

# Segmentation et Décodage d'un Patron de Lumière Structurée

David Fofi<sup>1</sup>, El Mustapha Mouaddib<sup>1</sup>, Joaquim Salvi<sup>2</sup>

<sup>1</sup>CREA, Université de Picardie Jules Verne, Amiens, France, david.fofi@u-picardie.fr

<sup>2</sup>IiA, Universitat de Girona, Girona, Espagne, qsalvi@eia.udg.es

**Mots-clefs** : lumière structurée, segmentation, espace des couleurs, coalescence.

## Résumé

Cet article présente une méthode de segmentation d'images issues d'un capteur de vision en lumière structurée, allant du seuillage local auto-adaptatif au décodage du motif. Ce dernier est basé sur un algorithme de coalescence initialisé grâce à la topologie de l'espace des couleurs utilisé, le CIE-Lab, proche de la perception humaine.

## Abstract

In this article, we present a segmentation method adapted structured light images, from the local self-adaptive thresholding to the pattern decoding. The latter is based on a clustering algorithm initialized thanks to the topology of the CIE-Lab colour space we use, closer to the human perception.

## 1. Introduction

La vision en lumière structurée consiste à remplacer, dans un système stéréoscopique classique, l'une des caméras par une source lumineuse projetant sur l'environnement un motif lumineux connu. Quand ce motif porte un code permettant d'identifier chacun de ses éléments, on parle de *lumière structurée et codée*. Elle est une alternative avantageuse à la stéréovision pour la mesure de scènes faiblement texturées ou produisant des images contenant un nombre limité de points d'intérêts : elle permet de structurer artificiellement la scène et de résoudre le problème de la mise en correspondance [1]. Les images issues de tels capteurs ont une morphologie particulière qui dépend étroitement du *motif structurant* (ou *patron de lumière*) projeté. Nous proposons, dans cette communication, une méthode de segmentation et de décodage d'une grille lumineuse codée par la couleur. Dans la section 2, nous rappelons brièvement la forme et le codage du motif que nous utilisons ; dans la section 3, nous décrivons la chaîne des traitements nécessaires à la segmentation et au décodage. Nous donnons quelques résultats expérimentaux en section 4, avant de conclure en section 5.

## 2. Codage du motif structurant

Salvi, Batlle et Mouaddib [5] ont proposé une grille composée de lignes horizontales et verticales régulièrement espacées et codée suivant les deux axes. Le codage est absolu et s'obtient à partir d'une unique projection, ce qui le rend utilisable pour la mesure de scènes dynamiques. Chaque ligne du motif est colorée de manière à ce que le triplet qu'elle forme avec ses deux plus proches voisins n'apparaisse qu'une fois dans l'ensemble du motif. Le codage des lignes horizontales et verticales repose sur des primitives de couleur différentes, ce qui permet de les segmenter indépendamment. Le rouge, le vert et le bleu servent à colorer les horizontales ; le cyan, le magenta et le jaune, les verticales. Les primitives ont été choisies bien espacées dans le cône TSI pour faciliter leur discrimination. Pour augmenter le nombre de points utiles du patron, il suffit d'incrémenter le nombre de primitives de couleur, au prix, bien sûr, d'un décodage plus malaisé. En utilisant six primitives, la taille maximale de la grille est de 29×29.

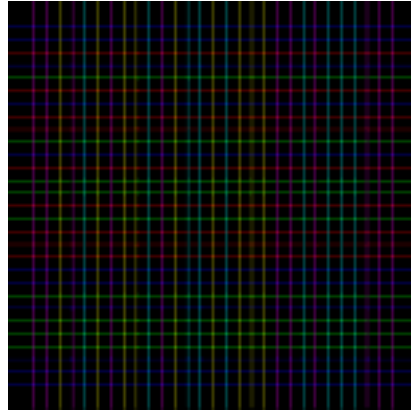


Figure 1 : Motif structurant

### 3. Segmentation du motif structurant

#### 3.1 Principe

Nous disposons d'images en lumière structurée colorées desquelles nous devons extraire les points d'intérêt (points d'intersection de la grille) et les couleurs des segments de droite qui leur servent de support. L'idée est de séparer les informations de luminance et de chrominance de l'image pour, sur la première, effectuer les traitements bas-niveau et, sur la seconde, analyser les couleurs pour décoder le motif. Le décodage des couleurs sera effectué sur le lieu des segments extraits. La Figure 2 récapitule les principales étapes de l'algorithme de segmentation.

