposé une solution utilisant la vision monoculaire. L'utilisation d'un système monoculaire pourrait répondre aux exigences du cahier des charges. En ce que concerne le temps du calcul, l'exploitation d'une seule image est beaucoup plus intéressante que celle de deux images d'un système binoculaire, ou de deux acquisitions d'une seule caméra avec deux points de vue différents. En ce que concerne le coût du système et la simplicité d'installation, l'utilisation d'une seule caméra est évidemment moins cher et moins encombrant qu'un système de numérisation tridimensionnelle. Enfin, pour le coût du système, nous envisageons que celui-ci soit autour de 10% de la valeur de la broche. Toutes ces contraintes industrielles, nous les retrouvons dans les différents chapitres du manuscrit.

Dans le premier chapitre, nous avons testé quelques systèmes commerciaux de numérisation tridimensionnelle. Malgré leur précision, ces systèmes ont un coût relativement élevé et nécessitent un temps de calcul assez important par rapport au gain de production souhaité. Notre solution monoculaire permet d'apporter une décision automatique et rapide sur la conformité du montage d'usinage, à partir d'une seule image. Nous avons proposé une configuration de la caméra au-dessus du montage d'usinage pour assurer une meilleure visibilité de la scène. Nous avons proposé deux configurations selon le type de machine.

- Pour le premier type de machine, nous proposons d'installer la caméra au-dessus du montage d'usinage à l'endroit de préparation de ce dernier.
- La seconde proposition concerne des machines pour lesquelles la préparation du montage d'usinage se fait directement dans l'espace de l'usinage, sur la table de l'usinage. Nous proposons d'installer la caméra dans le magasin des outils. Pour le contrôle de l'usinage et avant le lancement de la fabrication de pièce, l'opérateur n'a qu'à demander à la machine de chercher la caméra dans le magasin des outils, une opération qui ne nécessite qu'une ligne de code supplémentaire au début du programme de code G.

Cette configuration permet aussi de faciliter les traitements de la mise en correspondance pour le calcul de la pose et le recalage du modèle CAO sur l'image, la caméra étant installée dans la même position que la broche.

La mise en correspondance est effectuée par l'intermédiaire des graphes construits à partir des primitives simples, e.g. centre des cercles et intersection des droites. Ces primitives sont sélectionnées à partir de l'image des contours. Cette dernière est calculée par une nouvelle méthode s'appuyant sur la modélisation explicite de l'histogramme du module des gradients que nous avons proposée. Nous avons comparé cette nouvelle méthode par rapport aux indicateurs de performances donnés notamment par les courbes (*PR*), (*ROC*). Nous avons proposé aussi des critères de sélection des primitives à partir de l'image réelle du montage d'usinage. Ces critères permettent de regrouper les points pertinents afin de mieux représenter les relations topologiques entre les éléments de bridage de la scène y compris le brut.

Le recalage automatique du modèle CAO sur l'image réelle est donné par la mise en correspondance de graphes. Nous avons proposé une nouvelle méthode, basée sur l'information spatiale. Les graphes de proximité et l'utilisation du critère de voisin proche dans la recherche de la configuration Φ permettent d'assurer la meilleure correspondance, rendant la méthode proposée robuste et stable même pour des configurations réalisées "non conforme" du montage d'usinage.

De plus, la méthode demeure stable même pour des situations où le montage vu par la caméra est totalement différent du montage attendu par la machine. A l'issue de cette étape, les emplacements attendus des éléments de bridage, selon la définition CAO, sont localisés.

Une fois les régions d'intérêt déterminées, l'algorithme enchaîne avec une étape d'analyse locale pour la prise de décision. Nous avons utilisé un système expert flou. En effet, l'utilisation du système expert flou permet une analyse de mesure de similarité possibiliste conduite par les scores calculés sur les paramètres descriptifs de l'objet, e.g. le squelette, les contours et la dimension donnée par la surface. Nous avons montré le rôle important de la combinaison des trois paramètres pour la description de l'objet par la prise en compte de l'incertitude, dans la mesure de la similarité conduite sur les différentes étapes de l'algorithme, e.g. du calcul des contours jusqu'à la prise de décision. Nous avons proposé une modélisation des règles floues adaptée à notre problématique. Cette proposition satisfait la combinaison des paramètres descriptifs que nous avons établie et justifie ce choix.

Nous avons montré que, dans certains cas, l'intégration de l'information tridimensionnelle est nécessaire, à la fois pour répondre à des situations ambiguës, e.g. présence d'un élément en dehors des régions d'intérêt, objet qui n'a pas la bonne dimension et qui donne la même empreinte dans l'image et enfin, pour répondre au cas "orange" dans la chaîne numérique du traitement de l'usinage proposée, e.g. montage non conforme mais qui ne provoque pas de collision de l'outil avec les éléments de bridage. Ces deux situations nécessitent un test de simulation de la trajectoire de l'outil sur une carte dense de profondeur de la scène réelle, pour valider le montage réalisé.

Dans l'objectif de calculer une carte de profondeur à partir d'une seule acquisition d'image, nous avons proposé un nouveau principe permettant d'exploiter le flou optique. Ce principe s'articule sur la différence entre les réponses impulsionnelles d'une caméra couleurs, selon la longueur d'onde à partir des images de chrominances (*RGB*). Nous avons réalisé une étude permettant de montrer la relation entre l'éloignement de l'objet par rapport au plan image et la longueur d'onde. Ce résultat confirme la possibilité d'exploiter cette piste. Nous avons testé, pour l'instant, notre proposition de solution sur une méthode existante [Deschenes 2003]. Les résultats obtenus sont encourageants et donnent des perspectives. Notre objectif, dans le futur proche, est de proposer une méthode modélisant plus précisément ce phénomène et permettant un calcul rapide de la carte de profondeur.

7.2 Perspectives

Ce travail a permis de répondre à une grande partie de la question de conformité du montage d'usinage par rapport au modèle CAO. **La détection et la localisation** de l'anomalie dans la réalisation du montage d'usinage sont obtenues. Il est aussi important que le système soit capable d'indiquer à l'opérateur au pied de la machine, dans la phase d'optimisation de la trajectoire, la possibilité de déplacement ou d'utilisation des éléments de bridage. Cette possibilité permet d'intégrer les modifications de l'opérateur sans recours à une nouvelle saisie complète en *FAO* pour le calcul de la nouvelle trajectoire. Cependant, cette solution nécessite la présence de l'information 3D afin d'interpréter la trajectoire par rapport à la représentation 3D de la scène. Ce qui donne aussi la possibilité de fournir à l'opérateur, non seulement la position de l'anomalie, mais aussi l'instant de sa production, car la production de certaines pièces se fait sur plusieurs étapes d'usinage, voire même sur différentes MOCN.

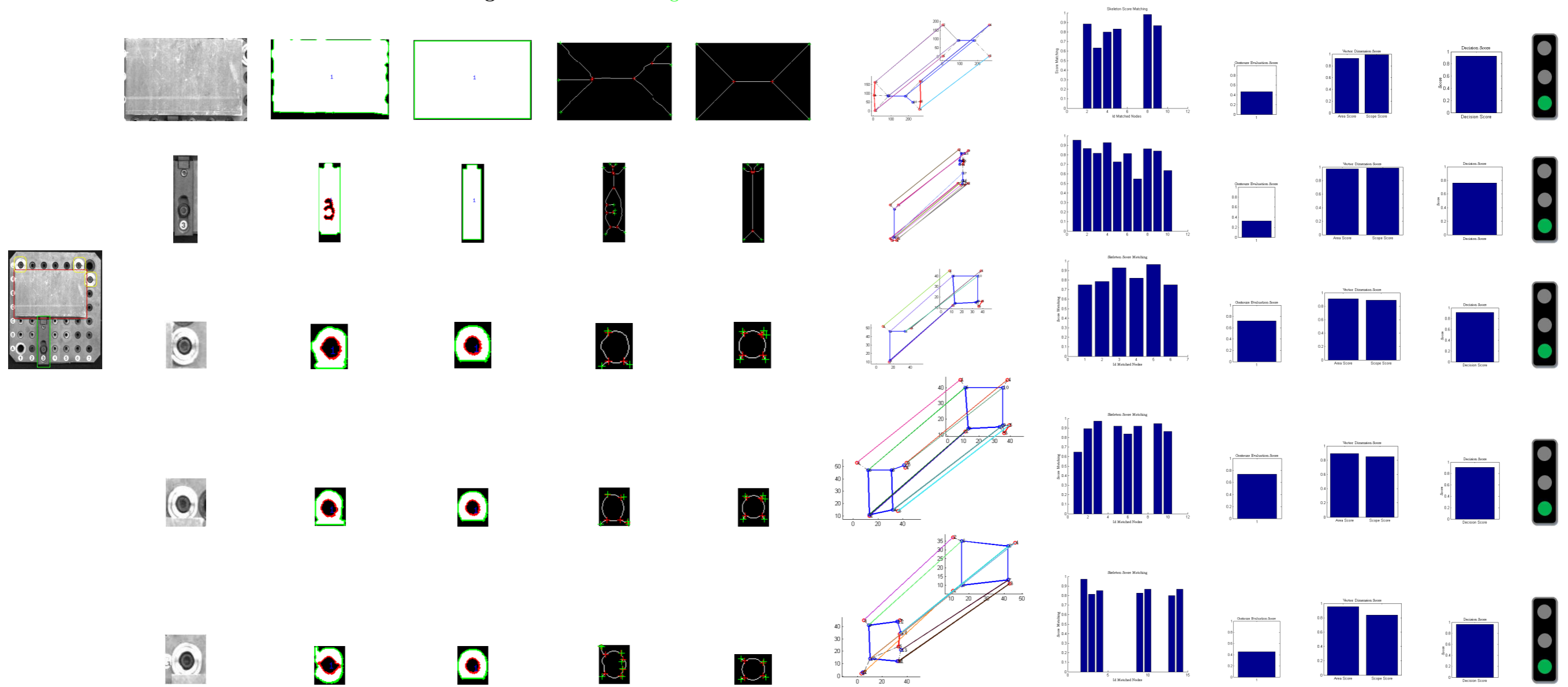
Le cœur de ce travail, repose sur l'utilisation de la vision par ordinateur pour répondre à une problématique industrielle. A chaque étape de solution, la recherche bibliographique nous a conduit à étudier différentes méthodes existantes et à proposer des méthodes nouvelles s'adaptant à notre problématique. L'intégration de la réalité virtuelle dans l'industrie est un point important permettant de sécuriser et augmenter la productivité des entreprises. Ce travail peut aussi être utilisé pour le contrôle et l'inspection de l'assemblage des pièces aéronautique, autre thèse en cours à l'ICA.

