

Reconnaissance d'objets sous éclairage non contrôlé par analyse d'histogrammes invariants et adaptés

Damien Muselet

Ludovic Macaire

Pierre Bonnet

Jack-Gérard Postaire

Laboratoire LAGIS UMR CNRS 8146
 Université des Sciences et Technologies de Lille - Cité Scientifique, Bâtiment P2 -
 59655 Villeneuve d'Ascq cedex, FRANCE

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

Résumé

Nous abordons le problème de la reconnaissance d'objets dans des images couleur à travers des changements d'éclairage entre les acquisitions de ces images. La majorité des méthodes rencontrées dans la littérature propose de comparer les images par l'intersection d'histogrammes couleur invariants qui seraient insensibles aux conditions d'éclairage lors de l'acquisition. A la différence des histogrammes couleur invariants qui analysent indépendamment chaque image, notre approche considère chaque couple constitué par l'image requête et l'une des images candidates. Elle détermine des histogrammes couleur adaptés à chacun de ces couples. La construction de ces histogrammes couleur adaptés s'appuie sur une procédure originale qui analyse les mesures de rang des pixels dans chaque image de composante. Des tests expérimentaux sur une base d'images publiques montrent l'efficacité des histogrammes couleur adaptés dans le domaine de la reconnaissance d'objets sous éclairage non contrôlé et ce, quel que soit le nombre de niveaux utilisés pour quantifier les coordonnées des cellules des histogrammes.

Mots clefs

Couleur, Reconnaissance d'objets, Éclairage, Histogrammes couleur invariants, Histogrammes couleur adaptés, Mesure de rang, Quantification.

1 Introduction

Dans cet article, nous traitons un problème particulier de l'indexation d'images, à savoir la reconnaissance d'objets par analyse de leurs couleurs, qui consiste à retrouver parmi toutes les images candidates d'une base, celles qui représentent le même objet que celui contenu dans l'image requête. Ces images contiennent un seul objet placé sur un fond uniforme afin de pouvoir extraire facilement les pixels qui représentent l'objet (cf. figure 1).

Nous considérons que la couleur de chaque pixel est représentée par un vecteur couleur \mathbf{c} , dont les coordonnées sont les niveaux discrets des trois composantes couleur, rouge (c^R), verte (c^V) et bleue (c^B). Chaque niveau est quantifié

sur N_0 niveaux, N_0 étant généralement fixé à 256.

La signature la plus utilisée pour représenter la distribution couleur d'une image \mathbf{I} est l'histogramme couleur $\mathbf{H}[\mathbf{I}]$ qui est composé de cellules repérées par leurs coordonnées sur les trois composantes, à savoir α^R , α^V et α^B . L'histogramme défini sur N_0^3 cellules étant une structure volumineuse et peu dense, ces coordonnées sont quantifiées sur N niveaux, N étant inférieur ou égal à N_0 . Cette sous-quantification des coordonnées des cellules nécessitent la définition de fonctions $f^k(c^k) = \alpha^k$, $k = R, V, B$. Ainsi, le contenu $\mathbf{H}[\mathbf{I}](\alpha^R, \alpha^V, \alpha^B)$ de chaque cellule de coordonnées $(\alpha^R, \alpha^V, \alpha^B)^T$ indique le nombre de pixels dans l'image analysée, caractérisés par les couleurs $(c^R, c^V, c^B)^T$ telles que :

$$(f^R(c^R) = \alpha^R, f^V(c^V) = \alpha^V, f^B(c^B) = \alpha^B)^T. \quad (1)$$

La mesure de similarité retenue pour la reconnaissance d'objets est l'intersection entre les histogrammes $\mathbf{H}[\mathbf{I}_{req}]$ d'une image requête \mathbf{I}_{req} et $\mathbf{H}[\mathbf{I}_{cand}]$ d'une image candidate \mathbf{I}_{cand} . Cette intersection, issue de l'intersection proposée par Swain [1], est exprimée de la manière suivante :

$$\begin{aligned} Inter(\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}]) = \\ \sum_{\alpha^R=0}^{(N-1)} \sum_{\alpha^V=0}^{(N-1)} \sum_{\alpha^B=0}^{(N-1)} \\ \min(h[\mathbf{I}_{req}](\alpha^R, \alpha^V, \alpha^B), h[\mathbf{I}_{cand}](\alpha^R, \alpha^V, \alpha^B)) \end{aligned} \quad (2)$$

avec

$$\begin{aligned} h[\mathbf{I}](\alpha^R, \alpha^V, \alpha^B) = \\ \frac{\mathbf{H}[\mathbf{I}](\alpha^R, \alpha^V, \alpha^B)}{\sum_{\alpha'^R=0}^{(N-1)} \sum_{\alpha'^V=0}^{(N-1)} \sum_{\alpha'^B=0}^{(N-1)} \mathbf{H}[\mathbf{I}](\alpha'^R, \alpha'^V, \alpha'^B)} \end{aligned} \quad (3)$$

où $h[\mathbf{I}](\alpha^R, \alpha^V, \alpha^B)$ est la valeur de la cellule $\mathbf{H}[\mathbf{I}](\alpha^R, \alpha^V, \alpha^B)$ normalisée par le nombre total de pixels qui représentent l'objet dans l'image.

