

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

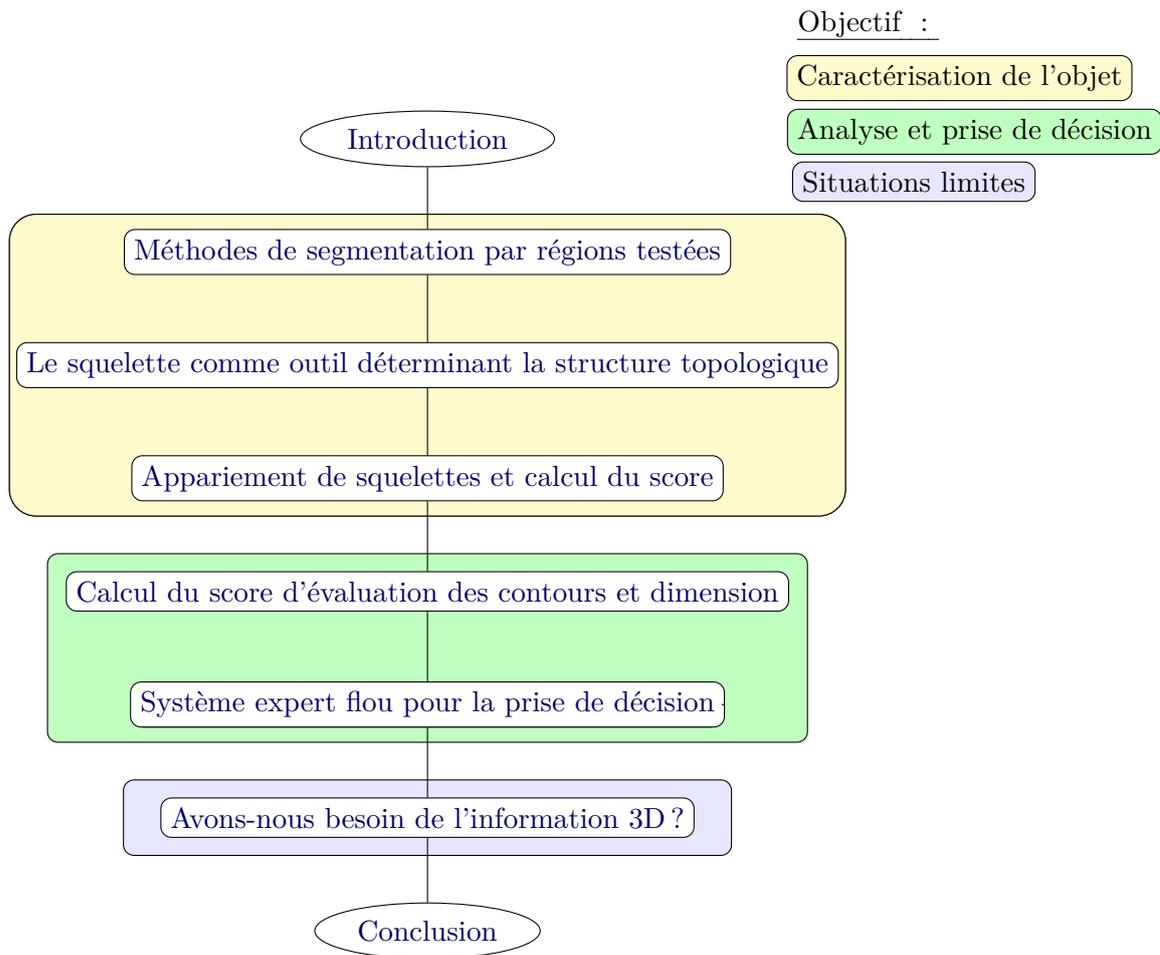


Figure 5.1 – Organigramme du chapitre 5 et objectif de chaque partie.

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

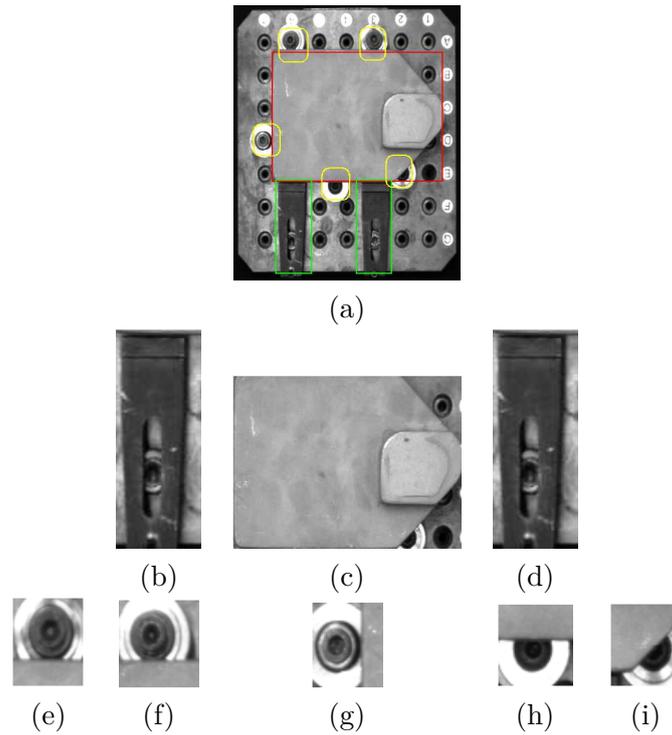


Figure 5.2 – Décomposition du montage d'usinage en régions d'intérêt des éléments de bridage pour la reconnaissance de formes. (a) Image du montage d'usinage, (b), (c), (d), (e), (f), (g), (h) et (i) imagettes des régions d : brut, Bride gauche, bride droite, pion 1, pion 2, pion 3, pion 4 et pion 5,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

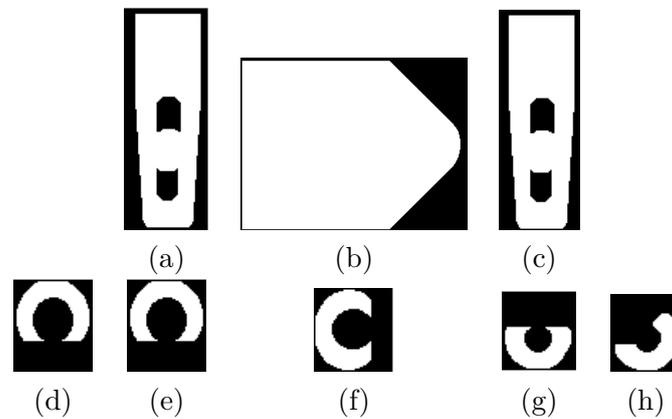


Figure 5.3 – Images binaires de référence des éléments de bridage du montage Semelle. (a) image binaire de la bride1, (b) image binaire du brut (c) image binaire de la bride2, (d) image binaire du pion1, (e) image binaire du pion2, (f) image binaire du pion3, (g) image binaire du pion4 et (h) image binaire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