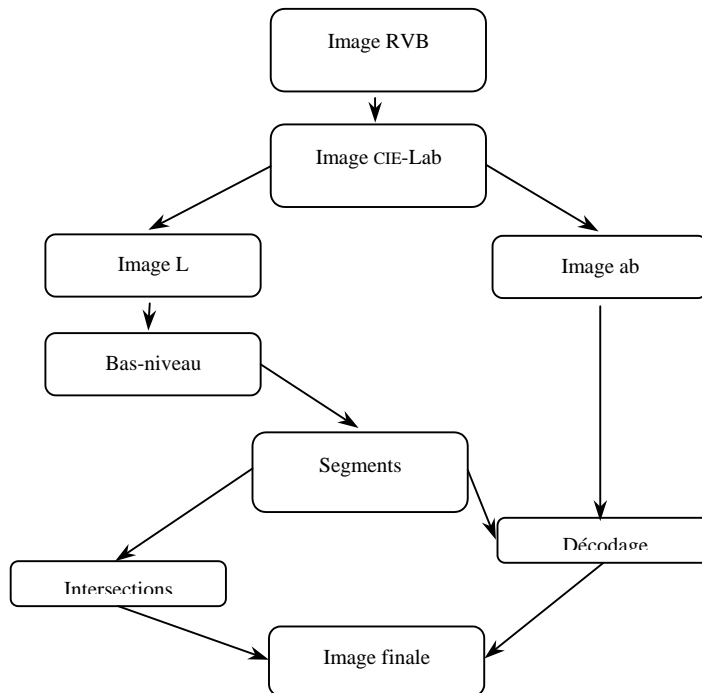


Figure 2 : Algorithme de segmentation

#### 3.2 Extraction des points d'intérêt

##### 3.2.1 Seuillage local auto-adaptatif

Considérons un pixel  $s$  de l'image, et  $V_s$  un voisinage de ce pixel. La moyenne des intensités des pixels sur ce voisinage est donnée par :

$$\bar{m}(s) = \frac{1}{\text{card}(V_s)} \cdot \sum_{i \in V_s} I(i) \quad (1)$$

Etant donné que les régions que nous désirons discriminer sont d'orientation proche de la verticale ou proche de l'horizontale, le voisinage que nous avons choisi favorise ces deux directions. En fait, deux masques sont appliqués sur l'image d'origine ; les deux images résultant de ces opérations sont ensuite additionnées. Détaillons le processus. Pour chaque pixel, la moyenne est simultanément calculée sur un voisinage 16×16 horizontal et vertical ; notons  $\bar{m}_H$  la moyenne horizontale et  $\bar{m}_V$  la moyenne verticale. Nous obtenons deux images résultantes notées respectivement  $I_H$  et  $I_V$ , avec :

$$I_H(s) = \begin{cases} 0 & \text{si } s < \bar{m}_H(s) \\ 1 & \text{si } s \geq \bar{m}_H(s) \end{cases}, \quad I_V(s) = \begin{cases} 0 & \text{si } s < \bar{m}_V(s) \\ 1 & \text{si } s \geq \bar{m}_V(s) \end{cases} \quad (2)$$

où le 1 correspond au niveau de gris maximal et le 0 au niveau de gris minimal. Pour obtenir l'image seuillée finale, nous additionnons les deux images. Une illustration des résultats du seuillage est donnée par la figure ci-dessous.