ANNEXE A

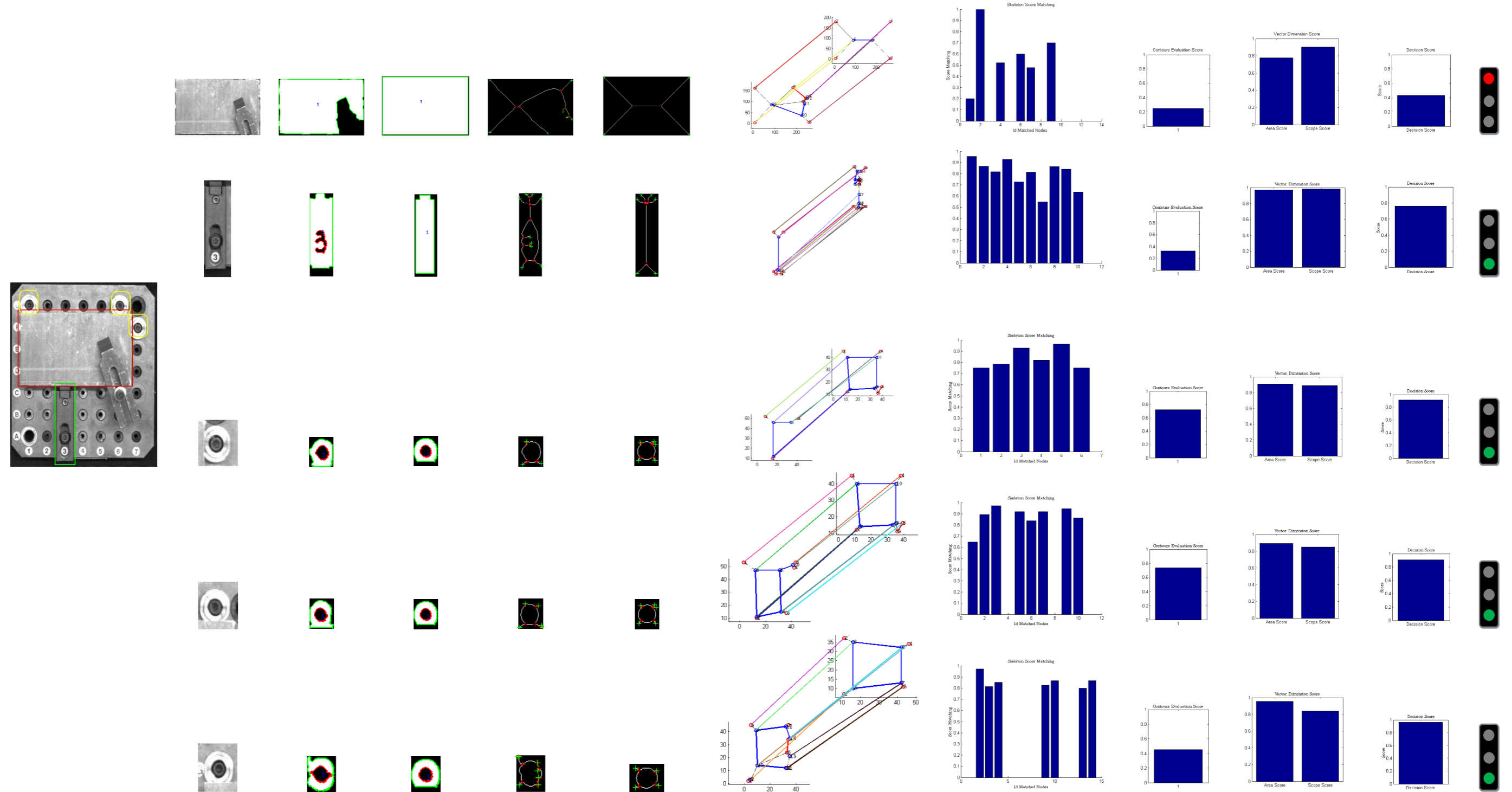
RÉSULTATS

A.1 Les résultats sont illustrés sur des pages de format A3

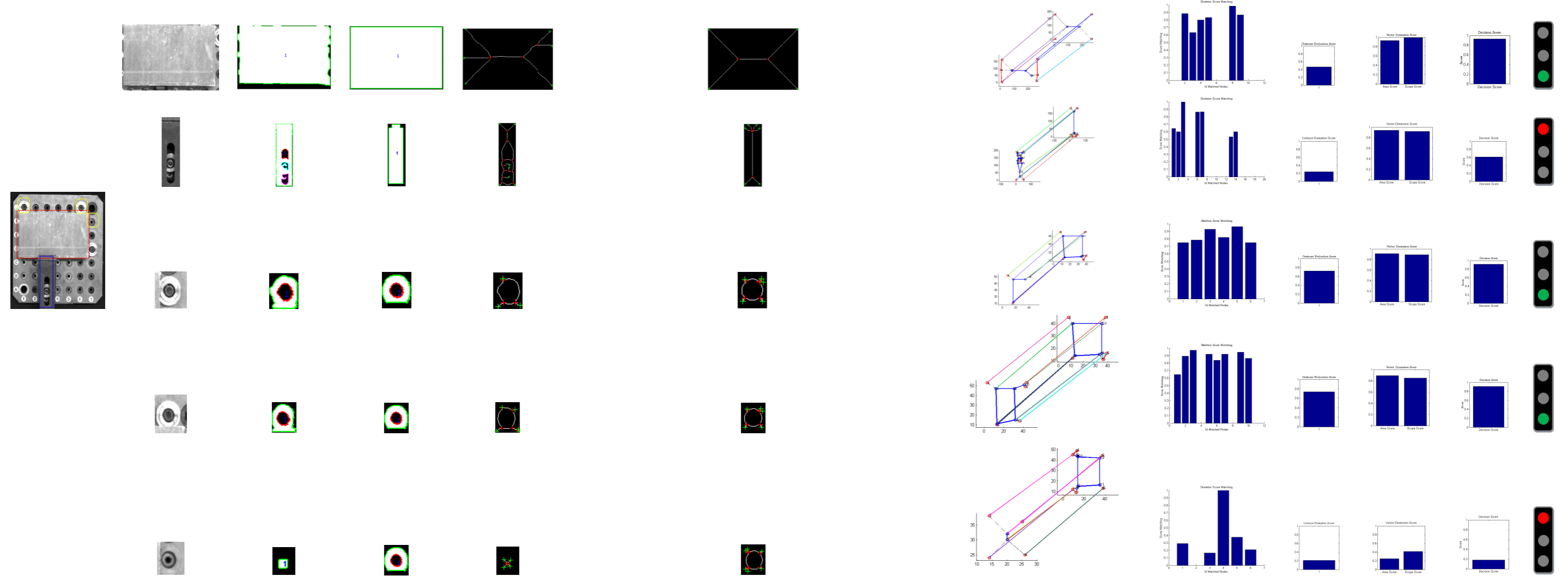
Montage Norelem : Configuration conforme