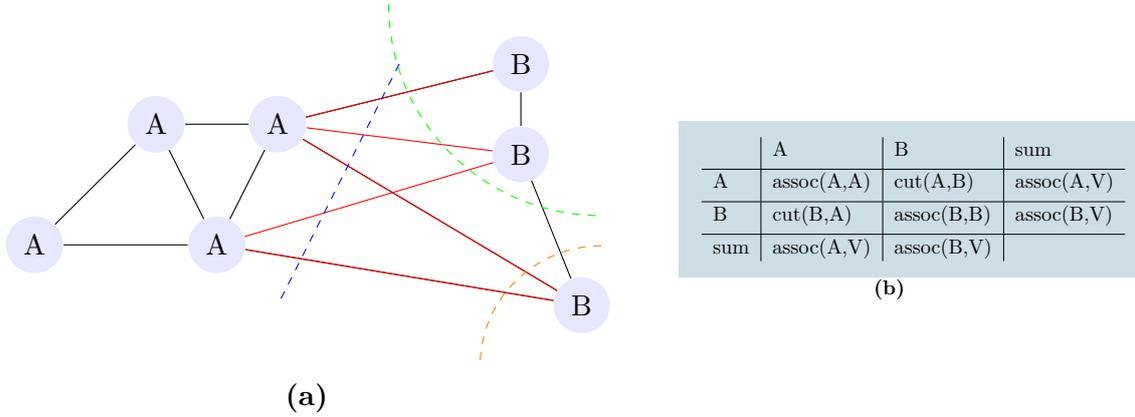


Figure 5.4 – Principe de l’algorithme Ncut [Shi 2000] : (A) un simple graphe et la coupe de poids minimale; (B) tableau des associations et des coupes pour ce graphe. *assoc* et *cut* sont calculés à partir de la matrice de poids W .

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

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.

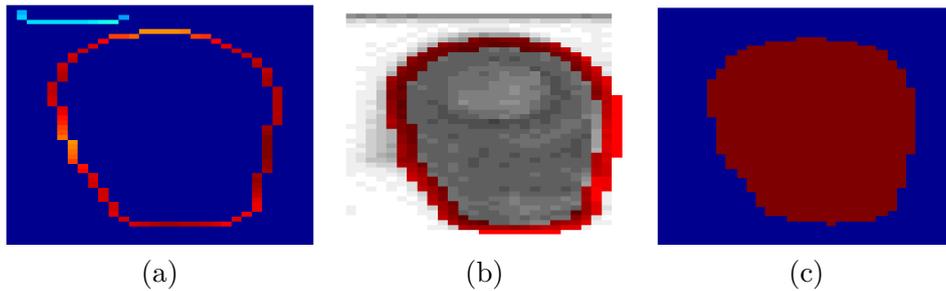


Figure 5.5 – Exemple d'application de l'algorithme *NCut* sur l'image du pion du montage *Minimanche*. (a) le degré de partitionnement calculé sur les contours, (b) les contours de partitionnement et (c) les régions de l'image.

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).

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

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*).

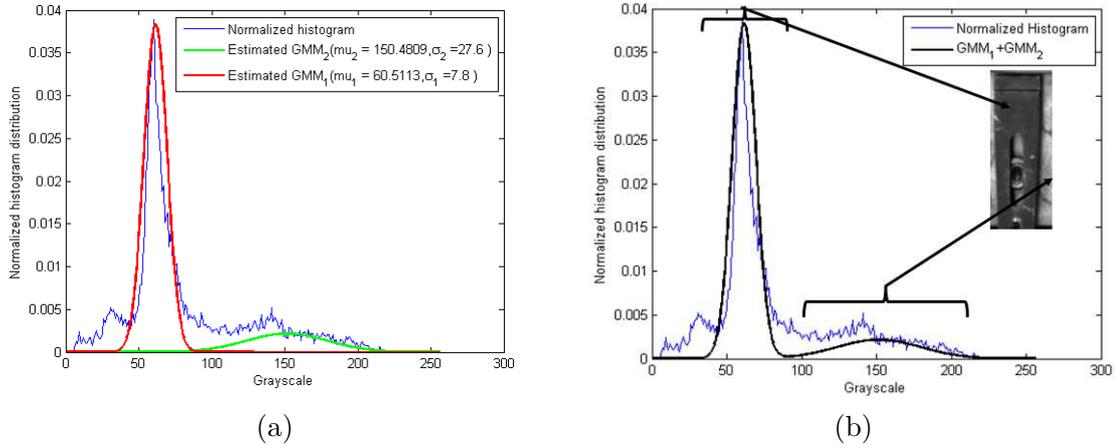


Figure 5.6 – Modélisation de l'histogramme de l'image du pion du montage Semelle par l'algorithme GMM-EM. Application à la segmentation par régions

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.

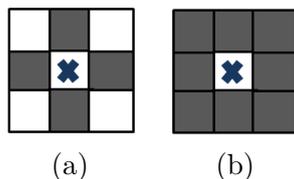


Figure 5.7 – éléments structurants utilisés dans la morphologie d'image. (a) de voisinage 4, (b) de voisinage 8.

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.

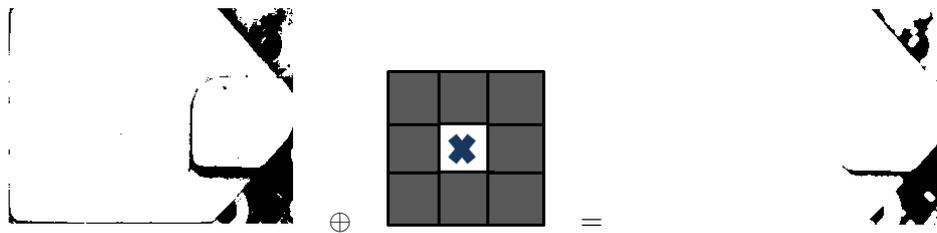


Figure 5.8 – Opération de dilatation en utilisant un filtre de voisinage 8×8 .

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.

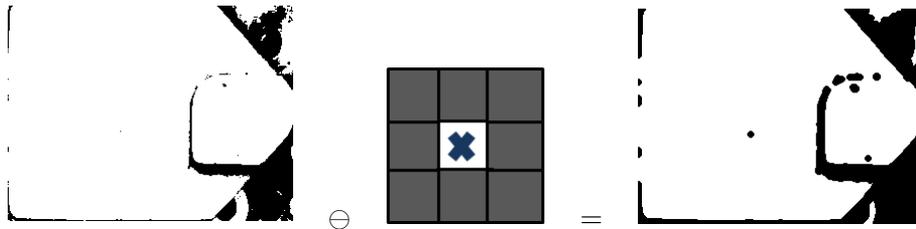


Figure 5.9 – Opération de dilatation en utilisant un filtre de voisinage 8×8 .