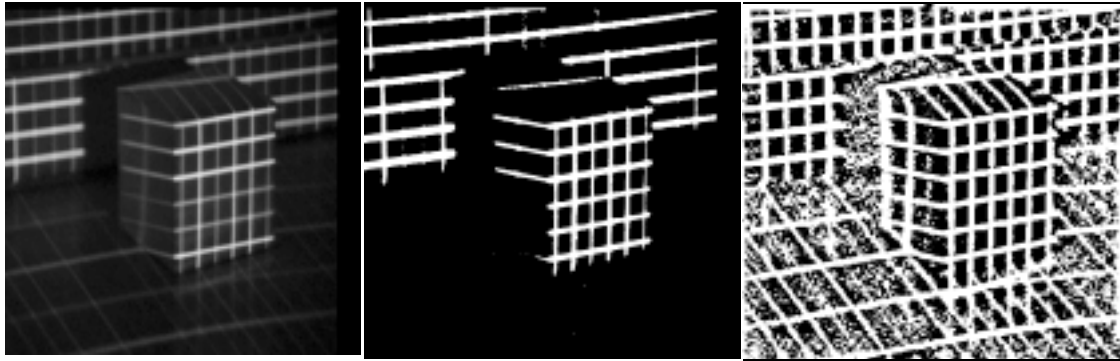


Figure 3 : Seuillage. A gauche : image originale. Au centre : seuillage global. A droite : seuillage adaptatif

### 3.2.2 Squelettisation

Le squelette d'une forme est défini comme l'ensemble des points dont la distance au point le plus proche du contour est localement maximum. A ce stade du traitement, nous disposons d'une image binaire représentant, en traits épais, les éléments du motif perçu. Le squelette de chacun de ces traits est une ligne, dont l'épaisseur est d'un pixel, situé au centre du trait épais. Une méthode bien connue de squelettisation est donnée par la *morphologie mathématique*. Cette discipline, basée sur la théorie des ensembles, offre un ensemble d'opérateurs jouant sur la structure des images [6]. Pour extraire le squelette, c'est l'opération d'*amincissement* qui sera utilisée, jusqu'à stabilisation (ou *idempotence*).

### 3.2.3 Transformation de Hough

Considérons un ensemble de points du squelette  $m_1, m_2, \dots, m_n$ , de coordonnées respectives  $(x_i, y_i)_{1 \leq i \leq n}$ . Imaginons un vecteur dont la valeur de la  $j^{\text{ème}}$  composante est égale au nombre de points pour lesquels  $\frac{y_i}{x_i} = j$ . Si une majorité de points sont alignés sur une droite de pente  $a$ , alors la  $a^{\text{ème}}$  composante devrait présenter un maximum et nous aurons détecté une ligne dans l'image, même dans le cas où les  $n$  points ne sont pas connexes. Généralisons ce résultat. Dans l'espace des paramètres  $(\rho, \theta)$ , une droite est représentée par un point. Il passe une infinité de droites d'équations par un point  $(x_i, y_i)$ . Pour un second point  $(x_j, y_j)$ , on aura une autre infinité de droites de paramètres  $(\rho, \theta)$ . Pour un ensemble de points alignés, ces droites se

couperont en un unique point de l'espace des paramètres ; ces paramètres sont ceux de la droite liant les points. L'algorithme tiré de ce raisonnement est appelé *transformation de Hough* [4].

Il convient maintenant d'étudier la particularité des images en lumière structurée pour fixer les bornes et le pas de l'espace des paramètres. On sait que les verticales projetées garderont un aspect quasi-vertical dans l'image et que les horizontales projetées garderont un aspect quasi-horizontale. De ce fait, nous scinderons l'ensemble des paramètres  $\theta$  en deux intervalles, l'un proche de zéro, l'autre proche de  $\pi/2$ . Sur la série d'images de test, il est apparu qu'une tolérance de  $\pm 10^\circ$  ( $\pm \pi/18$ ) était suffisante et ce, aussi bien pour les verticales que pour les horizontales. Le pas d'échantillonnage fixé est de  $1^\circ$ . En pratique, on utilise deux accumulateurs, l'un pour la détection des quasi-verticales, l'autre pour la détection des quasi-horizontales ; le second étant obtenu en ajoutant  $\pi/2$  aux valeurs du premier. Le paramètre  $\rho$  est choisi de sorte que l'ensemble de l'image soit balayée. Pour une image  $512 \times 512$ , par exemple, on prendra  $\rho$  allant de 0 à la diagonale, soit  $\sqrt{2} \cdot 512 \cong 724$ . Pour extraire les maxima locaux, nous passons une fenêtre  $10 \times 10$  sur les accumulateurs et nous ne gardons, au sein de cette fenêtre, que la valeur du tableau la plus élevée.