Nous nous focalisons sur les problèmes consécutifs à des conditions d'éclairage non contrôlées lors des acquisitions

des images. Dans le cadre de nos travaux, les acquisitions des images de chaque objet respectent les conditions suivantes :

- les images ne contiennent qu'un seul objet,
- les images sont acquises avec la même caméra réglée de manière identique,
- l'éclairage varie d'une image à l'autre par changement de la température ou de l'intensité de l'illuminant (cf. figure 1).

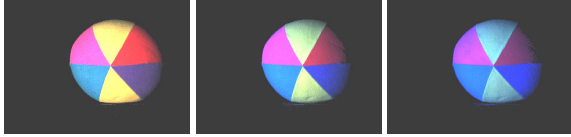


Figure 1 – Images couleur du même objet sous différents illuminants.

Les composantes couleur des pixels qui représentent les objets dépendant de l'éclairage, la reconnaissance d'objets basée sur l'intersection d'histogrammes couleur n'obtient pas de bons résultats quand l'éclairage est non contrôlé. Pour pallier ce problème, de nombreux auteurs proposent de comparer les images grâce à l'intersection d'histogrammes couleur invariants aux changements d'éclairage, présentés dans le second paragraphe. Ces histogrammes invariants sont issus de modèles de variation de la couleur qui sont trop simples pour décrire sans erreur les modifications des couleurs consécutives à un changement d'illuminant.

C'est pourquoi nous avons proposé une approche originale qui considère chaque couple constitué d'une image requête et de l'une des images candidates construit durant la recherche des images candidates représentant le même objet que celui contenu par l'image requête. De chaque couple d'images, nous examinons le couple d'histogrammes couleur ($\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}]$) et les adaptons de telle sorte à ce qu'ils soient identiques uniquement lorsque les deux images contiennent le même objet [2]. Cette approche est décrite succinctement dans le troisième paragraphe.

Les procédures de reconnaissance d'objets sont connues pour dépendre du nombre de niveaux utilisés pour quantifier les coordonnées des cellules des histogrammes. Il est généralement admis que les performances des algorithmes diminuent lorsque le nombre de niveaux diminue. L'objectif de cet article est d'analyser le comportement des intersections des histogrammes couleur invariants et adaptés lorsque le nombre N de niveaux des coordonnées des cellules varie. Cette analyse est basée sur des tests de reconnaissance d'objets effectués sur une base publique d'images couleur, dont les résultats sont présentés dans le quatrième paragraphe.

2 Les histogrammes couleur invariants

Les histogrammes couleur invariants sont calculés de manière à ce qu'ils soient aussi peu sensibles que possible aux conditions d'éclairage. Ils sont obtenus par transformation

des histogrammes couleur classiques [3].

L'évaluation de tels histogrammes couleur invariants s'appuie sur des modèles qui décrivent la variation du vecteur couleur de chaque pixel provoquée par un changement d'éclairage. La complexité de la relation réelle qui existe entre le niveau de chacune des composantes couleur d'un pixel, l'éclairage qui illumine l'élément de surface projeté sur ce pixel et chacune des fonctions de sensibilité des capteurs de la caméra qui observe l'objet est telle que ces modèles sont basés sur des hypothèses très restrictives.

Par exemple, le modèle diagonal, qui est l'un des modèles les plus utilisés, considère que les variations de vecteurs couleur provoquées par un changement d'éclairage se formalisent par une simple transformation linéaire. Ainsi, la normalisation "greyworld", basée sur ce modèle, propose de diviser les niveaux des composantes couleur de chaque pixel par la moyenne des niveaux de tous les pixels qui représentent l'objet dans l'image [1].

Par ailleurs, le modèle de Hordley suppose que l'ordre des niveaux des pixels dans une composante n'est pas modifié par un changement d'éclairage. Ce modèle conduit Hordley à considérer indépendamment les trois histogrammes 1D qui peuvent être extraits d'un histogramme couleur et à les transformer à l'aide d'une simple égalisation d'histogramme [4].