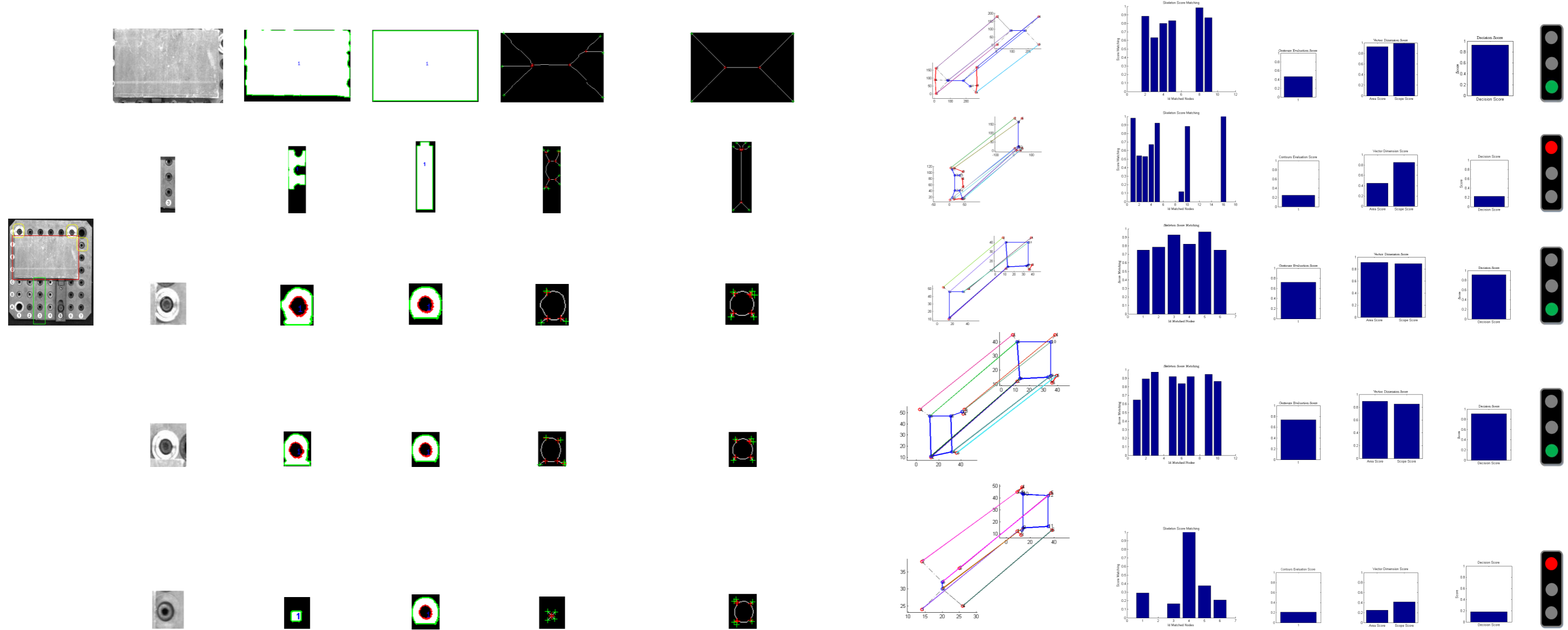
Montage Norelem : Configuration Bride en plus



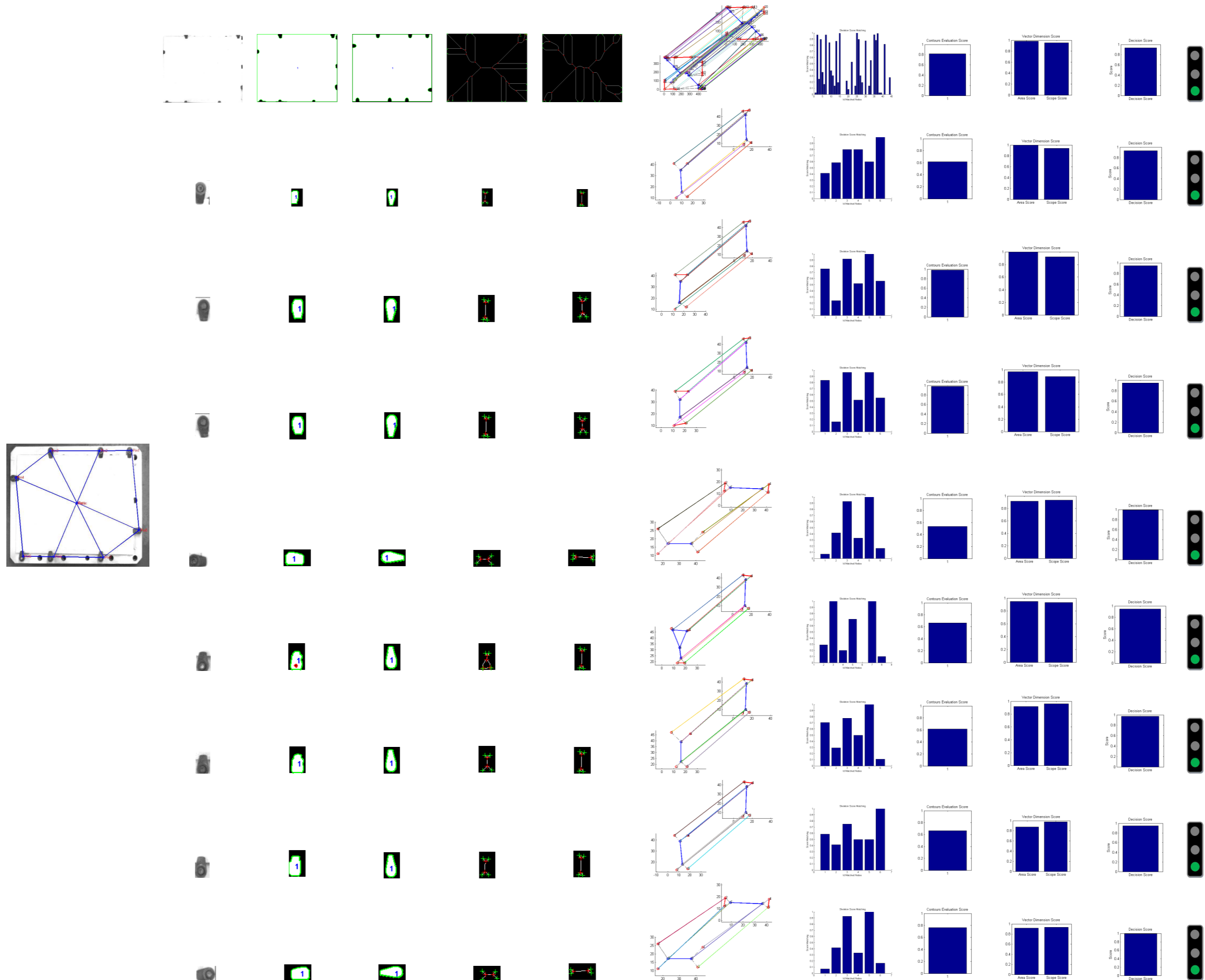
Montage Norelem : Configuration Bride qui n'a pas la bonne dimension et Pion en moins



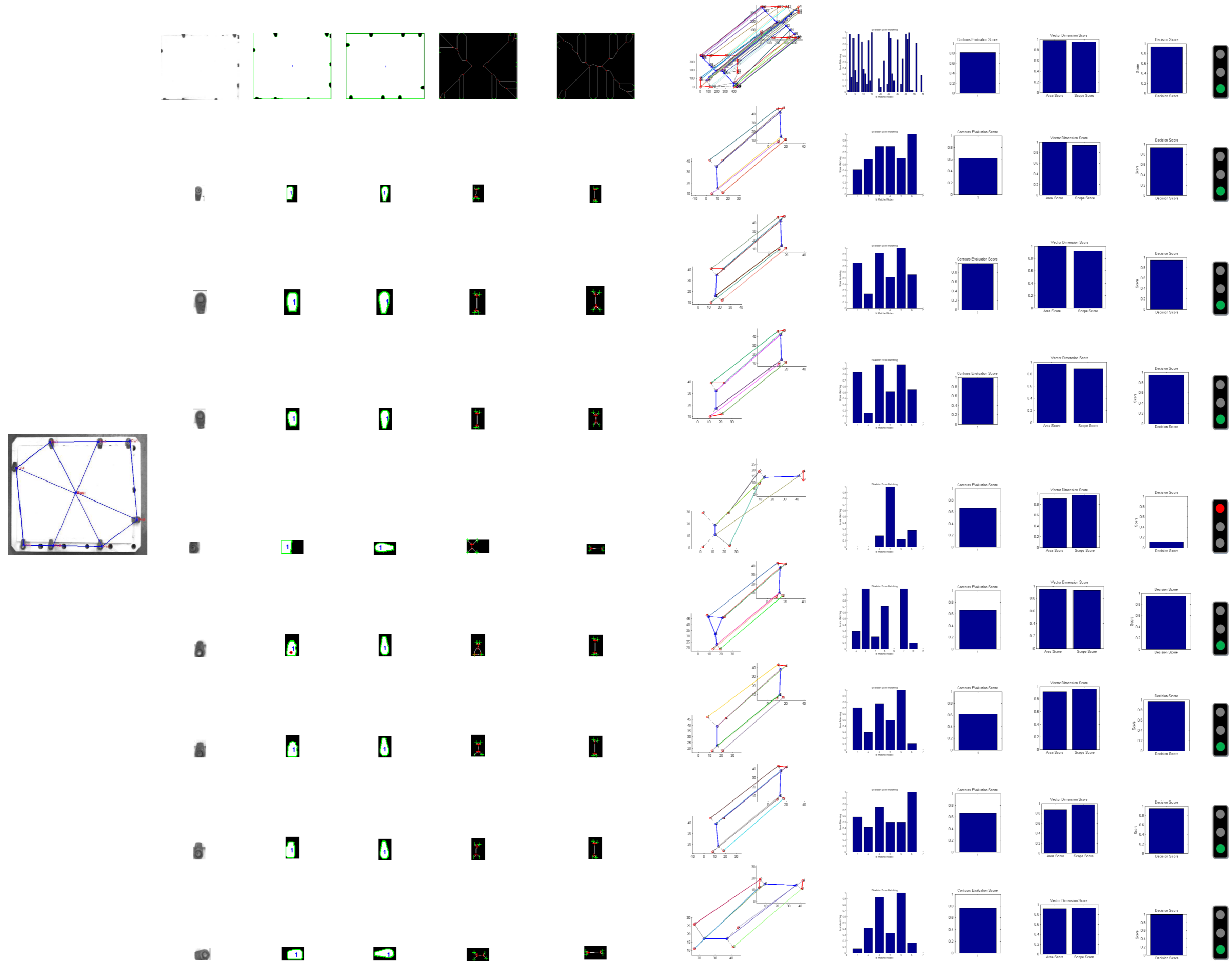
Montage Norelem : Configuration Bride dplace et Pion en moins