### 3.2.4 Extraction des points d'intérêt

Si nous calculions l'intersection des droites extraites en l'état, on trouverait bien évidemment des points d'intersection qui n'ont pas d'existence réelle. Il nous faut donc réduire les droites obtenues par la transformées de Hough au(x) segment(s) qui les compose(nt) avant de calculer l'intersection. Pour cela chaque droite est balayée jusqu'à ce qu'une extrémité soit rencontrée. Quelques précautions sont à prendre : puisque la transformée de Hough opère un échantillonnage sur l'orientation des droites, il faut tester la présence des extrémités sur un voisinage de deux ou trois pixels dans la direction perpendiculaire à l'orientation estimée. Pour éliminer les segments parasites provoqués par le bruit ou les barbules, on rejettera les segments dont la taille est inférieure à un seuil donné. Il suffit maintenant de calculer les points d'intersection entre les segments horizontaux et verticaux issus de la transformée de Hough pour obtenir les points d'intérêt de l'image. Il est à noter que l'algorithme de décodage sera effectué sur le lieu des segments ainsi extraits. La Figure 4 illustre les résultats obtenus.

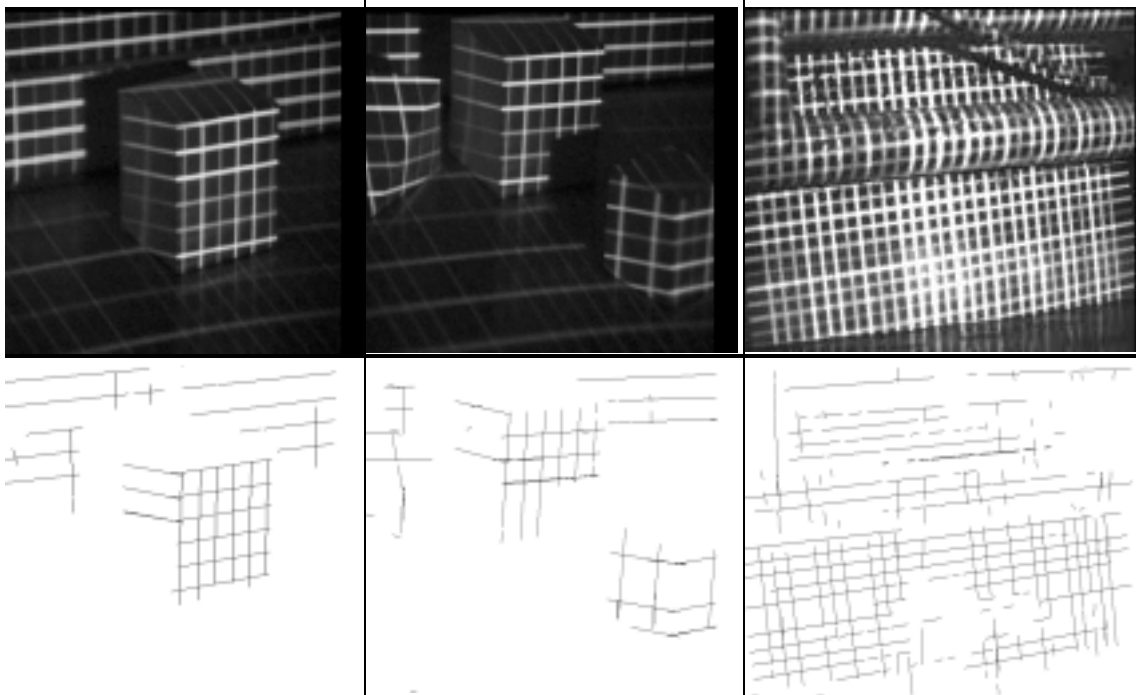


Figure 4 : Extraction des segments

### 3.3 Décodage du motif

#### 3.3.1 Conversion RVB-Lab

Pour convertir une couleur de l'espace RVB vers l'espace Lab, il faut passer par un espace intermédiaire, appelé XYZ, dont les bases théoriques résident dans les travaux de Wright [7] et Guild [3]. Ses couleurs primaires ont la particularité d'être imaginaires et définies telles que les coordonnées couleurs soient toujours positives ; la luminance de X et Z soient nulles, autrement dit Y concentre à elle seule l'information de luminance du signal ; le blanc soit représenté par un point défini par un mélange en quantités égales des trois primaires. Le passage de l'espace RVB à l'espace XYZ est donné par :S

