

Reconnaissance d'objets grace a l'analyse des composantes couleur adaptées au changement d'éclairage entre deux images

Damien MUSELET, Ludovic MACAIRE, Pierre BONNET, Jack-Gérard POSTAIRE

Laboratoire d'Automatique I³D – CNRS FRE 2497
Université des Sciences et Technologies de Lille
Cité Scientifique, Bâtiment P2
59655 Villeneuve d'Ascq cedex, FRANCE
Téléphone : (33) 3 20 43 41 69
Fax : (33) 3 20 43 65 67

dm@i3d.univ-lille1.fr, ludovic.macaire@univ-lille1.fr

Résumé – Dans le domaine de l'indexation d'images, les méthodes de reconnaissance d'objets couleur ont tendance à échouer lorsque les conditions d'éclairage lors des acquisitions diffèrent d'une image à l'autre. Dans cet article, nous proposons une nouvelle approche pour la recherche d'objets dans des bases d'images couleur qui permet de s'affranchir des variations d'éclairage. Pour cela, nous considérons qu'un changement d'illuminant ne perturbe que très légèrement l'ordre des niveaux des composantes couleur des pixels d'une même image. Pour comparer deux images, nous transformons les composantes couleur de manière spécifique à chaque couple formé par une image-modèle et une image-requête. Les composantes couleur des pixels de chaque couple d'images considéré sont transformées par une analyse spécifique des mesures de rang des pixels. Des tests effectués sur une base publique d'images montrent l'amélioration obtenue par notre méthode en terme de reconnaissance d'objets.

Abstract – The main problem of most of object recognition schemes is that they tend to fail in case of illumination changes between the color image acquisitions. In this paper, we propose new color features which are invariant across illumination changes. These features, which are adapted to query-target color images, are based on a new model of illumination change. Unlike the classical invariant features which are generally based on identical transformations for all the database images, we propose to consider each pair constituted by the query and one of the target images. Then, we evaluate the adapted target color features of each pixel in the target image so that they are similar to the adapted query color features of the pixel representing the same elementary surface in the query image. Experimental tests on a public image database show the effectiveness of our approach in the context of object recognition.

1 Introduction

Dans cet article, nous traitons un problème particulier de l'indexation d'images qui consiste à rechercher des objets dans une base d'images. Chaque image de la base contient un seul objet placé sur un fond uniforme. Le système d'indexation que nous proposons a pour objectif de retrouver les images modèles contenant le même objet que celui contenu dans l'image requête. Les acquisitions des images de chaque objet respectent les conditions suivantes (1):

- les images sont acquises avec la même caméra réglée de manière identique,
- l'objet peut subir une rotation ou une translation dans un plan perpendiculaire à l'axe optique de la caméra,
- l'éclairage change d'une image à l'autre.

Nous nous focalisons sur les problèmes induits par des conditions d'éclairage non contrôlées lors des acquisitions des images de la base. Pour cela, nous nous attachons à définir des signatures aussi peu sensibles que possible aux conditions d'éclairage de manière à pouvoir les exploiter pour mesurer la similarité entre les images à comparer.

Deux approches principales permettent de s'affranchir des variations d'éclairage entre deux images. La première, dénommée "color constancy", consiste à estimer la réflectance spectrale de chaque élément de surface représenté par chaque pixel, afin d'obtenir une caractérisation de l'objet contenu dans l'image indépendante des conditions d'éclairage [1, 2]. Cette estimation ne peut être que grossière car elle est effectuée à partir des trois composantes du vecteur couleur $(c^R(P), c^V(P), c^B(P))^T$ de chaque pixel P . En effet, elles sont les résultats des intégrations des stimuli générés par l'élément de surface représenté par P dans les domaines de sensibilité spectrale des capteurs R, V et B .