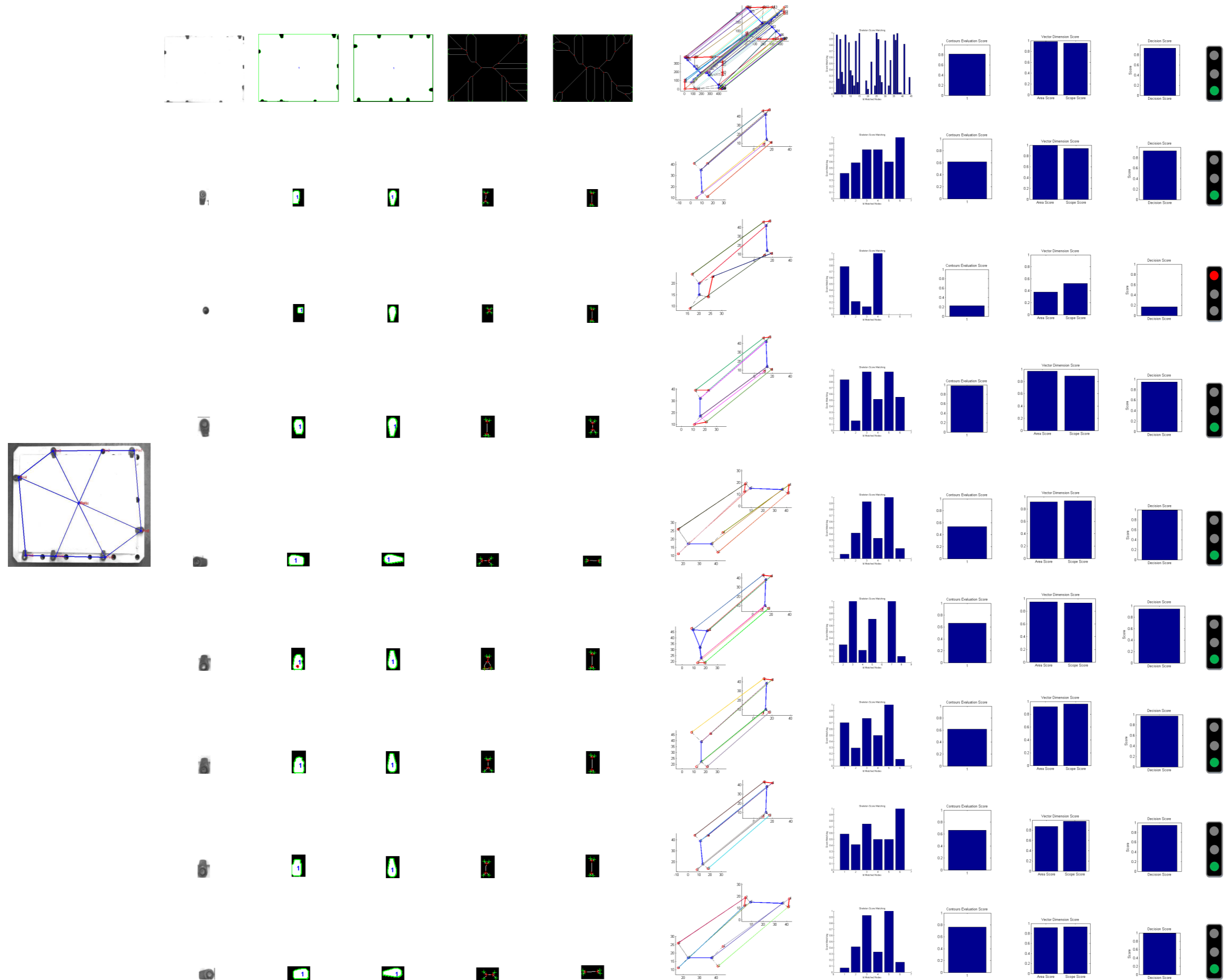
Montage Minimanche : Configuration conforme



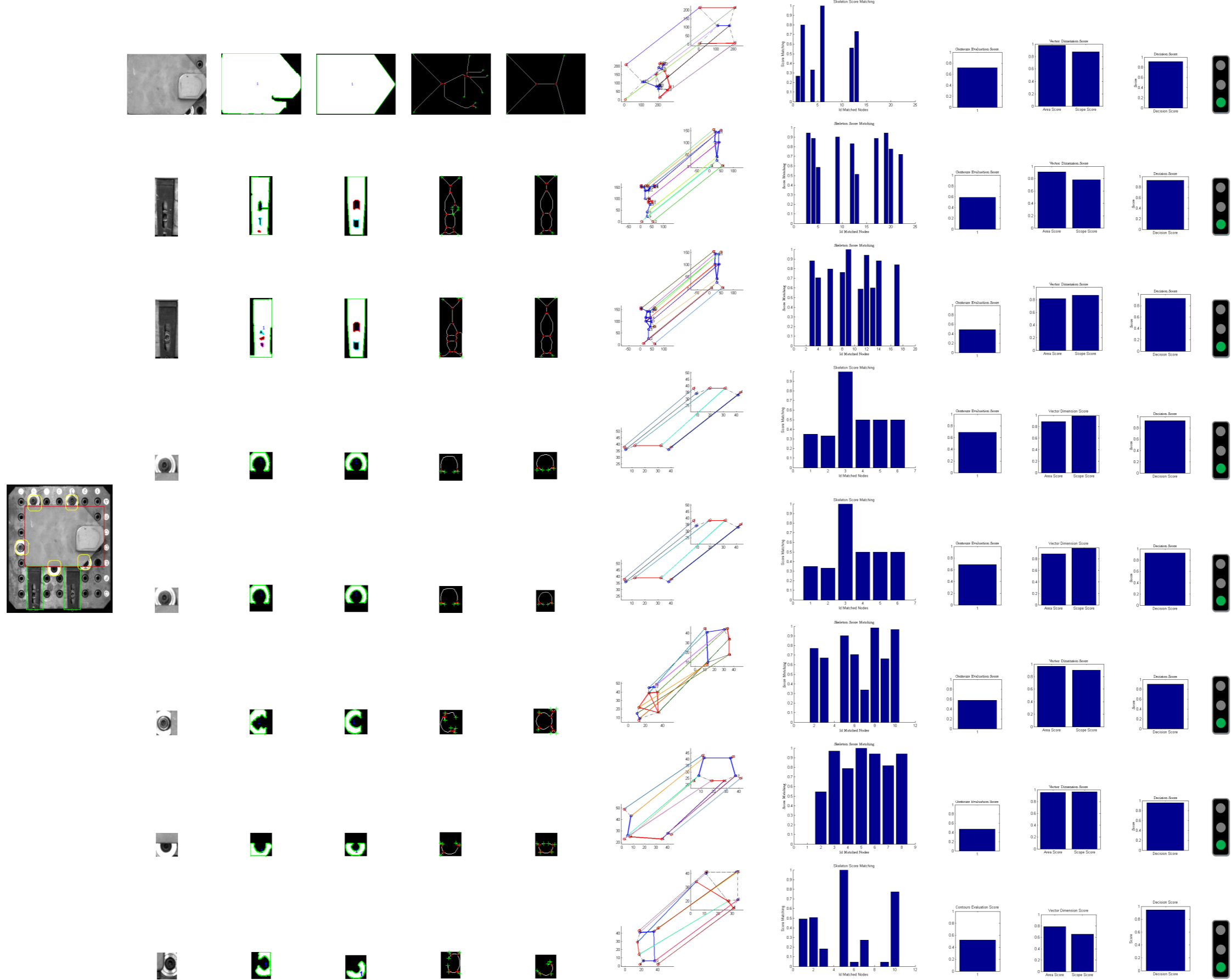
Montage Minimanche : Configuration Pion mal positionne