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

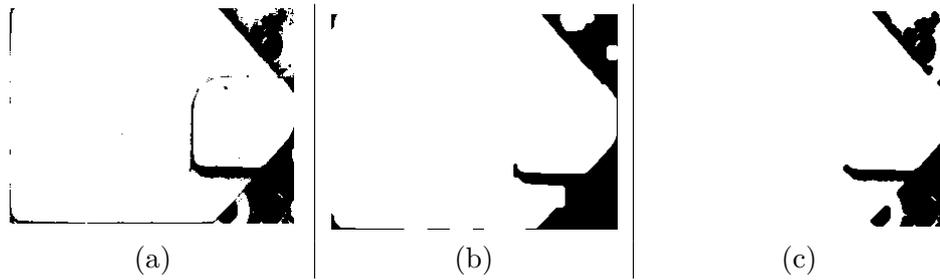


Figure 5.10 – Opérations *Opening* et *Closing* en utilisant un élément structurant de voisinage 4×4 .

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.

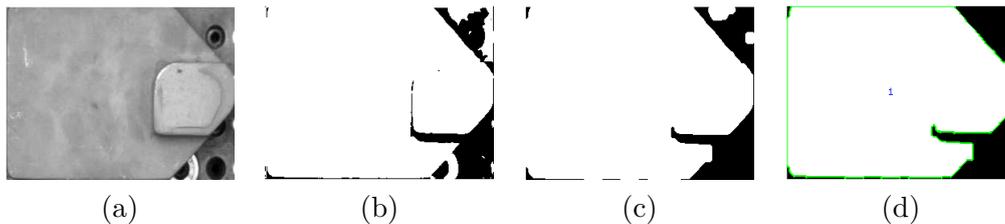


Figure 5.11 – Résultat de différentes étapes de traitement proposées pour le calcul de la forme de l'objet, (a) image de luminance de la région de brut, (b) binarisation par la méthode d'Otsu, (c) application de l'opération de morphologie mathématique, ouverture et (c) restitution de la forme de l'objet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

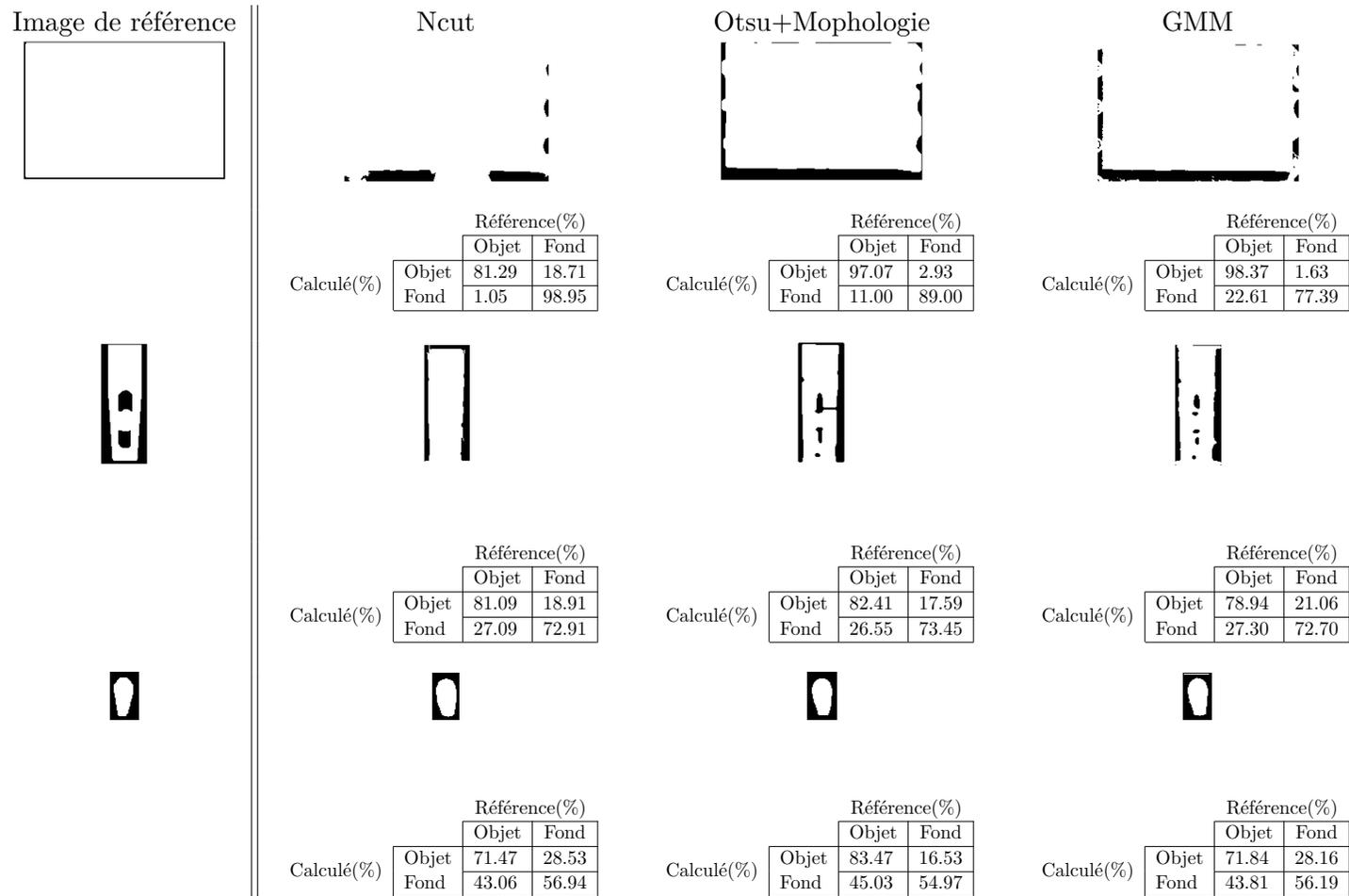


Figure 5.12 – Résultats de comparaison des méthodes de segmentation par régions testées.

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{T}), 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.

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

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.