C'est la raison pour laquelle nous adoptons la seconde approche qui consiste à déterminer, à partir du vecteur couleur $(c^R(P), c^V(P), c^B(P))^T$ de chaque pixel P , un vecteur d'invariants couleurs, noté $(l^R(P), l^V(P), l^B(P))^T$. Ces invariants sont des fonctions descriptives de la couleur de l'élément de surface représenté par P , indépendantes des conditions d'éclairage, sans toutefois tenter d'accéder à la réflectance spectrale de l'élément de surface considéré [3].

Comme de nombreux auteurs, la signature d'image que nous retenons est l'histogramme des vecteurs d'invariants couleur. Swain a montré que l'intersection des histogrammes couleur est une mesure de similarité pertinente, même dans le cas d'occlusion des objets présents dans les images [4]. Ainsi, la me-

sure de similarité entre deux images est l'intersection entre les histogrammes H_{req} d'une image requête I_{req} et H_{mod} d'une image modèle I_{mod} . Cette intersection est exprimée de la manière suivante:

$$Inter(H_{req}, H_{mod}) = \sum_{l^R=0}^{(n-1)} \sum_{l^V=0}^{(n-1)} \sum_{l^B=0}^{(n-1)} \min(H_{req}(l^R, l^V, l^B), H_{mod}(l^R, l^V, l^B)) \quad (1)$$

$$\times \frac{1}{\sum_{l^R=0}^{(n-1)} \sum_{l^V=0}^{(n-1)} \sum_{l^B=0}^{(n-1)} H_{mod}(l^R, l^V, l^B)},$$

où $H_{req}(l^R, l^V, l^B)$ et $H_{mod}(l^R, l^V, l^B)$ représentent les cellules de chacun des histogrammes des vecteurs d'invariants couleur. Plusieurs travaux ont comparé les résultats de reconnaissance d'objets obtenus par l'intersection d'histogrammes des vecteurs d'invariants couleur classiques [5]. Ils concluent qu'aucun des invariants couleur ne permet d'obtenir des résultats réellement satisfaisants pour la reconnaissance d'objets éclairés lors de l'acquisition de chaque image par différents illuminants.

Dans cet article, nous proposons une méthode de calcul de composantes couleur insensibles à un changement d'illuminant entre les acquisitions de deux images d'un même objet. Ces composantes, appelées "composantes couleur adaptées", sont issues d'un modèle original de transformation des niveaux des composantes couleur d'un pixel provoqué par un changement d'illuminant. Ce modèle, présenté dans la seconde partie de cet article, suppose que pour deux images contenant le même objet, le changement d'illuminant entre les acquisitions des deux images ne modifie que très légèrement l'ordre des niveaux d'une composante couleur des pixels. Pour comparer deux images qui sont acquises sous différentes conditions d'éclairage, nous proposons de ne pas suivre la démarche classique qui consiste à trouver une transformation unique qui fournit des invariants couleur insensibles à tous les changements possibles d'éclairage. Nous proposons d'adapter les niveaux des composantes couleur de manière spécifique à chaque couple formé par l'une des images modèles et l'image requête. L'algorithme de calcul des niveaux des composantes couleur adaptées des pixels de chaque couple d'images est présenté dans la troisième partie. Afin d'évaluer les performances de cette méthode, nous avons effectué des tests expérimentaux sur la base d'images publiée par l'Université de East Anglia [6] qui contient les images de 28 objets multicolores éclairés avec trois illuminants différents. Des méthodes d'indexation utilisant des invariants classiques ont été testées sur cette base et la comparaison des résultats fait l'objet de la discussion présentée dans la dernière partie de cet article.