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{bmatrix} 0.618 & 0.177 & 0.205 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.056 & 0.944 \end{bmatrix} \cdot \begin{pmatrix} R \\ V \\ B \end{pmatrix} \quad (3)$$

Il suffit maintenant d'appliquer les formules suivantes pour passer dans l'espace Lab :

$$\begin{cases} L = 116 \cdot \left(\frac{Y}{Y_0}\right)^{1/3} - 16 & \text{pour } \frac{Y}{Y_0} > 0.008856 \\ L = 903.3 \cdot \left(\frac{Y}{Y_0}\right) & \text{pour } \frac{Y}{Y_0} \leq 0.008856 \\ a = 500 \cdot \left( f\left(\frac{X}{X_0}\right) - f\left(\frac{Y}{Y_0}\right) \right) \\ b = 200 \cdot \left( f\left(\frac{Y}{Y_0}\right) - f\left(\frac{Z}{Z_0}\right) \right) \end{cases} \quad (4)$$

où  $(X_0, Y_0, Z_0)$  sont les coordonnées du blanc de référence, et la fonction  $f$  est donnée par :

$$\begin{cases} f(t) = \sqrt[3]{t} & \text{pour } t > 0.008856 \\ f(t) = 7.7787 \cdot t + \frac{16}{116} & \text{pour } t \leq 0.008856 \end{cases} \quad (5)$$

#### 3.3.2 Décodage par coalescence

Des mesures ont été effectuées sur une dizaine d'images où les composantes  $a$  et  $b$  des primitives de couleur ont été calculées. Nous avons ainsi cherché à vérifier le pouvoir discriminant d'un tel espace. La Figure 5 présente les résultats de mesures effectuées sur une seule de ces images.

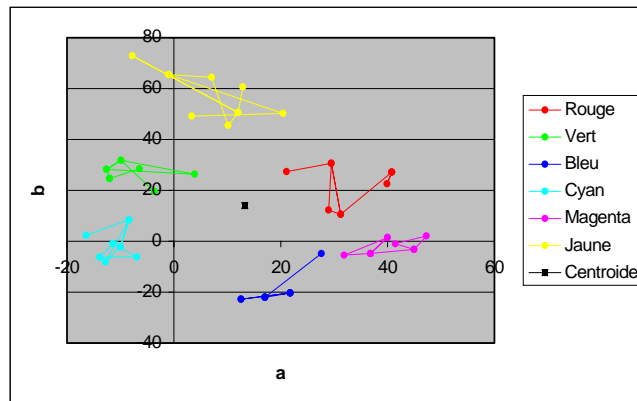


Figure 5 : Classification des primitives de couleur

Une première constatation s'impose : les classes, si elles sont bien discriminées, sont très éloignées des mesures de référence théoriques ; elles semblent toutes se rapprocher du point

d'achromatisme (0, 0), ce qui réduit leur pouvoir discriminant. Sur cet exemple, les coordonnées du barycentre des points sont approximativement (13, 14). En revanche, le positionnement relatif des six classes est respecté : le jaune est bien la classe dont la composante  $b$  est la plus forte, le bleu dont la composante  $b$  est la plus faible, le vert se trouve bien entre le jaune et le cyan, etc. L'idée est d'utiliser un algorithme de coalescence [2] pour partitionner les mesures en six classes représentant chacune une primitive de couleur et d'utiliser le positionnement relatif pour initialiser l'algorithme.