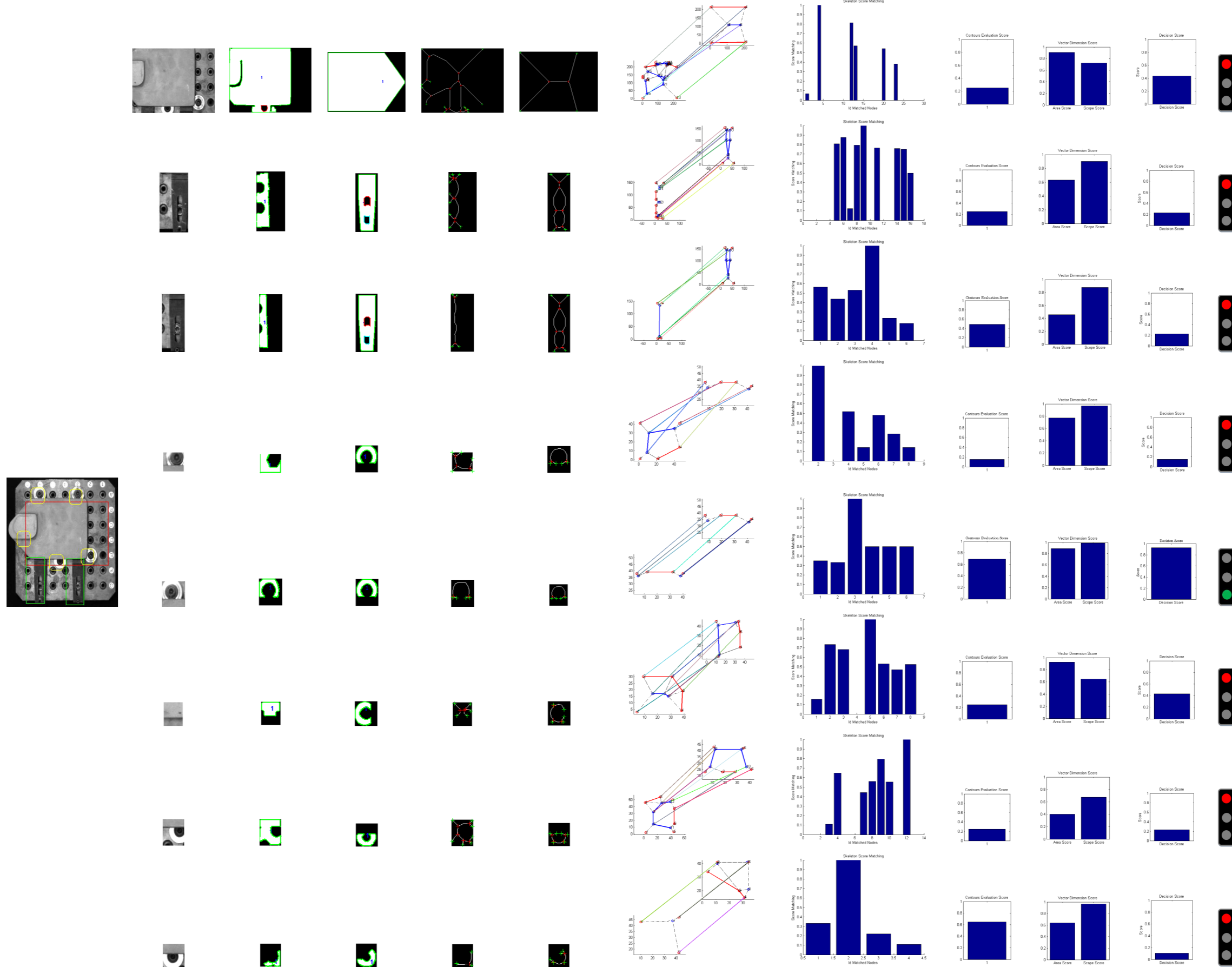
Montage Minimanche : Configuration Pion en moins



Montage Semelle : Configuration Conforme

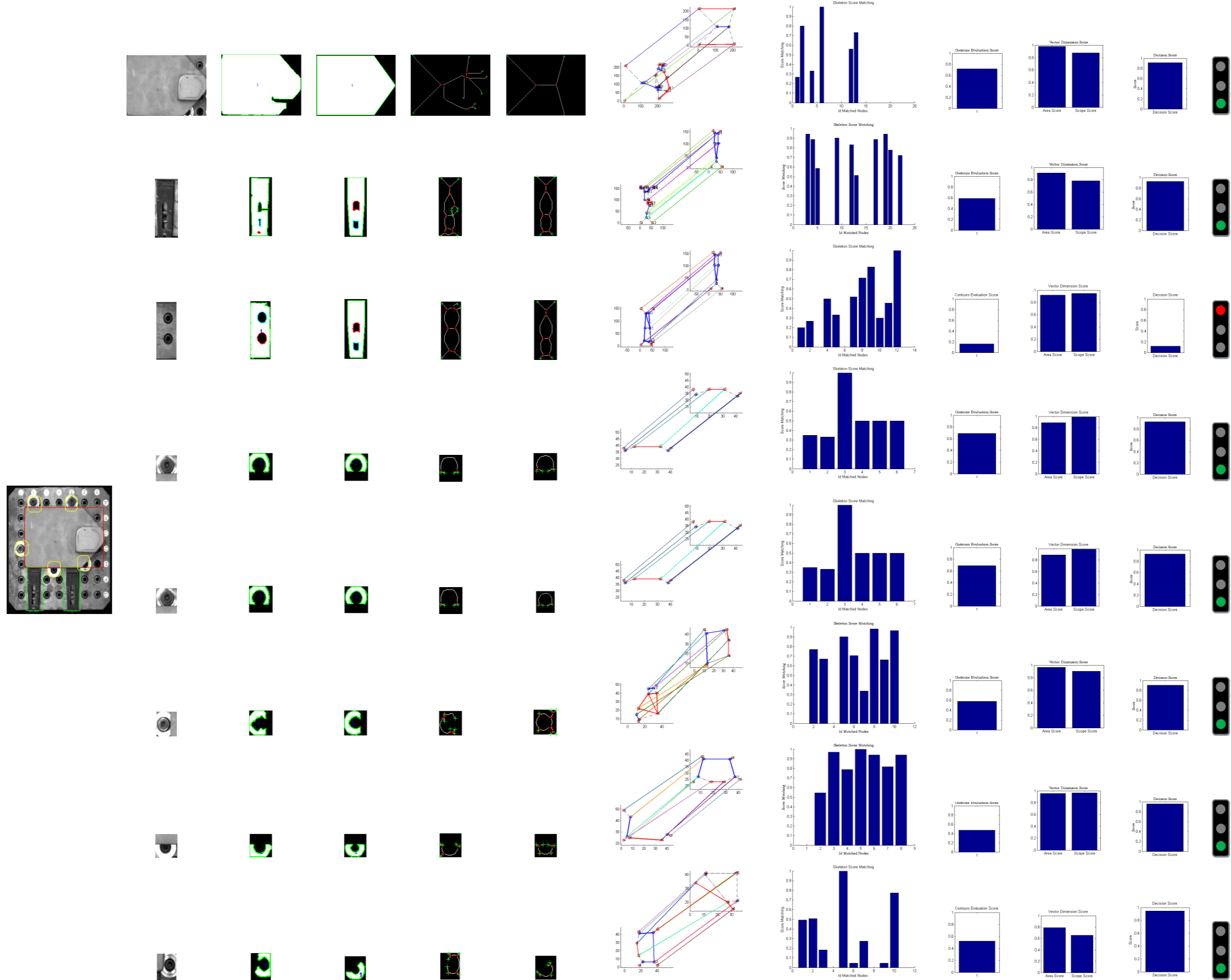


Montage Semelle : Configuration Inverse



Montage Semelle :

Configuration Bride en moins



Bibliographie

- [Ahmad 2010] R. Ahmad, S. Tichadou et J.-Y. Hascoët. Integration of Vision Based Image Processing for Multi-Axis CNC Machine Tool Safe and Efficient Trajectory Generation. *International Journal of Machine Engineering*, vol. 10, no. 4, 2010. (Cité en pages vii, 15 et 16.)
- [Arbelaez 2009] P. Arbelaez, M. Maire, C. C. Fowlkes et J. Malik. From Contours to Regions : An Empirical Evaluation. In *Conf. Computer Vision and Pattern Recognition, 2009.*, 2009. (Cité en page 112.)
- [Arcelli 1985a] C. Arcelli, K.-M. Lam et W.-C. Siu. Extraction of the Euclidean skeleton based on a connectivity criterion. *Pattern Recognition*, vol. 36, pages 721–729, 1985. (Cité en page 123.)
- [Arcelli 1985b] C. Arcelli, G. Sanniti di Baja et R. Perucchio. A Width Independent Fast Thinning Algorithm. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7, pages 463–474, 1985. (Cité en page 123.)
- [Arcelli 1989] C. Arcelli et G. S. Di Baja. A one-pass two-operation process to detect the skeletal pixels on the 4-distance transform. *Mathematical Morphology and its Applications to Image and Signal Processing*, vol. 11, no. 4, pages 411–414, 1989. (Cité en page 123.)
- [Bai 2007] X. Bai, L.J. Latecki et W.-Y. Liu. Skeleton Pruning by Contour Partitioning with Discrete Curve Evolution. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29(3), pages 449–462, 2007. (Cité en page 124.)
- [Bai 2008] X. Bai et L.J. Latecki. Path Similarity Skeleton Graph Matching. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 30(7), pages 1282–1292, 2008. (Cité en pages 124 et 125.)
- [Beveridge 1996] J.R. Beveridge, C. Graves et C. Leshner. Some Lessons Learned from Coding the Burns Line Extraction Algorithm in the Darpa Image Understanding Environment. Technical Report CS-96-125, Computer Science Dept. Colorado State Univ., 1996. (Cité en page 53.)
- [Bouguet 2004] J.-Y. Bouguet. Camera calibration toolbox for matlab, mar 2004. (Cité en pages 70, 89 et 90.)
- [Bouguila 2007] N. Bouguila, D. Ziou et R. I-Hammoud. A Bayesian Non-Gaussian Mixture Analysis : Application to Eye Modeling. *Conf. Computer Vision and Pattern Recognition, 2007.*, pages 1–8, June 2007. (Cité en page 116.)
- [Brandt 1992] J. W. Brandt et V. R. Algazi. Continuous skeleton computation by Voronoi diagram. *Computer Vision, Graphics, Image Processing*, vol. 55, pages 329–338, 1992. (Cité en page 123.)
- [Breu 1995] H. Breu, J. Gil, D. Kirkpatrick et M. Werman. Linear time euclidean distance transform algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, pages 529–533, 1995. (Cité en page 123.)