2 Modèles de transformation de la couleur due à un changement d'éclairage

2.1 Formation spectrale de la couleur

Comme nous ne disposons d'aucune information, ni sur la formation des images, ni sur les illuminants utilisés lors de leur acquisition, nous considérons un élément de surface Lambertienne plan, perpendiculaire à l'axe optique de la caméra et éclairé de manière uniforme. Sous ces hypothèses, $c^k(P)$, le niveau de la composante couleur k , avec $k = R, V, B$, du pixel

P représentant cet élément de surface dépend de $E(\lambda)$, la distribution spectrale de l'illuminant incident, de $\beta(P, \lambda)$, la réflectance spectrale de l'élément de surface représenté par P , et de $S^k(\lambda)$, l'une des fonctions de sensibilité spectrale des trois capteurs R, V, B de la caméra:

$$c^k(P) = \int_{\lambda=380}^{\lambda=780} S^k(\lambda) \beta(P, \lambda) E(\lambda) d\lambda, \quad k = R, V, B. \quad (2)$$

Cette équation montre qu'un vecteur couleur ($c^R(P), c^V(P), c^B(P)$)^T ne peut pas être considéré comme représentant la propriété intrinsèque de l'élément de surface représenté par le pixel P .

2.2 Le modèle diagonal

Le modèle spectral de formation de la couleur étant trop complexe pour définir la relation unissant la couleur et les conditions d'éclairage, de nombreux auteurs préfèrent utiliser le modèle diagonal de transformation de la couleur par changement d'éclairage [7]. Ce modèle suppose que les capteurs, définis par $S^k(\lambda)$, ne sont sensibles que sur des intervalles étroits de longueurs d'onde. Cette hypothèse permet de simplifier l'équation (2) afin d'exprimer la transformation de la couleur des pixels de deux images acquises selon les conditions d'acquisition décrites en (1) et contenant le même objet. Ce modèle stipule que $c^k(P_{req})$, le niveau de la composante couleur k du pixel P_{req} représentant un élément de surface éclairé par l'illuminant de l'image requête, peut être exprimé à partir de $c^k(P_{mod})$, le niveau de la composante couleur k du pixel P_{mod} représentant le même élément de surface que P_{req} , mais éclairé par l'illuminant de l'image modèle par :

$$c^k(P_{req}) = \alpha^k \times c^k(P_{mod}), \quad k = R, V, B, \quad (3)$$

où α^k est un paramètre constant pour tous les pixels des deux images considérées et déterminé pour chaque composante couleur k . Les hypothèses sur lesquelles sont fondées ce modèle diagonal sont trop restrictives pour qu'il décrive sans erreur la transformation de la couleur provoquée par un changement d'éclairage [8].

2.3 Le modèle de Hordley

Hordley propose un modèle fondé sur des hypothèses moins restrictives que le modèle diagonal. Ce modèle exprime le niveau $c^k(P_{req})$ en fonction du niveau $c^k(P_{mod})$ à l'aide d'une fonction monotone croissante f^k [9]:

$$c^k(P_{req}) = f^k(c^k(P_{mod})), \quad k = R, V, B. \quad (4)$$

La fonction f^k n'étant pas forcément linéaire, ce modèle de transformation est moins restrictif et donc beaucoup mieux adapté que le modèle diagonal à une transformation de la couleur provoquée par un changement d'éclairage [9]. Ce modèle suppose que tous les pixels caractérisés par un même niveau pour la composante couleur k dans l'image modèle I_{mod} soient aussi caractérisés par un même niveau pour la composante couleur k dans l'image requête I_{req} . Cependant cette hypothèse est contredite par l'équation (2) de la formation spectrale de la couleur qui implique que les niveaux d'une composante couleur de deux pixels peuvent être égaux lorsqu'un illuminant éclaire les deux éléments de surface considérés, et peuvent être différents avec un autre illuminant.