On a pu constater avec ce type d'algorithme que le partitionnement final dépend de l'initialisation des centres de classe. Au lieu de choisir arbitrairement  $n$  points parmi les mesures comme centres de classe initiaux, nous utilisons le positionnement relatif des six classes, connu *a priori*, pour déduire les coordonnées des six centres de classe. On calcule tout d'abord le centre de gravité de tous les points mesurés dans l'espace  $(a, b)$  ; notons  $\bar{m}$  ce point. On extrait ensuite des mesures le point dont la composante  $b$  est la plus élevée ; on suppose que ce point est membre de la classe jaune. On calcule l'écart chromatique  $d$  entre ces deux points. Puisque l'on connaît les valeurs théoriques des six primitives de couleur, on va placer le centre  $m_i$  de classe  $i$  de la manière suivante :

$$m_i = \frac{d}{\|a_i^2 + b_i^2\|} \cdot \begin{pmatrix} a_i \\ b_i \end{pmatrix} + \bar{m} \quad (6)$$

où  $a_i$  et  $b_i$  représente les valeurs théoriques. La première classe est initialisée dans la direction théorique du rouge, la deuxième dans la direction théorique du bleu, etc. Grâce à l'initialisation, à la connaissance *a priori* du positionnement relatif des classes et à l'algorithme de coalescence, on obtient directement la classification des points mesurés dans les six classes représentant chacune une primitive de couleur. Cet algorithme peut être affiné en deux points : il est possible d'éliminer les mauvaises classifications en prenant en compte le fait que tous les points appartenant à un même élément de motif (vertical ou horizontal) sont de même couleur ; si l'on connaît *a priori* la géométrie du système, il est possible de partitionner indépendamment les primitives utilisées pour le codage des verticales et des horizontales.

## 4. Résultats expérimentaux

### 4.1 Segmentation de l'image

L'ensemble du processus de traitement des images en lumière structurée a été implémenté. Nous avons voulu comparer les résultats obtenus par la méthode que nous proposons avec des résultats obtenus par extraction manuelle des points d'intersection et aussi, avec des résultats obtenus par corrélation d'une imagerie de référence. Nous avons testé notre algorithme sur quatre images d'une grille 5×5 prises sous différents angles et à différentes distances. Pour la segmentation manuelle, nous avons simplement cliqué sur le point qui nous semblait correspondre au point d'intersection ; la mise en correspondance d'imagerie a été réalisée en utilisant la fonction de *pattern recognition* du logiciel Matrox Inspector<sup>®</sup> : une intersection du motif est utilisée comme modèle (imagerie 20×20) et balaye l'image, l'indice de ressemblance est fixé à 60% et l'imagerie de référence peut subir une rotation de ±5°. Les quatre coins du motif, à cause de leur forme particulière, ne sont jamais reconnus par cette méthode. En outre, sur des images moins polygonales, où les intersections apparaissent avec de plus grandes distorsions, cet algorithme offre un taux de reconnaissance très faible.

Nous désignons la méthode que nous proposons par **LS**, la segmentation manuelle par **Manu** et la corrélation d'imageries par **Insp**. D'une manière générale, on peut constater que les écarts entre les trois méthodes comparées ici sont relativement faibles ; notre méthode ne s'écartant pas plus d'un pixel des deux autres, statistiquement. Il est illusoire, bien entendu, de s'attacher à l'une ou l'autre des méthodes et prétendre qu'elle fournit une mesure "absolue" ou même "plus précise". Cependant, ces tests nous ont permis de mesurer la dérive de notre méthode par rapport à une méthode déjà connue (**Insp**) et à une autre qui possède l'avantage d'être robuste au bruit, aux déformations et aux changements d'illumination, nous donnant une mesure certes

difficilement analysable, mais toujours "proche de la vérité" (**Manu**). Les résultats obtenus, sur des scènes polygonales, valident notre méthode en termes d'efficacité et de précision.

	LS-Manu		LS-Insp		Manu-Insp	
<b>Moyenne</b>	-0,09	-0,53	0,73	0,56	0,76	1,02
<b>Ecart-type</b>	2,69	1,41	2,61	1,26	0,77	1,17
<b>Distance moyenne</b>	0,54		0,92		1,27	

Table 1 : Comparaison des méthodes de segmentation. Statistiques globales.