Ces deux modèles classiques étant basés sur des hypothèses très restrictives, nous proposons de calculer des histogrammes couleur adaptés dont l'évaluation est fondée sur un modèle plus réaliste.

3 Les histogrammes couleur adaptés

La principale originalité de notre approche est de traiter le problème de changements d'éclairage, non pas en considérant de manière indépendante les histogrammes couleur des images requête et candidate, mais plutôt en considérant chaque couple d'histogrammes candidat et requête pour construire le couple d'histogrammes couleur adaptés. Ces deux histogrammes couleur sont adaptés de telle manière à ce que leur intersection soit élevée seulement si les objets contenus dans les deux images comparées sont similaires [2]. Pour ce faire, nous nous basons sur un modèle de changement d'éclairage original appelé "modèle des plus proches mesures de rang" et présenté dans la partie suivante.

3.1 Le modèle des plus proches mesures de rangs

Ce modèle, issu du modèle de Hordley, considère chacune des trois images de composante I^k , $k = R, V, B$, qui constituent l'image couleur et dans lesquelles les pixels P sont respectivement caractérisés par les niveaux de composante $c^k(P)$, $k = R, V, B$. Dans chaque image de composante I^k , les pixels P sont triés dans l'ordre croissant de leurs niveaux et sont associés à des mesures de rang $R^k[\mathbf{I}](P)$, telles que les premiers classés (niveaux les plus faibles) obtiennent une mesure de rang proche de 0 et les

derniers (niveaux les plus élevés), une mesure de rang égale à 1 :

$$R^k[\mathbf{I}](P) = \frac{\sum_{i=0}^{c^k(P)} H^k[\mathbf{I}](i)}{\sum_{i=0}^{N-1} H^k[\mathbf{I}](i)}, \quad k = R, V, B, \quad (4)$$

où $H^k[\mathbf{I}](i)$ est le nombre de pixels caractérisés par le niveau i dans l'image de composante I^k .

Comme les mesures de rang de tous les pixels caractérisés par le même niveau dans un image de composante sont égales, par souci de simplification de notation, la mesure de rang d'un pixel correspond, par la suite, à la mesure de rang du niveau de ces pixel.

Considérons le cas où l'image requête \mathbf{I}_{req} et l'image candidate \mathbf{I}_{cand} contiennent le même objet observé dans les mêmes conditions mais éclairé par des illuminants différents. Soient les pixels P_{req} et P_{cand} qui représentent respectivement les mêmes éléments de surface dans les images requête et candidate. Le modèle des plus proches mesures de rangs suppose que la mesure de rang de P_{req} dans l'image de composante couleur I_{req}^k est, parmi toutes les mesures de rang des pixels de I_{req}^k , celle qui est la plus proche de la mesure de rang de P_{cand} dans I_{cand}^k :

$$\begin{aligned} & \|R^k[\mathbf{I}_{cand}](P_{cand}) - R^k[\mathbf{I}_{req}](P_{req})\| = \\ & \min_{P \in I_{req}^k} \|R^k[\mathbf{I}_{cand}](P_{cand}) - R^k[\mathbf{I}_{req}](P)\|. \end{aligned} \quad (5)$$

Pour réduire les effets de changements d'illuminant entre les images requête et candidates, nous proposons d'adapter leurs histogrammes par modification des niveaux de gris de façon à ce que, dans les histogrammes adaptés, les rangs soient les plus proches, à niveau de gris égal.

3.2 Les histogrammes couleur adaptés

Adaptation d'histogrammes. Pour chaque couple d'images ($\mathbf{I}_{req}, \mathbf{I}_{cand}$), l'approche consiste à transformer le couple d'histogrammes couleur ($\mathbf{H}[\mathbf{I}_{req}], \mathbf{H}[\mathbf{I}_{cand}]$) en un couple d'histogrammes couleur adaptés ($\mathbf{H}_{adapt}[\mathbf{I}_{req}], \mathbf{H}_{adapt}[\mathbf{I}_{cand}]$). Cette adaptation s'applique indépendamment pour chaque composante couleur. Elle revient donc à transformer, de manière indépendante, les trois couples d'histogrammes 1D ($H^k[\mathbf{I}_{req}], H^k[\mathbf{I}_{cand}]$), $k = R, V, B$ qui peuvent être extraits du couple d'histogrammes couleur en trois couples d'histogrammes 1D adaptés ($H_{adapt}^k[\mathbf{I}_{req}], H_{adapt}^k[\mathbf{I}_{cand}]$), $k = R, V, B$.