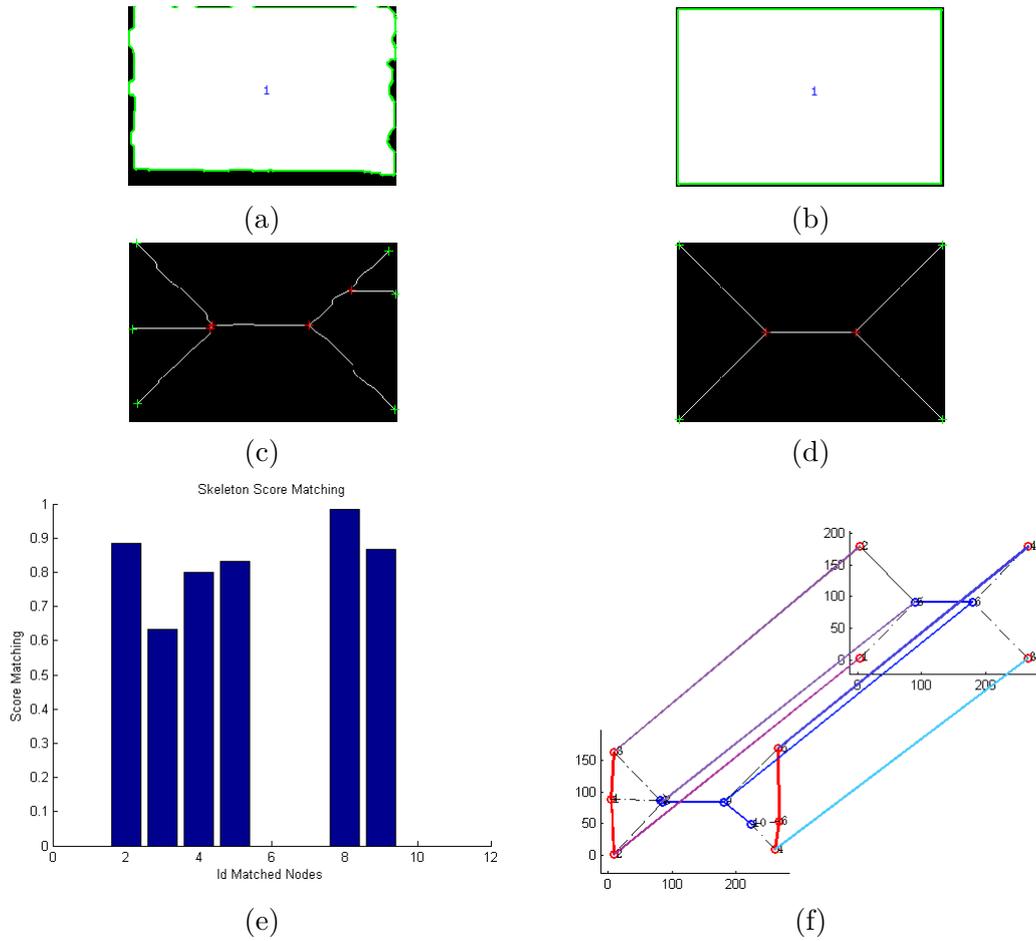


Figure 5.13 – Mise en correspondance de squelettes du brut. (a) image du brut réel calculé, (b) image de référence du brut théorique, (c) squelette correspondant à (a), (d) squelette correspondant à (b), (e) score d'appariement pour chaque nœud du graphe réel, (f) appariement des graphes

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.

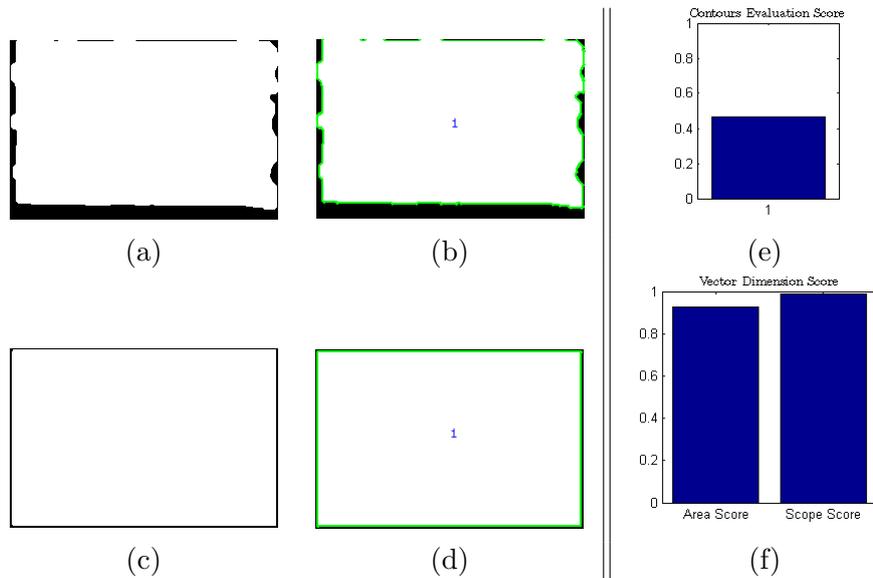


Figure 5.14 – Calcul du score d'évaluation de contours et le score de dimension. (a) image binaire du brut réel calculé, (b) image des contours du brut réel, (c) image binaire du brut de référence, (d) image des contours du brut de référence, (e) score d'évaluation de contours, (f) score de dimension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

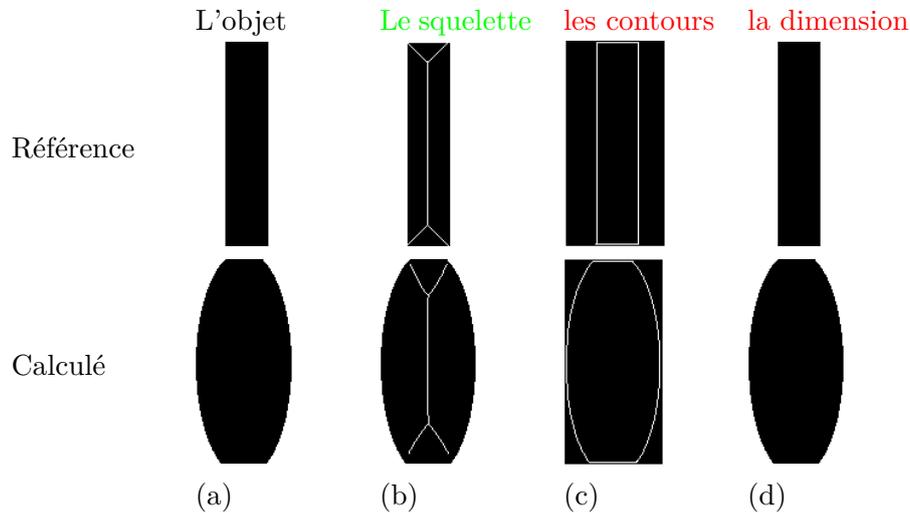


Figure 5.15 – Deux formes ayant le même squelette, mais elles n'ont pas la même dimension ni les mêmes contours.

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.

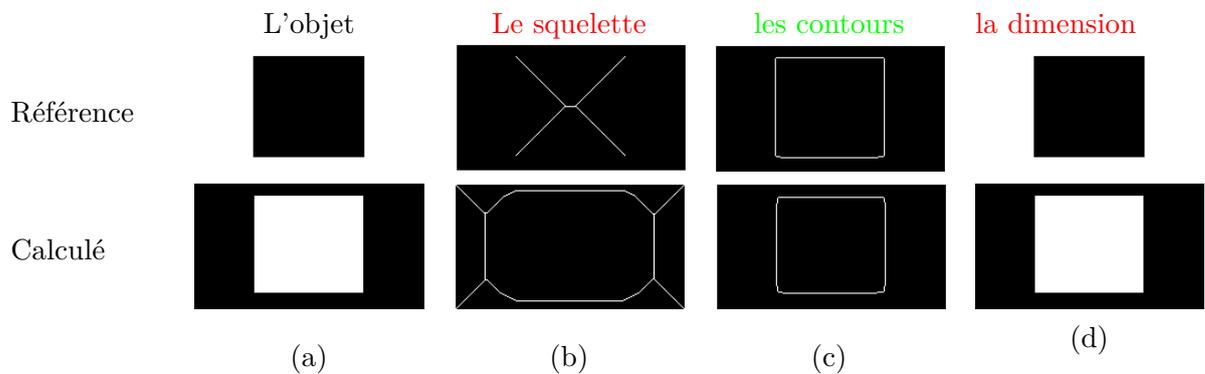


Figure 5.16 – Deux formes ayant les mêmes contours, mais elles n'ont pas la même dimension ni le même squelette.