2.4 Modèle proposé

Nous proposons un modèle dérivé du modèle de Hordley qui tient compte de ce phénomène:

$$c^k(P_{req}) = f^k(c^k(P_{mod})) + \phi^k(P_{mod}), k = R, V, B, \quad (5)$$

où, pour chaque composante couleur k , f^k est une fonction monotone croissante et ϕ^k est une fonction représentant la dispersion des niveaux qui peut être causée par la formation spectrale de la couleur décrite par l'équation (2). Considérons deux éléments de surface représentés respectivement par deux pixels P_{mod}, Q_{mod} dans l'image modèle et par P_{req}, Q_{req} dans l'image requête, les deux images contenant le même objet et étant acquises selon les conditions décrites en (1). La relation suivante est vérifiée par la fonction f^k , monotone croissante:

$$c^k(P_{mod}) \geq c^k(Q_{mod}) \Rightarrow f^k(c^k(P_{req})) \geq f^k(c^k(Q_{req})), \\ k = R, V, B. \quad (6)$$

Par contre, bien que $c^k(P_{mod})$ soit supérieur ou égal à $c^k(Q_{mod})$, nous ne pouvons pas conclure que $c^k(P_{req})$ est supérieur ou égal à $c^k(Q_{req})$, à cause de la fonction ϕ^k .

2.5 Mesure de rang des pixels

Ce modèle peut être exploité en triant les pixels d'une image I par ordre croissant de leurs niveaux pour la composante couleur k , $k = R, V, B$. On peut alors associer à chaque pixel P de l'image une mesure de rang notée $R_I^k(P)$ de telle sorte que la mesure de rang des pixels qui ont le niveau le plus faible c^k soit égale à 0, tandis que celle des pixels qui ont le niveau le plus élevé soit égale à 1. Soit $h^k(x)$ le nombre de pixels P dans l'image I dont le niveau $c^k(P)$ de la composante k est égal à x . La mesure de rang d'un pixel P dans cette image I pour la composante k est évaluée de la manière suivante:

$$R_I^k(P) = \frac{\sum_{i=0}^{c^k(P)} h^k(i)}{\sum_{i=0}^N h^k(i)}, \quad (7)$$

où N indique le nombre de niveaux avec lequel la composante k est quantifiée. Soit le couple de pixels (P_{req}, P_{mod}) qui représentent respectivement le même élément de surface dans les deux images I_{req} et I_{mod} , ces deux images contenant le même objet et étant acquises selon les conditions décrites en (1). Alors que le modèle de Hordley suppose que:

$$R_{req}^k(P_{req}) = R_{mod}^k(P_{mod}), \quad (8)$$

notre modèle est basé sur l'hypothèse moins restrictive suivante:

$$|R_{req}^k(P_{req}) - R_{mod}^k(P_{mod})| = \min_{P \in I_{req}} |R_{req}^k(P) - R_{mod}^k(P_{mod})|. \quad (9)$$

Ainsi, le modèle proposé de transformation de la couleur due à un changement d'éclairage suppose que pour deux images contenant le même objet et acquises selon les conditions décrites en (1), la mesure de rang d'un pixel dans I_{req} est la plus proche de la mesure de rang du pixel représentant le même élément de surface dans I_{mod} . Nous en déduisons que l'ordre des niveaux d'une composante couleur entre deux pixels d'une même image est conservé, à un biais près, dans l'autre image, et ce quels que soient les illuminants utilisés lors des acquisitions des images.