Pour une composante k donnée, cette transformation est basée sur un couple de tables 1D de correspondance ($LUT_{adapt}^k[\mathbf{I}_{req}], LUT_{adapt}^k[\mathbf{I}_{cand}]$) telles que :

$$\begin{cases} H_{req, cand}^k[\mathbf{I}_{req}](x) = \sum_{LUT_{adapt}^k[\mathbf{I}_{req}](c) = x} H^k[\mathbf{I}_{req}](c), \\ \text{et} \\ H_{req, cand}^k[\mathbf{I}_{cand}](x) = \sum_{LUT_{adapt}^k[\mathbf{I}_{cand}](c) = x} H^k[\mathbf{I}_{cand}](c), \end{cases} \quad (6)$$

avec $k = R, V, B$.

Ainsi, la cellule de l'histogramme 1D adapté $H_{adapt}^k[\mathbf{I}_{req}]$ associée au niveau x regroupe les contenus des cellules de l'histogramme 1D $H^k[\mathbf{I}_{req}]$ associées aux niveaux c , tels que $LUT_{adapt}^k[\mathbf{I}_{req}](c)$ est égale à x .

De la même manière, la cellule de l'histogramme 1D adapté $H_{adapt}^k[\mathbf{I}_{cand}]$ associée au niveau x regroupe les contenus des cellules de l'histogramme 1D $H^k[\mathbf{I}_{cand}]$ associées aux niveaux c , tels que $LUT_{adapt}^k[\mathbf{I}_{cand}](c)$ est égale à x .

La détermination du couple d'histogrammes 1D adaptés repose sur le couple de tables 1D de correspondance des niveaux de gris.

Couple de tables de correspondance. Les histogrammes 1D requête et candidat adaptés doivent contenir au maximum N cellules non vides. Par conséquent, la procédure regroupe les contenus des cellules de chaque histogramme 1D qui sont associées à des niveaux contigus pour déterminer les N cellules de l'histogramme 1D adapté qui en découle.

Pour construire l'histogramme 1D requête (respectivement candidat) adapté, la procédure détermine un ensemble de $(N + 1)$ niveaux parmi les N_0 possibles, notés X_{req}^k (respectivement notés X_{cand}^k) qui délimitent N intervalles. Les contenus des cellules de l'histogramme 1D requête (respectivement candidat) associées aux niveaux appartenant aux mêmes intervalles sont regroupés dans les N cellules de l'histogramme 1D requête (respectivement candidat) adapté.

Le couple de tables de correspondance 1D ($LUT_{adapt}^k[\mathbf{I}_{req}], LUT_{adapt}^k[\mathbf{I}_{cand}]$) s'exprime à partir des couples de niveaux retenus ($X_{req}^k(z), X_{cand}^k(z)$) comme :

$$\begin{cases} LUT_{adapt}^k[\mathbf{I}_{req}](x) = z \\ \quad \text{pour } x \in [X_{req}^k(z), X_{req}^k(z + 1)[, \\ \text{et} \\ LUT_{adapt}^k[\mathbf{I}_{cand}](x) = z \\ \quad \text{pour } x \in [X_{cand}^k(z), X_{cand}^k(z + 1)[, \end{cases} \quad (7)$$

avec $z = 0, \dots, N - 1$, $(X_{req}^k(0), X_{cand}^k(0)) = (0, 0)$ et $(X_{req}^k(N), X_{cand}^k(N)) = (N_0, N_0)$.

L'équation (7) montre que les sous-ensembles de pixels de I_{req}^k dont les niveaux appartiennent à $[X_{req}^k(z), X_{req}^k(z + 1)[$ et les sous-ensembles de pixels de I_{cand}^k dont les niveaux appartiennent à $[X_{cand}^k(z), X_{cand}^k(z + 1)[$, sont tous les deux représentés par des cellules de mêmes coordonnées dans les histogrammes 1D adaptés. La détermination de ces couples de niveaux ($X_{req}^k(z), X_{cand}^k(z)$), avec $z = 1, \dots, N - 1$, influence donc directement l'intersection entre les histogrammes adaptés.

Détermination des couples de niveaux pertinents. Les couples de niveaux sont déterminés de manière à réduire les effets des changements d'illuminant se produisant entre

les acquisitions des images requête et candidate sur les histogrammes couleur requête et candidat adaptés. Chaque couple de niveaux doit caractériser respectivement dans chacune des images de composante des pixels qui observent les mêmes éléments de surface.