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.

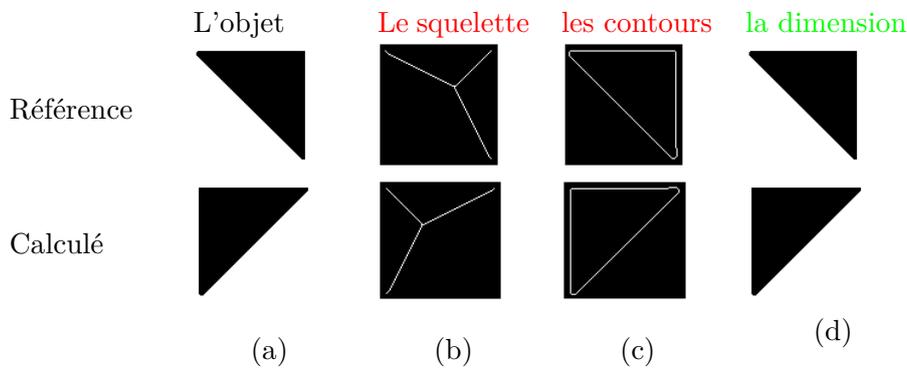


Figure 5.17 – Deux formes ayant la même dimension, mais elles n'ont pas le même squelette, ni les mêmes contours.

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.

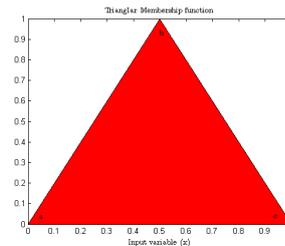
5.7.1 Définition des ensembles flous

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 :

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



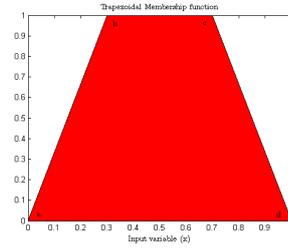
En utilisant les fonctions min et max, nous obtenons une expression de la variable x dans le triangle (a, b, c) :

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

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 :

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



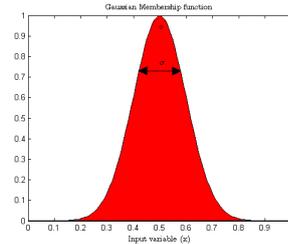
En utilisant les fonctions min et max, nous obtenons une expression de la variable x dans le trapézoïde (a, b, c, d) :

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

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.*

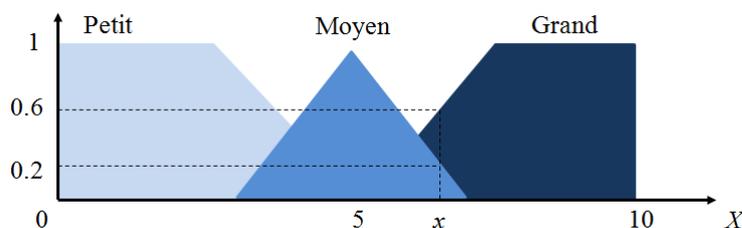


Figure 5.18 – Définition du degré d'appartenance.

- **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.

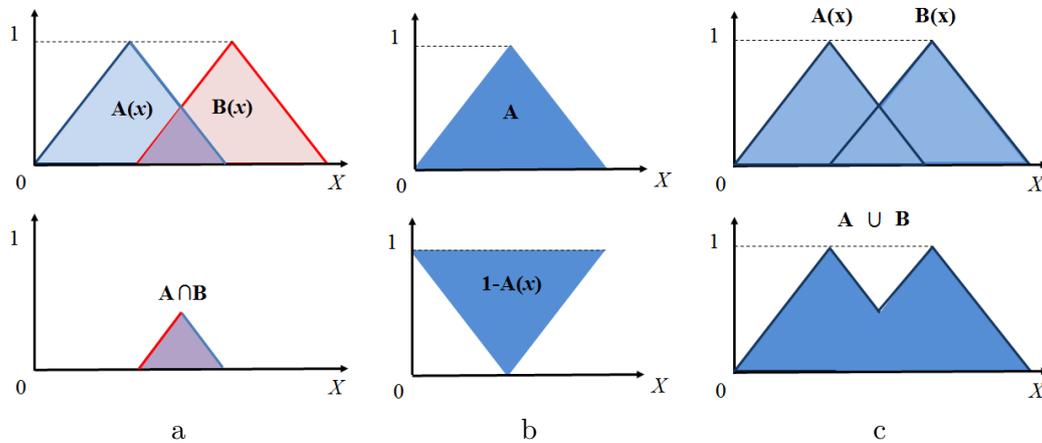


Figure 5.19 – Exemple d'opérations sur les ensembles flous : (a) intersection floue, (b) complément flou et (c) union floue

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 :