3 Composantes couleur adaptées

Afin de comparer les contenus des images du couple formé par l'une des images modèles et l'image requête, nous proposons d'éliminer les conséquences du changement d'illuminant entre ces deux images. Les niveaux des composantes couleur $c^k(P_{req})$ et $c^k(P_{mod})$ des pixels P_{req} dans l'image requête et P_{mod} dans l'image modèle I_{mod} sont respectivement transformés en niveaux des composantes couleur adaptées $l_{adapt_req}^k(P_{req})$ et $l_{adapt_mod}^k(P_{mod})$. Pour éliminer l'effet du changement d'illuminant entre les deux images, nous proposons de transformer les niveaux des composantes couleur des pixels de telle sorte que $l_{adapt_mod}^k(P_{mod})$ et $l_{adapt_req}^k(P_{req})$ soient égaux lorsque les pixels P_{mod} et P_{req} représentent le même élément de surface respectivement dans les images modèle et requête. Comme les objets peuvent être déplacés d'une image à l'autre, le principal problème est de mettre en correspondance chaque pixel P_{mod} de l'image modèle avec le pixel P_{req} de l'image requête qui représente le même élément de surface que P_{mod} . Nous supposons que le changement d'illuminant ne perturbe que très légèrement les mesures de rang des pixels P_{req} et P_{mod} pour chaque composante couleur. Ainsi, pour chaque composante couleur k , le pixel P_{mod} de l'image modèle est mis en correspondance avec le pixel P_{req} de l'image requête dont la mesure de rang dans l'image requête est la plus proche de celle de P_{mod} dans l'image modèle. En généralisant cette règle à tous les pixels du couple d'images considéré, le calcul de chaque composante couleur adaptée analyse chaque composante couleur k des deux images. Elle s'exécute selon les trois étapes successives suivantes afin que pour chaque composante couleur d'un couple d'images, les histogrammes des niveaux des composantes adaptées soient les plus semblables possibles.

La première étape détermine les niveaux de la composante couleur adaptée $l_{adapt_mod}^k$ des pixels de l'image modèle. Elle recherche pour chaque pixel P_{mod} de l'image modèle, le pixel correspondant P_{req} de l'image requête qui minimise l'écart entre les rangs:

$$|R_{req}^k(P_{req}) - R_{tar}^k(P_{mod})| = \min_{P \in I_{req}} |R_{req}^k(P) - R_{mod}^k(P_{mod})|. \quad (10)$$

Le niveau $l_{adapt_mod}^k(P_{mod})$ est alors rendu égal à $c^k(P_{req})$:

$$l_{adapt_mod}^k(P_{mod}) = c^k(P_{req}). \quad (11)$$

Lors de la seconde étape, les niveaux de chaque composante couleur adaptée $l_{adapt_req}^k(P_{req})$ des pixels P_{req} de l'image requête qui ont été mis en correspondance lors de la première étape, sont rendus égaux aux niveaux $c^k(P_{req})$ de la composante couleur k . Ainsi, les niveaux de chaque composante couleur adaptée des couples de pixels mis en correspondance lors de l'étape 1 sont égaux.

Enfin, la troisième étape détermine les niveaux $l_{adapt_req}^k(P'_{req})$ de chaque composante couleur adaptée des pixels P'_{req} de l'image requête qui n'ont pas été mis en correspondance lors de la première étape. Le niveau $l_{adapt_req}^k(P'_{req})$ d'un tel pixel P'_{req} est rendu égal au niveau du pixel mis en correspondance lors de la première étape, immédiatement supérieur à $c^k(P'_{req})$.

4 Résultats

Afin d'évaluer les performances de cette méthode, nous avons effectué des tests expérimentaux sur la base d'images publiée par l'Université de East Anglia [6] qui contient les images acquises selon les conditions décrites en (1) de 28 objets multicolores éclairés avec l'un des trois illuminants différents. La figure 1 présente six images de deux objets éclairés par les trois illuminants disponibles. Le tableau 1 montre les résul-

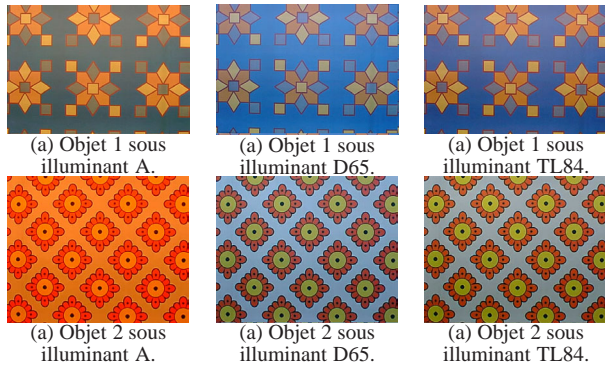


FIG. 1: Deux objets éclairés par l'un des trois illuminants différents.