Cette détermination est basée sur notre modèle de changement d'éclairage qui suppose que les pixels dont les mesures de rang sont les plus proches dans chaque couple d'images de composante représentent les mêmes éléments de surface. Pour chaque couple d'images de composante (I_{req}^k, I_{cand}^k) , les $(N - 1)$ couples de niveaux $(X_{req}^k(z), X_{cand}^k(z))$ retenus par la procédure seront donc ceux dont les mesures de rang sont parmi les plus proches dans ces images. Ainsi, les couples de niveaux sélectionnés sont ceux qui caractérisent le mieux des pixels observant les mêmes éléments de surface dans le couple d'images de composante considéré.

4 Résultats expérimentaux

Nous allons ici comparer les performances atteintes par notre méthode à celles atteintes par la normalisation du greyworld [3] et l'égalisation d'histogrammes [4] dans le domaine de la reconnaissance d'objets.

Les tests ont été effectués sur la base publique de l'université de Simon Fraser constituée de 11 objets éclairés par l'un des 5 illuminants disponibles (55 images au total) [1]. La figure 2 représente les pourcentages de cas où l'image candidate qui représente le même objet que l'image requête a été classée première. Ces pourcentages de bons résultats sont représentés en fonction du nombre N de niveaux avec lesquels sont quantifiées les coordonnées des cellules des histogrammes invariants ou adaptés.

Sur la figure 2, nous constatons que les résultats de reconnaissance d'objets obtenus par l'intersection d'histogrammes invariants diminuent lorsque le nombre N augmente, tandis que ceux obtenus par l'intersection d'histogrammes couleur adaptés restent stables. Plus la couleur est quantifiée, moins bonnes sont les performances des intersections d'histogrammes invariants.

La base d'images utilisée est réduite. Elle ne permet pas de juger complètement de l'efficacité d'une méthode pour la reconnaissance d'objets à travers des changements d'éclairage. Elle permet néanmoins de mettre en évidence que les histogrammes invariants sont très sensibles au nombre de niveaux sur lesquels sont quantifiés les coordonnées des cellules des histogrammes. Ce phénomène s'explique par les hypothèses très restrictives sur lesquelles s'appuient les calculs des histogrammes invariants. Par ailleurs, les résultats obtenus par notre méthode illustrent sa robustesse à cette quantification.

Ils montrent l'intérêt de l'intersection des histogrammes couleur adaptés pour la reconnaissance d'objets en cas de changements d'éclairage.

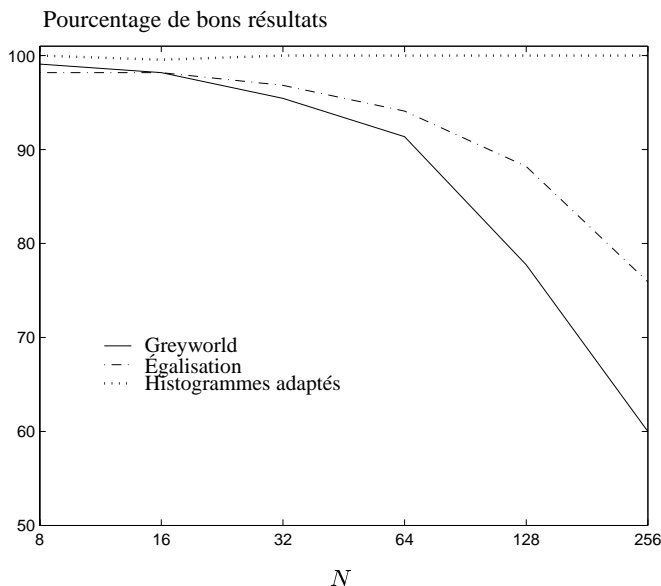


Figure 2 – Évolution des résultats de reconnaissance d'objets sur la base de l'université de Simon Fraser en fonction du nombre N de niveaux avec lesquels sont quantifiées les coordonnées des cellules des histogrammes.

Références

- [1] B.V. Funt, K. Barnard, et L. Martin. Is machine colour constancy good enough? Dans *Proceedings of the 5th European Conference on Computer Vision*, pages 445–459, 1998.
- [2] D. Muselet, L. Macaire, J.G. Postaire, et P. Bonnet. Reconnaissance d'objets grâce à l'analyse des composantes couleur adaptées au changement d'éclairage entre deux images. Dans *19e Colloque GRETSI sur le Traitement du Signal et des Images*, volume 3, pages 30–34, Paris France, 2003.
- [3] G.D. Finlayson et G. Schaefer. Colour indexing across devices and viewing conditions. Dans *Proceedings of 2nd Int. Workshop on Content-based MultiMedia Indexing*, pages 215–221, Brescia, Italy, 2001.
- [4] S.D. Hordley, G.D. Finlayson, G. Schaefer, et 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.