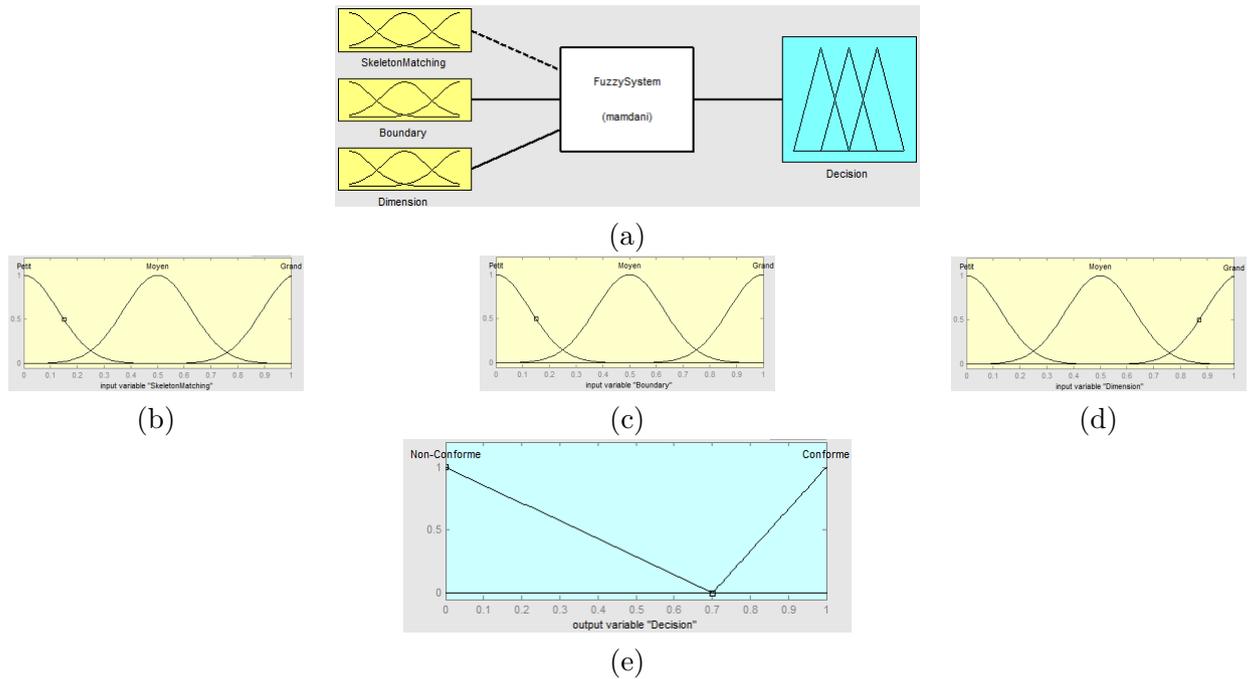


Figure 5.20 – Définition des règles floues, (a) architecture général du système expert flou, (b) règle de score d'appariement de squelettes, (c) règle de score CE, (d) règle de score de dimension et (e) règle de la décision

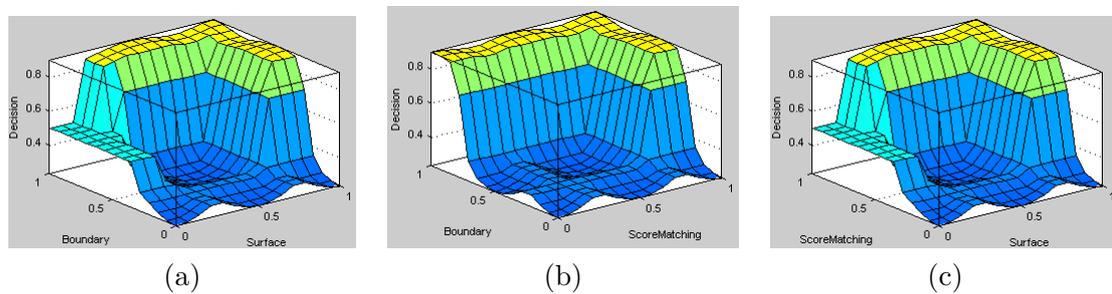


Figure 5.21 – Relation 2D entre les règles floues, (a) relation 2D entre l'ensemble flou **CE** et l'ensemble flou **Dimension**, (b) relation 2D entre l'ensemble flou **CE** et l'ensemble flou **Score d'appariement de squelettes**, (c) relation 2D entre l'ensemble flou **Dimension** et l'ensemble flou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

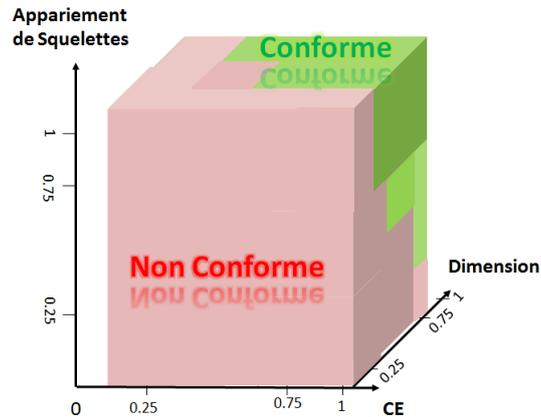


Figure 5.22 – Relation 3D entre les variables d'entrées et la variable de sortie floues

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.

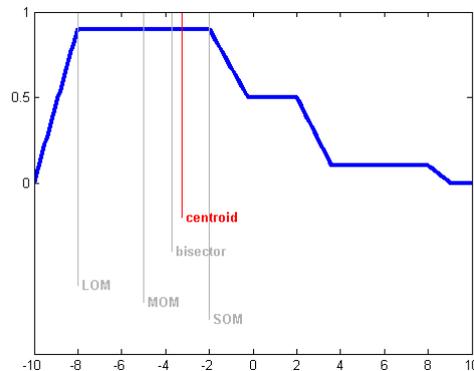


Figure 5.23 – Relation entre les méthodes de fuzzification et la forme des règles floues. Exemple de www.mathworks.com

- 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).

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

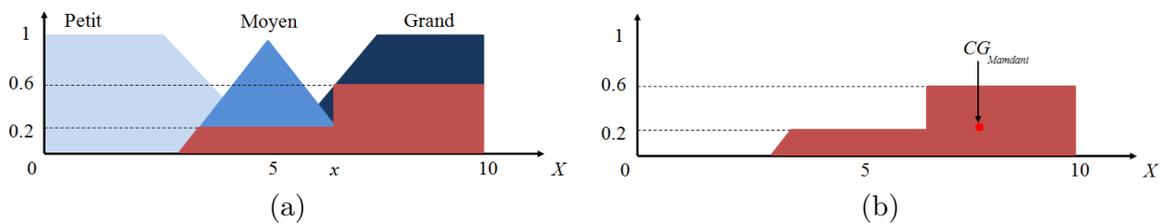


Figure 5.24 – Principe de l'inférence Mamdani et agrégation des règles de sortie, (a) calcul de taux d'appartenance, (c) inférence et calcul de la valeur de la sortie.

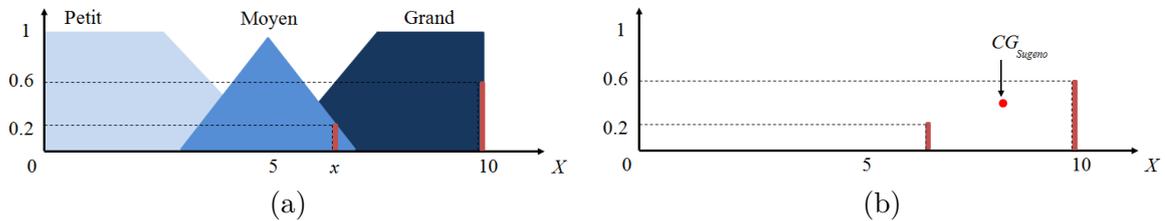


Figure 5.25 – Inférence Sugeno et agrégation des règles de sortie, (a) calcul de taux d'appartenance, (c) inférence et calcul de la valeur de la sortie.