- [Brox 2005] T. Brox, B. Rosenhahn et J. Weickert. Three-Dimensional Shape Knowledge for Joint Image Segmentation and Pose Estimation. In *Pattern Recognition*, Springer LNCS 3663, pages 109–116, 2005. (Cité en page 69.)
- [Bunke 1983] H. Bunke et G. Allermann. Inexact graph matching for structural pattern recognition. *Pattern Recognition*, vol. 1, no. 4, pages 245–253, 1983. (Cité en page 70.)
- [Burns 1986] J.B. Burns, A.R. Hanson et E.M. Riseman. Extracting Straight Lines. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 4, pages 425–455, 1986. (Cité en page 53.)
- [Caelli 2004] T. Caelli et S. Kosinov. An eigenspace projection clustering method for inexact graph matching. *IEEE Trans. Pattern Analysis and Machine Intelligence.*, vol. 26(4), pages 515–519, April 2004. (Cité en page 69.)
- [Caux 2012] M. Caux. Méthodologie de maîtrise des variations géométriques des produits en conception, fabrication et contrôle dans le contexte de l’usine numérique. PhD thesis, ENS CACHAN, 29 Juin 2012. (Cité en page 3.)
- [Colombo 2004] Carlo . Colombo, Dario Comanducci, Alberto Del Bimbo et Alberto Del Federico Pernici. Accurate Automatic Localization of Surfaces of revolution for self-calibration and metric reconstruction. In *In Proc. IEEE Workshop on Perceptual Organization in Computer Vision*, page 67, 2004. (Cité en page 18.)
- [Cox 2005] D. Cox, J. Little et D. O’Shea. Using Algebraic Geometry, Second Edition. *Graduate Texts in Mathematics*, Springer, vol. 185, pages 364–374, 2005. (Cité en page 70.)
- [Daniel 2000] P. Daniel. Peut-on extraire le relief d’une seule image ? PhD thesis, Institut de Recherche en Informatique de Toulouse, 28 Janvier 2000. (Cité en pages 150 et 167.)
- [Debashis 2010] S. Debashis et K. Pal Sankar. Gradient histogram : thresholding in region of interest for edge detection. *Image and vision computing journal*, vol. 28, pages 677–695, 2010. (Cité en pages 37, 38, 39 et 40.)
- [Delaunay 1928] B. Delaunay. Sur la sphère vide. volume 1, pages 223–228, 1928. (Cité en page 74.)
- [Dementhon 1995] D. Dementhon et L.S. Davis. Model-based object pose in 25 lines of code. volume 15(1/2), pages 123–141, 1995. (Cité en pages 70 et 103.)
- [Deschenes 2003] F. Deschenes, D. Ziou et P. Fuchs. Improved Estimation of defocus blur and spatial shifts in spatial domain : A homotopy-based approach. *Pattern Recognition*, vol. 36, pages 2105–2125, 2003. (Cité en pages v, 28, 149, 150, 151, 160, 163, 166 et 177.)
- [Deschenes 2004] F Deschenes, D Ziou et M.-F. Auclair-Fortier. Detection of lines, line junctions and line terminations. *International Journal of Remote Sensing*, vol. 25, no. 3, pages 511–535, feb 2004. (Cité en page 53.)

- [Desolneux 2000] A. Desolneux, L. Moisan et J-M. Morel. Meaningful Alignments. International Journal of Computer Vision, vol. 40, no. 1, pages 7–23, 2000. (Cit  en page 54.)
- [Dey 2002] N. Dey, A. Boucher et M. Thonnat. Image formation model of a 3D translucent object observed in light microscopy. volume 2, pages 469–472, 2002. (Cit  en page 153.)
- [Dimitrov 2000] P. Dimitrov, C. Phillips et K. Siddiqi. Robust and Efficient Skeletal Graphs. Conf. Computer Vision and Pattern Recognition, 2000. (Cit  en page 124.)
- [Dimitrov 2003] P. Dimitrov, J. N. Damon et K. Siddiqi. Flux Invariants for Shape. Conf. Computer Vision and Pattern Recognition, 2003. (Cit  en page 124.)
- [Dong-Qing 2004] Zhang Dong-Qing et Chang Shih-Fu. Stochastic Attributed Relational Graph Matching for Image Near-Duplicate Detection. Rapport technique 1-15, 2004. (Cit  en page 70.)
- [Duchenne 2011] O. Duchenne, F. Bach, In-So Kweon et J. Ponce. A Tensor-Based Algorithm for High-Order Graph Matching. IEEE Trans. Pattern Analysis and Machine Intelligence., vol. 33, no. Dec. 2011, pages 2383–2395, 2011. (Cit  en pages 69 et 73.)
- [Faugeras 1993] O. Faugeras, B. Hotz, M. Herve, T. Vi ville, Z. Zhang, P. Fua, E. Th ron, L. Moll, G. Berry, J. Vuillemin, P. Bertin et C. Proy. Real time correlation-based stereo : algorithm, implementations and applications. Rapport technique 1-49, INRIA Rhones-alpes, 1993. (Cit  en page 21.)
- [Filliat 2012] D. Filliat. RGBD object recognition and visual texture classification for indoor semantic mapping. pages 127–132, 2012. (Cit  en page 169.)
- [Forsyth 2002] D.A. Forsyth et J. Ponce. Computer Vision : A Modern Approach. 2002. (Cit  en pages 110 et 117.)
- [Frei 1977] W. Frei et C. Chen. Fast boundary detection : A generalization and a new algorithm. IEEE Trans.on Computer, vol. 26, no. 10, pages 988–998, 1977. (Cit  en page 53.)
- [Gabriel 1969] Ruben K. Gabriel et Robert R. Sokal. A New Statistical Approach to Geographic Variation Analysis. Systematic Zoology, vol. 18, no. 3, pages 259–278, sep 1969. (Cit  en pages 70, 74 et 82.)
- [Ge 1996] Y. Ge et J.M. Fitzpatrick. On the Generation of Skeletons from Discrete Euclidean Distance Maps. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18(11), pages 1055–1066, 1996. (Cit  en page 123.)
- [Gori 2005] M. Gori, M. Maggini et L. Sarti. Exact and Approximate Graph Matching Using Random Walks. IEEE Trans. Pattern Analysis and Machine. Intelligence, vol. 27, no. 7, pages 1100–1111, 2005. (Cit  en page 70.)
- [Grompone von Gioi 2010] R. Grompone von Gioi, J. Jakubowicz, J-M. Morel et G. Randall. LSD : A Fast Line Segment Detector with a False Detection Control. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 32, no. 4, pages 722–732, 2010. (Cit  en pages ix, 52, 53, 54, 55 et 62.)