## 4.2 Décodage du motif

Pour ces résultats, nous avons choisi des points appartenant aux segments extraits, en évitant les points d'intersection. Nous avons mesuré dix points différents par primitive de couleur ; les points ont été choisis de manière à couvrir toute l'étendue de l'image et le plus d'éléments de motif possible. Sur la Figure 6 (en haut), les soixante points mesurés ont été classés convenablement. Le mouvement des centres de classe nous montre que l'initialisation permet d'aller du "plus" discriminant (plus on s'éloigne de l'origine, plus les couleurs peuvent être discriminées facilement) vers le "moins" discriminant ; c'est un avantage puisque le mouvement inverse serait susceptible de faire dévier les membres d'une classe vers une classe voisine. On constate également que les centres de classe restent à tout moment assez éloignés les uns des autres. Sur le second exemple au bas de la Figure 6, on constate, à l'œil, que le bleu et le magenta sont difficilement discriminables et qu'il en est de même pour le vert et le jaune. Les résultats donnés par l'algorithme corroborent cette impression visuelle. Si la classe rouge est entièrement reconnue, deux points verts sont classés parmi les jaunes, huit points magentas passent chez les bleus et deux points magentas chez les cyans. En analysant le mouvement des centres de classe, on voit que le centre de la classe jaune se dirige clairement vers la classe verte, ce qui explique les confusions générées par l'algorithme. D'autre part, le magenta est complètement absorbé par le bleu et n'est pas reconnu. En fait, l'algorithme donne des résultats tout à fait comparables à l'interprétation visuelle que l'on peut avoir des images. Pour rendre la reconnaissance des couleurs plus robustes, on peut décoder indépendamment les verticales et les horizontales de l'image.

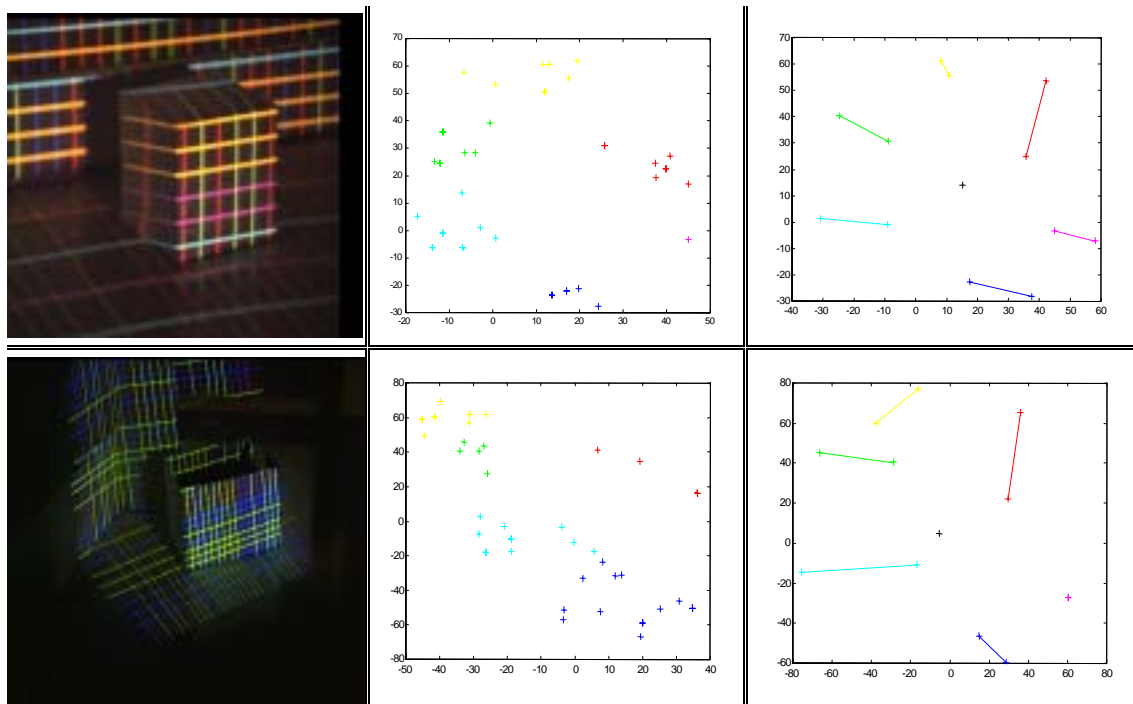


Figure 6 : Décodage. A gauche : l'image d'origine. Au centre : les points classés. A droite : le mouvement des centres de classe