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

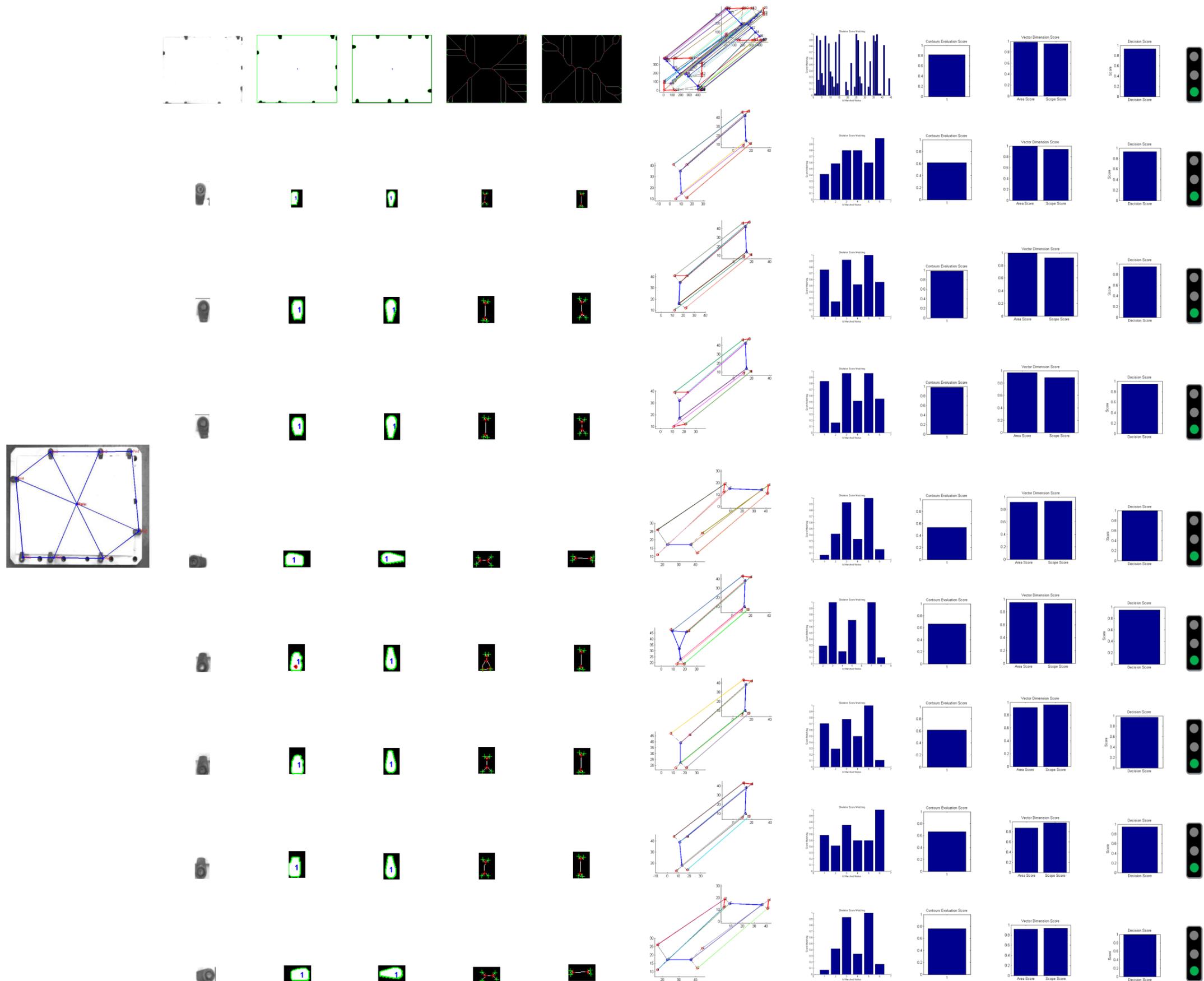
par notre méthode, e.g. Otsu et opérations morphologies pour l'objet réel, la troisième colonne l'image théorique de l'objet de référence, la quatrième et la cinquième colonne les squelettes (G) et (H) calculés sur l'image binaire réel et théorique, successivement, la sixième colonne l'appariement de 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.

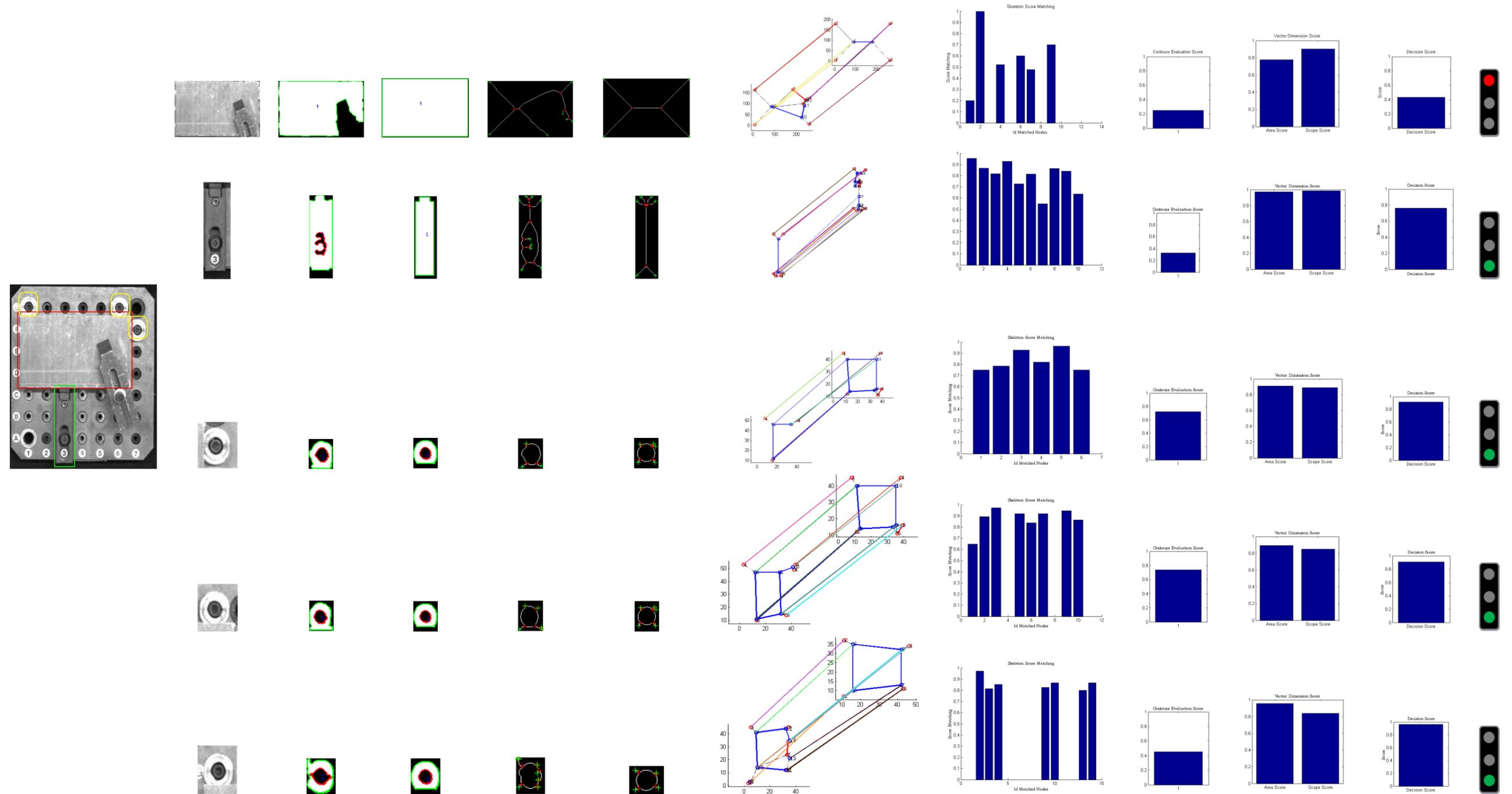
Cependant, le pion décalé n'est pas localisé! Car la définition CAO ne donne pas un pion à cet endroit. Nous nous sommes interrogés sur l'analyse de la plaque pour répondre à ce cas de figure. Si un objet est ajouté sur la plaque en dehors des régions d'intérêt, comment l'extraire à partir du reste³ 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.