- [Haralick 1991] R.M. Haralick, C. Lee, K. Ottenberg et M. Nölle. Analysis and solutions of the three point perspective pose estimation problem. pages 592–598, 1991. (Cité en pages 96 et 98.)
- [Hartley 2004] R. I. Hartley et A. Zisserman. Multiple View Geometry in Computer Vision. 2004. (Cité en page 60.)
- [Holz 2012] D. Holz, R. Holzer S. Bogdan.-Rusu et S. Behnke. Real-Time Plane Segmentation using RGB-D Cameras. pages 306–317, 2012. (Cité en page 169.)
- [Horaud 1993] R. Horaud et O. Monga. Vision par ordinateur, outils fondamentaux. 1993. (Cité en page 102.)
- [Huijsmans 1993] D.P. Huijsmans et N. Sebe. Extended performance graphs for cluster retrieval. IEEE Conference on Computer Vision and Pattern Recognition (CV-PR'01), feb Dec.1993. (Cité en page 45.)
- [Kapur 1985] J. N. Kapur, P. K. Sahoo et K. C. Wong. A new method for gray-level picture thresholding using the entropy of the histogram. Graph. Models Image Process., vol. 29, pages 273–285, 1985. (Cité en pages viii, 38, 40, 47 et 49.)
- [Karabagli 2011] B Karabagli, T Simon et Orteu J.-J. Détermination automatique du seuil de binarisation des modules des gradients par modélisation de leur histogramme. In Proceedings of ORASIS 2011, Praz-sur-Arly, France, pages 2–8, juin 2011. (Cité en page 38.)
- [Karabagli 2013] B. Karabagli, M.-F. Auclair-Fortier, T. Simon et J.-J. Orteu. Depth Map Estimation from a Single RGB image : Principale and Calibration. Stage à l'université de Sherbrooke, Québec, présenté à EDSys, Tarbes, 2013. (Cité en page 28.)
- [Kass 1988] M. Kass, A. Witkin et D. Terzopoulos. Snakes : Active contour models. International Journal of Computer Vision 1, pages 321–331, 1988. (Cité en pages 37 et 38.)
- [Keller 1999] P. Keller J.M.and Matsakis. Aspects of high level computer vision using fuzzy sets. In Conf. Fuzzy Systems,IEEE, volume 2, pages 847 – 852, 1999. (Cité en page 131.)
- [Kimmel 2003] R. Kimmel et AM. Bruckstein. Regularized Laplacian Zero Crossings as Optimal Edge Integrators. International Journal of Computer Vision, vol. 53, no. 3, pages 225–243, 2003. (Cité en page 38.)
- [Klette 1998] R. Klette, K. Schlüns et A. Koschan. COMPUTER VISION : Three-Dimensional Data from Images. page 392, 1998. (Cité en page 60.)
- [Kourosh 2012] Khoshelham Kourosh et Oude Elberink Sander. Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications. sensors, vol. 12, pages 1437–1454, 2012. (Cité en pages 23 et 27.)
- [Ku 2005] C.L. Ku. On improvement of the computation speed of Otsu's image thresholding. J. Electron. Imaging., vol. 14(2)023011-023011-12., 2005. (Cité en pages viii, 37, 38, 39, 40, 47 et 49.)

- [Kutulakos 1998] KN Kutulakos et JR Vallino. Calibration-free augmented reality. IEEE Transactions on Visualization and Computer Graphics, no. 4(1), pages 1–20, 1998. (Cité en page 20.)
- [Lankford 1969] P.M. Lankford. Regionalization : Theory and alternative algorithms. Geographical Analysis, vol. 1, pages 169–212, 1969. (Cité en page 74.)
- [Li 1989] H. Li et H.S. Yang. Fast and reliable image enhancement using fuzzy relaxation technique. IEEE Trans. Systems, Man and Cybernetics, vol. 19(5), pages 1276–1281, 1989. (Cité en page 131.)
- [Li 1999] Chang-Tsun Li et Lou Der-Chyuan. Edge detection based on the multiresolution Fourier transform. IEEE Workshop on Signal Processing Systems, SiPS 99., pages 686–693, 1999. (Cité en page 38.)
- [Liu 1990] Y. Liu, T.S Huang et O.D. Faugeras. Determination of camera location from 2D to 3D line and point correspondences. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pages 28–37, 1990. (Cité en pages 102 et 103.)
- [Liu 2010] M-Y. Liu, O. Tuzel, A. Veeraraghavan et R. Chellappa. Fast directional chamfer matching. In Conf. Computer Vision and Pattern Recognition. IEEE Computer Society, 2010. (Cité en page 127.)
- [Lowe 2004] G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, vol. 60, no. 2, pages 91–110, 2004. (Cité en page 69.)
- [Lu 1989] Y. Lu et R.C. Jain. Behavior of edges in scale space. IEEE T-PAMI, vol. 11(4), pages 337–356, 1989. (Cité en page 38.)
- [Maire 2009] Michael Randolph Maire. Contour Detection and Image Segmentation. PhD thesis, UNIVERSITY OF CALIFORNIA, BERKELEY,, 2009. (Cité en page 46.)
- [Matula 1980] D. W. Matula et R. R. Sokal. Properties Of Gabriel Graphs Relevant To Geographic Variation Research And The Clustering Of Points In The Plane. Geographical Analysis, vol. 12, pages 205–222, 1980. (Cité en page 74.)
- [Medasani 2001] S. Medasani et Y. Krishnapuram R. Choi. Graph Matching by Relaxation of Fuzzy Assignments. IEEE Trans. on Fuzzy Systems,, vol. 9(1), pages 173–182, 2001. (Cité en page 72.)
- [Meijster 2000] A. Meijster, J. Roerdink et H. Hesselink. A general algorithm for computing distance transforms in linear time. Mathematical Morphology and its Applications to Image and Signal Processing, pages 331–340, 2000. (Cité en page 123.)
- [Moreno-Noguer 2007] F. Moreno-Noguer, V. Lepetit et P. Fua. Accurate non-iterative $o(n)$ solution to the pnp problem. IEEE 11th International Conference on Computer Vision., pages 1–8, 14-21 Oct. 2007. (Cité en page 96.)
- [Murphy 2012] K.-P. Murphy. Machine Learning : A Probabilistic Perspective. 2012. (Cité en page 117.)

- [Nacereddine 2009] N. Nacereddine, S. Tabbone, D. Ziou et L. Hamami. L'algorithme EM et le Modèle de Mélanges de Gaussiennes Généralisées pour la Segmentation d'images. Application au contrôle des joints soudés par radiographie. Conf. Traitement et Analyse de l'Information : Méthodes et Applications, pages 217–222, 2009. (Cité en pages 112 et 117.)
- [Namuduri 1992] K. R. Namuduri, R. Mehrotra et N. Ranganathan. Edge Detection Models Based on Gabor Filters. International Conference on Image, Speech and Signal Analysis, vol. 3, 1992. (Cité en page 37.)
- [Nister 2004] D. Nister. An efficient solution to the five-point relative pose problem. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26(6), pages 756–770, 2004. (Cité en page 96.)
- [Ogniewicz 1995] R. L. Ogniewicz et O. Kùbler. Hierarchic Voronoi skeletons. Pattern Recognition, vol. 28(3), pages 343–359, 1995. (Cité en page 123.)
- [Pao 1989] Y.-H. Pao. Vague features and Vagues Decision Rules : The Fuzzy-Set Approach. 1989. (Cité en page 131.)
- [Paolo 2007] Favaro Paolo et Soatto Stefano. 3-D Shape Estimation and Image Restoration - Exploiting Defocus and Motion Blur. pages 1–249, 2007. (Cité en pages 28 et 152.)
- [Papari 2011] G. Papari, N. Petkov et W. C. Regli. Edge and line oriented contour detection : State of the art. Image and Vision Computing, vol. 29, pages 79–103, 2011. (Cité en pages viii, 37 et 38.)
- [Pechard 2009] P.-Y. Pechard. Génération de trajectoires d'usinage grande vitesse 5 axes par flanc d'outil : intégration d'un critère de fluidité. PhD thesis, ENS CACHAN, 07 Juillet 2009. (Cité en page 4.)
- [Pentland 1987] P-A. Pentland. A New Sense for Depth of Field. volume 9, pages 523–531, 1987. (Cité en page 152.)
- [Peyrucat 2013] J-F. Peyrucat, D. Torre, S. Carras, C. Corleto, L. Gonzalez et S. Durand. La mesure dimensionnelle par palpation. Magazine Contrôles, Essais, Mesures, no. 43, pages 33–41, 2013. (Cité en pages vii, 13 et 14.)
- [Pudney 1998] C Pudney. Distance-Ordered Homotopic Thinning : A Skeletonization Algorithm for 3D Digital Images. Computer Vision and Image Understanding, vol. 72, no. 3, pages 404–413, 1998. (Cité en page 123.)
- [Quan 1999] L. Quan et Z.D. Lan. Linear n-point camera pose determination. IEEE Trans. Pattern Analysis and Machine Intelligence,, vol. 21(8), pages 774–780, 1999. (Cité en page 70.)
- [Reinhard 2004] Klett Reinhard et Rosenfeld Azriel. Digital geometry : Geometric methods for digital picture analysis. The Morgan Kaufmann series in computer graphics and geometric modeling, page 656, 2004. (Cité en page 17.)
- [Richtsfeld 2012] T. Richtsfeld A.and Morwald, J. Prankl, M. Zillich et M. Vincze. Segmentation of Unknown Objects in Indoor Environments. In IEEE RSJ International Conf. Intelligent Robots and Systems (IROS), 2012. (Cité en pages 143 et 169.)