## 5. Conclusions

Cet article décrit la chaîne des traitements nécessaires à la segmentation et au décodage des images en lumière structurée. Nous avons montré que la précision donnée pour l'extraction des points d'intérêt du motif était satisfaisante et qu'un décodage fiable était possible sans prise en compte du processus complet de formation de la couleur. Nous avons également proposé une méthode de seuillage dynamique originale adaptée aux images en lumière structurée. La perte d'intensité lumineuse, la différence d'albedo des surfaces, leur teinte, fait que les éléments de motif apparaissent avec beaucoup de différences d'une région à l'autre de l'image, même voisines. Un seuillage global ne permet donc pas de faire ressortir l'ensemble du motif, ainsi, une perte d'information importante altère les images. Un seuillage moyennant, sur les lignes et les colonnes de l'image, permet d'atténuer les disparités d'intensité lumineuse de l'image, et d'adapter la valeur du seuil en chaque point à un voisinage restreint. Des différences notables, entre le seuillage global et le seuillage adaptatif que nous proposons, peuvent être soulignées. Une remarque peut être faite pour l'ensemble du processus de segmentation des images : les différents seuils (accumulateur de Hough, fenêtre de filtrage de Hough, taille des segments, etc.) devront être ajustés au type d'images que l'on désire analyser et aux conditions de prise de vue.

Les différentes expérimentations effectuées montrent qu'il est possible de décoder le motif structurant à partir d'une seule image en utilisant un algorithme de coalescence classique. L'idée est d'abord de convertir l'image RVB fournie par la caméra dans un espace des couleurs proche de la perception humaine, le CIE-Lab, puis d'utiliser une connaissance *a priori* sur la position relative des classes de couleurs dans cet espace pour initialiser l'algorithme. Nous avons montré que dans certaines conditions de prise de vue, l'algorithme pouvait être mis en défaut. Il est par exemple courant que le jaune et le vert ou le bleu et le magenta soient difficilement distinguables. En théorie, chaque couleur primaire est susceptible d'être confondue avec l'une des deux couleurs secondaires qui l'avoisinent, et vice-versa. Mais, puisque les couleurs primaires sont utilisées pour coder les éléments verticaux du motif et les couleurs secondaires pour coder les éléments horizontaux, il est possible de découpler leur décodage et, ainsi, de rendre l'algorithme très robuste. Un point important de la méthode de traitement de nos images est la séparation de la segmentation, réalisée sur une image en niveaux de gris, et du décodage, réalisée sur une image des couleurs. On pourrait ainsi imaginer, dans un futur proche, un traitement des images sur une architecture parallèle de segmentation et de décodage.

## 6. Références

- [1] J. Batlle, E. Mouaddib, J. Salvi, "Recent Progress in Coded Structured Light as a Technique to Solve the Correspondence Problem: a Survey", *Pattern Recognition*, vol.31, n°7, pp. 963-982, 1998.
- [2] B. Dubuisson, *Diagnostic et reconnaissance des formes*, Editions Hermes, 1990.
- [3] J. Guild, "The colorimetric properties of the spectrum", *Philosophical Transactions of the Royal Society of London*, A230, pp. 149-187, 1931.
- [4] P.V.C. Hough, "Methods and means for recognizing complex patterns", *U.S. Patent*, 3 069 654, 1962.
- [5] J. Salvi, J. Batlle, E. Mouaddib, "A Robust-Coded Pattern Projection for Dynamic 3D Scene Measurement", *Pattern Recognition Letters*, 19(1998), pp. 1055-1065, 1998.
- [6] J. Serra, *Image Analysis and Mathematical Morphology*, vol.1, Ac. Press, London, 1982.
- [7] W.D. Wright, "A re-determination of the trichromatic coefficients of the spectral colours", *Transactions of the Optical Society*, 30, pp. 141-164, 1928.