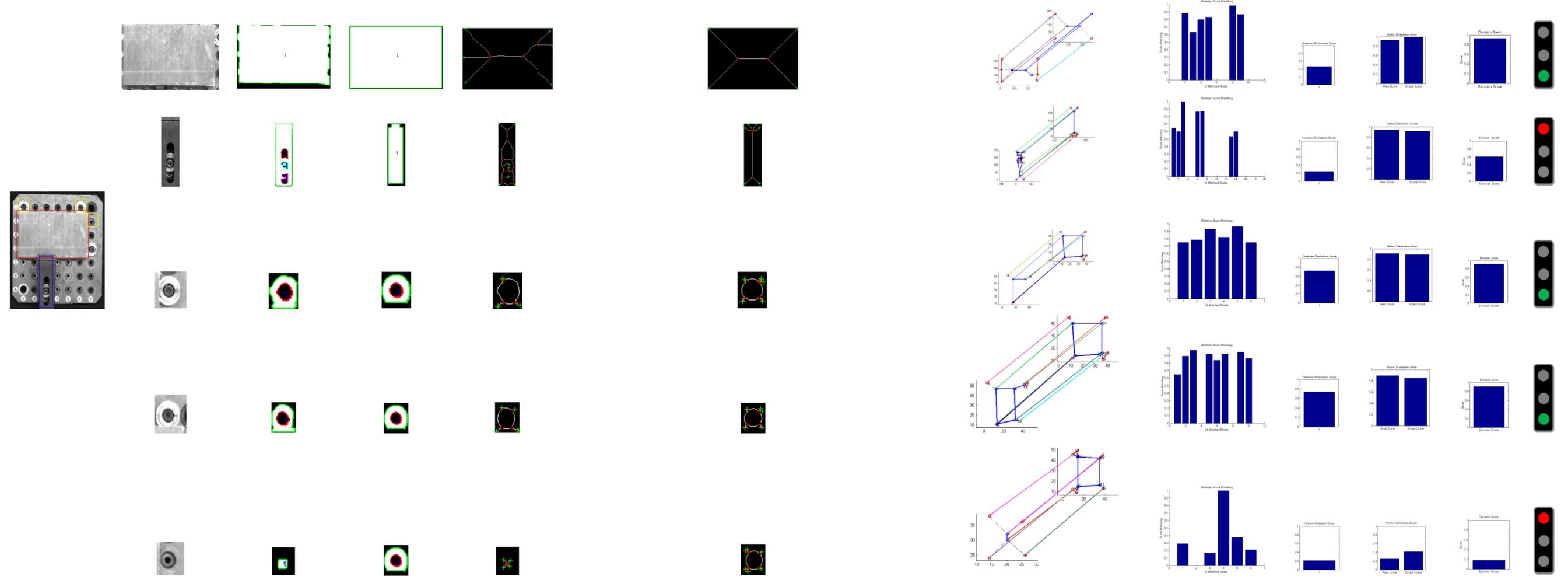
Montage Minimanche : Configuration conforme



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

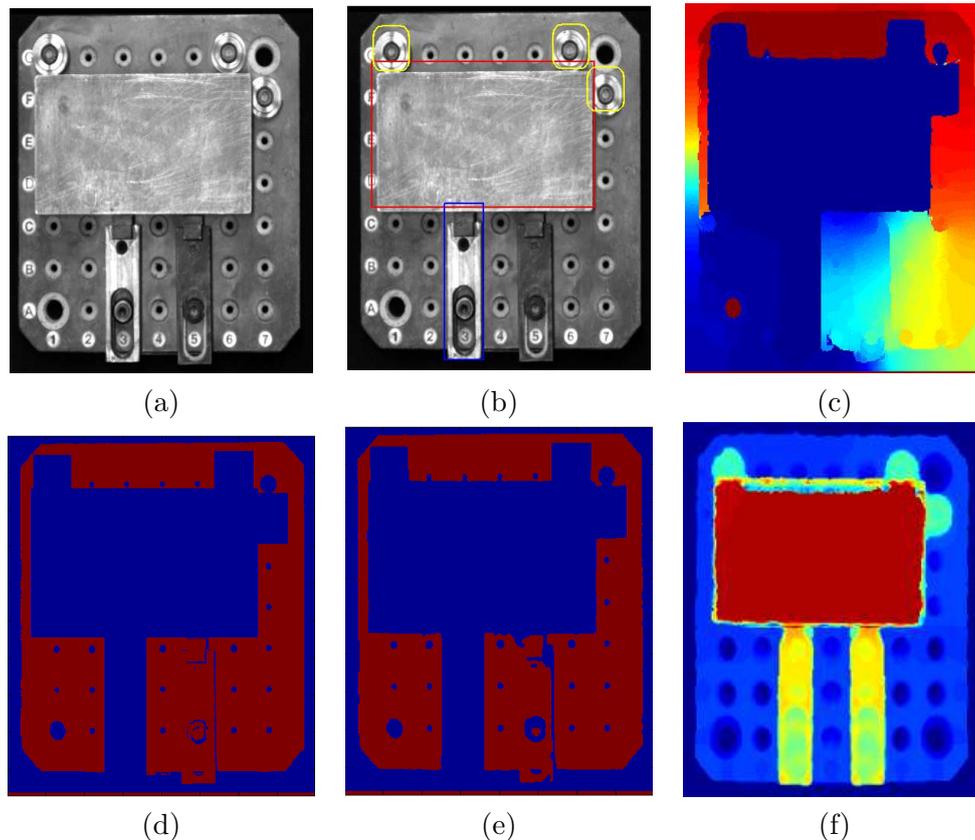


Figure 5.26 – Nécessité de l'information 3D, (a) Image du montage Norelem sans les éléments de bridage analysés, (b) localisation des régions d'intérêt, (c) Segmentation par la méthode *Ncut*, (d) Segmentation par la méthode proposée, (e) segmentation par la méthode *GMM* et (f) carte de profondeur de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