- [Roberto M. 2005] Cesar Jr. Roberto M., E. Bengoetxea, I. Bloch et P. Larrañaga. Inexact graph matching for model-based recognition : Evaluation and comparison of optimization algorithms. *Pattern Recognition*, vol. 38, no. 11, pages 2099–2113, 2005. (Cité en pages 70 et 73.)
- [Sandhu 2011] Romeil Sandhu, Samuel Dambreville, Anthony Yezzi et Allen Tannenbaum. A Nonrigid Kernel-Based Framework for 2D-3D Pose Estimation and 2D Image Segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence.*, vol. 33, no. 6, pages 1098–1115, jun 2011. (Cité en pages 69 et 103.)
- [Schmid 1997] C. Schmid et R. Mohr. Local grayvalue invariants for image retrieval. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pages 530–535, 1997. (Cité en page 69.)
- [Schneider 1994] G. Schneider, B. Heit, J. Honig et J. Bremont. Monocular depth perception by evaluation of the blur in defocused images. volume 2, pages 116–119, 1994. (Cité en page 152.)
- [Shaked 1998] D. Shaked et A. M. Bruckstein. Pruning Medial Axes. *Computer Vision and Image Understanding*, vol. 69(2), pages 156–169, 1998. (Cité en page 124.)
- [Shakhnarovich 2005] G. Shakhnarovich, T. Darrell et P. Indyk. In Nearest-Neighbor Methods in Learning And Vision : Theory And Practice, page 252, 2005. (Cité en page 83.)
- [Shi 2000] J. Shi et M. Jitendra. Normalized Cuts and Image Segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22(8), pages 888–905, 2000. (Cité en pages xii, 73, 111, 113 et 114.)
- [Simon 1994] T. Simon, M Khoudeir, B. Heit et J. Bernard. Image segmentation by setting thresholds for the magnitude of gradient. *ICECS'94-IEEE Electronics, Circuits and Systems*, pages 40–45, 1994. (Cité en pages 37 et 40.)
- [Simon 2004] C. Simon, F. Bicking et T. Simon. Depth Estimation Based On Thick Oriented Edges in Images. volume 1, pages 135–140, 2004. (Cité en page 150.)
- [Stokseth 1969] P.A. Stokseth. Proprerties of a defocused optical system. volume 59(10), pages 1314–1321, 1969. (Cité en pages xiii, 153 et 154.)
- [Subbarao 1988] M. Subbarao et G. Surya. Depth Recovery from Blurred Edges. pages 498–503, 1988. (Cité en page 152.)
- [Subbarao 1994] M. Subbarao et G. Surya. Depth From Defocus : A spatial Domain Approach. volume 13, pages 271–297, 1994. (Cité en page 160.)
- [Swets 1996] J.A. Swets. Signal detection theory and ROC analysis in psychology and diagnostics. *Collected papers*. Mahwah, NJ : Erlbaum., 1996. (Cité en page 45.)
- [Tian 2002] X. Tian, W. Hui, Z. Xionghui et R. Xueyu. Object modeling of multiple views using dual quaternion in reverse engineering. *The International Journal of Advanced Manufacturing Technology*, vol. 20, no. 7, pages 495–502, sep 2002. (Cité en page 21.)
- [Tian 2010] X. Tian, X. Zhang, K. Yamazaki et A. Hansel. A study on three-dimensional vision system for machining setup verification. *Robotics and*

- Computer-Integrated Manufacturing, vol. 26, no. 1, pages 46–55, feb 2010. (Cité en pages vii, 19 et 20.)
- [Tournier 2010] C. Tournier et et coll. Usinage à Grande Vitesse. page 400, 2010. (Cité en page 4.)
- [Toussaint 1988] G.T. Toussaint. A graph theoretical primal sketch, In G. Toussaint (editor). Computational Morphology - a Computational Geometric Approach to the Analysis of Form, pages 229–260, 1988. (Cité en page 73.)
- [Trevor 1988] D. Trevor et K. Wohn. Calculating Depth From Focus Using A Pyramid Architecture. volume 57, pages 271–297, 1988. (Cité en page 150.)
- [Tsai 1986] R.Y. Tsai. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pages 364–374, 1986. (Cité en page 70.)
- [Vasilevskiy 2002] A. Vasilevskiy et K. Siddiqi. Flux Maximizing Geometric Flows. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24(12), pages 1565–1578, 2002. (Cité en page 124.)
- [Wrobel 2001] BernhardP. Wrobel. Calibration and Orientation of Cameras in Computer Vision. volume 34, pages 7–62. 2001. (Cité en page 92.)
- [Wu 1993] Z. Wu et R. Leahy. An Optimal Graph Theoretic Approach to Data Clustering : Theory and Its Application to Image Segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, no. 11, pages 1101–1113, 1993. (Cité en pages 70 et 114.)
- [Xie 2003] W. Xie, R.P. Thompson et R. Perucchio. A topology-preserving parallel 3D thinning algorithm for extracting the curve skeleton. Pattern Recognition, vol. 36, pages 1529–1544, 2003. (Cité en page 123.)
- [Zedeh 1965] L.A. Zedeh. Fuzzy Sets. Information and Control, vol. 8, pages 338–353, 1965. (Cité en pages 131 et 133.)
- [Zhang 2000] Z. Zhang. A flexible new technique for camera calibration. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pages 1330–1334, 2000. (Cité en page 70.)
- [Zhang 2006] J. Zhang, A.C. Berg, M. Maire et J. Malik. SVM-KNN Discriminative Nearest Neighbor Classification for Visual Category Recognition. Proc. IEEE Conf. Computer Vision and Pattern Recognition,, 2006. (Cité en page 82.)
- [Zhang 2007a] J. Zhang, S. Marszalek et C. Schmid. Local Features and Kernels for Classification of Texture and Object Categories : An In-Depth Study. int'l J. Computer Vision and Pattern Recognition, no. 2, pages 213–238, 2007. (Cité en page 82.)
- [Zhang 2007b] X. Zhang, X. Tian, Y. Kazuo et F. Makoto. 3D Reconstruction and Pose Determination of the Cutting Tool from a Single View. In ISVC (2)'07, pages 377–386, 2007. (Cité en page 17.)
- [Zhang 2008] X. Zhang, W. Tsang, X. Tian, Y. Kazuo et M. Masahiko. Automatic Segmentation of the Apparent Contour for 3D Modeling of Cutting Tools from Single View. In ISVC (2)'08, pages 772–781, 2008. (Cité en page 17.)

- [Zhang 2011] X. Zhang, W-M. Tsang, Y. Kazuo et M. Masahiko. A study on automatic on-machine inspection system for 3D modeling and measurement of cutting tools. *Journal of Intelligent Manufacturing*, may 2011. (Cité en pages vii, 17, 18 et 19.)
- [Zheng 2006] Y. Zheng et D. Doermann. Robust Point Matching for Nonrigid Shapes by Preserving Local Neighborhood Structures. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pages 643–649, 2006. (Cité en pages 69 et 82.)
- [Zhou 2009] C. Zhou, S. Lin et S. K. Nayar. Coded Aperture Pairs for Depth from Defocus. In *IEEE International Conference on Computer Vision (ICCV)*, Oct 2009. (Cité en page 150.)
- [Zhuo 2011] S. Zhuo et T. Sim. Defocus Map Estimation from a Single Defocused Image. *Pattern Recognition*, vol. 44, no. 9, pages 1852–1858, 2011. (Cité en pages v, xiii, 28, 149, 150, 152, 157, 158, 159 et 160.)

Vérification automatique des montages d'usinage par vision.
Application à la sécurisation de l'usinage.

Résumé : Le terme "usinage à porte fermée", fréquemment employé par les PME de l'aéronautique et de l'automobile, désigne l'automatisation sécurisée du processus d'usinage des pièces mécaniques. Dans le cadre de notre travail, nous nous focalisons sur la vérification du montage d'usinage, avant de lancer la phase d'usinage proprement dite. Nous proposons une solution sans contact, basée sur la vision monoculaire (une caméra), permettant de reconnaître automatiquement les éléments du montage (brut à usiner, pions de positionnement, tiges de fixation, etc.), de vérifier que leur implantation réelle (réalisée par l'opérateur) est conforme au modèle 3D numérique de montage souhaité (modèle CAO), afin de prévenir tout risque de collision avec l'outil d'usinage.

Mots clés : Machine outil à commande numérique, détection de contours, segmentation d'images, estimation de pose, mise en correspondance des graphes, mise en correspondance de squelettes, système expert flou, le flou optique.

Vision-based automatic verification of machining setup. Application to machine tools safety.

Abstract : In High Speed Machining it is of key importance to avoid any collision between the machining tool and the machining setup. If the machining setup has not been assembled correctly by the operator and is not conform to the 3D CAD model sent to the machining unit, such collisions can occur. We have developed a vision system, that utilizes a single camera, to automatically check the conformity of the actual machining setup within the desired 3D CAD model, before launching the machining operation.

First, we propose a configuration of the camera within the machining setup to ensure a best acquisition of the scene.

In the aim to segmente the image in regions of interest, e.g. regions of the clamping elements and piece, based-on 3D CAD model, we realise a matching between graphes, theoretical and real graphe computed from theoretical image of 3D-CAD model and real image given by real camera. The graphs are constructed from a simple feature, such as circles and lines, that are manely present in the machining setup.

In the aim to define the regions of interest (ROI) in real image within ROI given by 3D CAD model, we project a 3D CAD model in the real image, e.g. augmented reality. To automatically check the accordance between every region defined, we propose to compute three parametres, such as skeleton to represente the form, edges to represent a geometry and Area to represent dimension. We compute a score of accordance between three parameters that will be analyzed in fuzzy system to get a decision of conformity of the clamping element within it definition given in the CAD model.

Some cases of machining setup configurations require 3D information to test the trajectory of the machine tool. To get out this situation, we have proposed a new depth from defocus based-method to compute a depth map of the scene.

Finally, we present the result of our solution and we show the feasibility and robustness of the proposed solution in differents case of machining setup.

Keywords : Machining setup, edges detection, image segmentation, pose estimation, graphs matching, skeleton matching, fuzzy logic, Depth From Defocus.