tats de reconnaissance d'objets sur cette base d'images obtenus par la méthode d'indexation utilisant soit l'un des deux vecteurs d'invariants couleur classiques retenus pour leur efficacité dans [5, 9], à savoir les invariants par Greyworld [10] et les invariants par égalisation d'histogrammes [9], soit le vecteur des composantes couleur adaptées de chaque pixel. La valeur contenue dans la case correspondant au rang i indique le pourcentage de recherches d'images où l'image modèle qui représente le même objet que l'image requête a obtenu au moins le i^{me} rang du classement des images modèles. Ainsi, l'exploitation d'un invariant permet d'obtenir une recherche parfaite si la valeur correspondant au rang 1 est égale à 100%. La comparaison des résultats indique que la méthode d'indexation utilisant les composantes couleur adaptées obtient les meilleurs résultats pour la reconnaissance des 28 objets éclairés lors de l'acquisition de chaque image par l'un des trois illuminants disponibles. Deux raisons principales expliquent l'amélioration des résul-

Vecteurs d'invariants testés	1 ^{er} rang	≤ 2 ^{eme} rang	≤ 3 ^{eme} rang
Invariants couleur par Greyworld [10]	67.26	77.70	82.59
Invariants par égalisation d'histogramme [9]	86.01	90.92	92.71
Composantes couleur adaptées	89.29	94.35	95.83

TAB. 1: Résultats de reconnaissance d'objets obtenus par 3 vecteurs d'invariants couleur avec la base d'images de l'Université de East Anglia.

tats de reconnaissance d'objets. D'une part, les invariants sont évalués pour chaque couple d'image modèle - image requête afin de déterminer les effets provoqués par des changements d'éclairage entre les acquisitions. D'autre part, le modèle pro-

posé, sur lequel sont basées les composantes adaptées respecte le modèle spectral de formation de la couleur.

Références

- [1] B. Funt, V. C. Cardei, and K. Barnard, "Method of estimating chromaticity of illumination using neural networks," in *United States Patent*, USA, 1999, vol. 5,907,629.
- [2] G. D. Finlayson, S. D. Hordley, and P. M. Hubel, "Color by correlation: a simple, unifying framework for color constancy," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23(11), pp. 1209–1221, 2001.
- [3] B. Funt, F. Ciurea, and J. McCann, "Retinex in matlab," in *Eighth Color Imaging Conf.*, Scottsdale USA, 2000, pp. 112–121.
- [4] M. J. Swain and D. H. Ballard, "Color indexing," *Int. Jour. of Computer Vision*, vol. 7(1), pp. 11–32, 1991.
- [5] G. D. Finlayson and G. Schaefer, "Colour indexing across devices and viewing conditions," in *Proceedings of 2nd International Workshop on Content-based Multi-Media Indexing (CBM01)*, Brescia, Italy, 2001.
- [6] "http://vision.doc.ntu.ac.uk/research/catsi/database.html," .
- [7] G. Finlayson, M.S. Drew, and B. Funt, "Diagonal transforms suffice for color constancy," *IEEE Proceedings: Int. Conf. on Computer Vision*, pp. 164–171, 1993.
- [8] B.V. Funt, K. Barnard, and L. Martin, "Is machine colour constancy good enough?," in *Proceedings of the 5th European Conference on Computer Vision*, 1998, pp. 445–459.
- [9] S.D. Hordley, G.D. Finlayson, G. Schaefer, and G. Y. Tian, *Illuminant and Device Invariant Colour Using Histogram Equalisation*, Technical report sys-c02-16, School of Information Systems, University of East Anglia, Norwich, United Kingdom, 2002.
- [10] R. Gershon, A. D. Jepson, and J. K. Tsotsos, "From [r,g,b] to surface reflectance: computing color constant descriptors in images," *Perception*, pp. 755–758, 1988.