INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Indexation et appariement d'images par modèle de mélange gaussien des couleurs

Christophe Biernacki et Roger Mohr

No 3600

Janvier 1999

THÈME 3

Rapport de recherche

ISSN 0249-6399

Indexation et appariement d'images par modèle de mélange gaussien des couleurs

Christophe Biernacki* et Roger Mohr*

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Movi

Rapport de recherche n 3600 — Janvier 1999 — 15 pages

Résumé : Nous proposons d'indexer un objet, ou partie d'image, par le paramètre d'un mélange gaussien qui approxime la distribution des couleurs, le nombre de classes étant choisi automatiquement par des critères bayésiens. Tout objet requête sera alors apparié à l'objet indexé le plus vraisemblable parmi tous les objets indexés de la base d'apprentissage. Cette technique est ensuite comparée à celle de l'indexation par histogrammes utilisant la dissimilarité du chi-deux pour apparier.
L'avantage des mélanges gaussiens repose principalement sur un gain important de la taille de l'index par rapport aux histogrammes surtout lorsque la dimension augmente, tout en gardant une qualité d'appariement du même ordre.

Mots-clé : Taille d'indexation, taux d'appariement, mélange de distributions, nombre de classes.

(Abstract: pto)

* laboratoire Gravir-Imag et projet Movi

Indexing and Matching Images with a Gaussian Mixture Model of Colors

Abstract: An object, it means a part of an image, is indexed by the Gaussian mixture parameter which approximates the color distribution, the number of clusters being automatically choosen by Bayesian criteria. Every new object is then matched with the learned object having the largest likelihood. This method is compared to histograms using chi-deux dissimilarity to perform matching.
Gaussian mixture models both decrease drastically the size of the index, especially in higher dimension, and keep a good matching capability.

Key-words: Index size, matching rate, distributions mixture, number of clusters.

1 Approximation de la distribution des couleurs

Un objet $\mathcal{O}$ est composé de n pixels, chacun étant décrit par une position x_i dans l'image et une couleur c_i ($i = 1, \ldots, n$). On entend ici par le terme couleur c_i aussi bien le vecteur couleur (r_i, g_i, b_i) que le niveau de gris associé ou bien encore la fonction de normalisation interne de (r_i, g_i, b_i).
L'histogramme des couleurs est une approximation de la distribution des couleurs (Swain et Ballard, 1991). Cette information est globale puisqu'elle élimine l'information de position. Cette façon d'indexer un objet $\mathcal{O}$ peut s'interpréter en terme de mélange probabiliste. Cela nous conduira naturellement à introduire le modèle de mélange gaussien comme une alternative possible.

1.1 Histogramme

Considérer un histogramme en K classes revient en effet à supposer implicitement que $c_1, \ldots, c_n$ sont n réalisations i.i.d. d'une distribution mélange f de K lois uniformes. On a

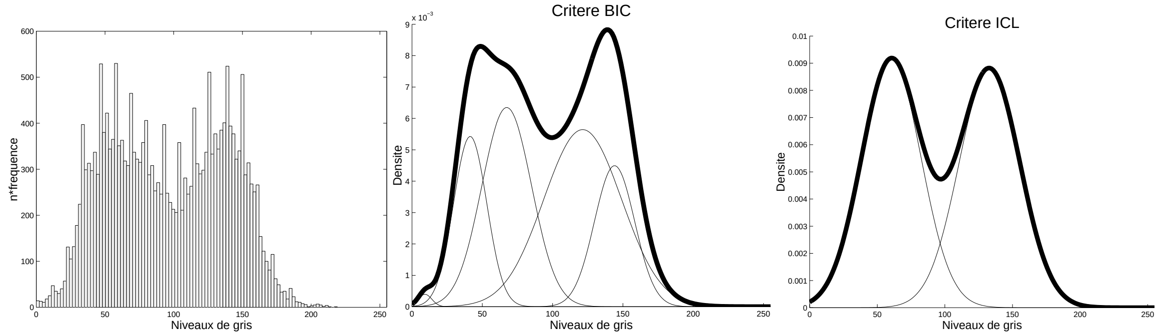
$$f(c|\theta) = \sum_{k=1}^{K} p_k U(c|a_k)$$

où p_k est la proportion de la classe k ($p_k \geq 0$ et $\sum_k p_k = 1$), $U(\cdot|a_k)$ est la loi uniforme de support a_k et le paramètre du mélange θ regroupe l'ensemble des p_k et a_k.
Généralement, l'espace des couleurs est partitionné *a priori* en K régions. Ainsi, les supports a_k sont fixés avant l'expérimentation et disjoints. En fait, ces K régions forment presque toujours un maillage régulier de l'espace des couleurs. Quoiqu'il en soit, les seuls paramètres à estimer sont les proportions p_k et l'estimateur du maximum de vraisemblance associé est simplement la fréquence de chaque couleur dans chaque classe. On obtient ainsi un histogramme dit «normalisé».

1.2 Mélange gaussien

1.2.1 Justification

À ce stade, deux raisons plaident pour l'emploi d'un mélange de K lois normales plutôt qu'uniformes. Tout d'abord, les paramètres a_k de position-dispersion ne peuvent être estimés dans le cas uniforme à cause de la non-identifiabilité du mélange ainsi induite. D'autre part, et de manière plus empirique, la distribution des couleurs, «à l'oeil», semble effectivement composée de une à plusieurs lois normales. Cela est visible sur la figure 1(a) qui donne l'histogramme des niveaux de gris du premier objet de la séquence violetta (voir l'expérimentation plus loin). Tous les autres objets donnent cette même impression.
Chaque gaussienne est composée par un centre, une matrice de variance et une proportion. Chacune de ces trois grandeurs caractérisent une *couleur dominante*, c'est-à-dire une couleur particulièrement présente dans l'objet. Le centre indique la valeur moyenne de cette couleur, la matrice de variance indique la dispersion de cette couleur autour de ce centre et la proportion donne le taux de présence de cette couleur par rapport aux autres couleurs dominantes.



(a) (b) (c)

FIG. 1 – *Premier objet de la séquence violetta : (a) histogramme des niveaux de gris, (b) composantes gaussiennes (trait fin) et densité mélange (trait épais) obtenus avec le critère BIC, (c) même principe avec le critère ICL.*

1.2.2 Formalisation

Formellement, le mélange gaussien f s'écrit

$$f(c|\theta) = \sum_{k=1}^{K} p_k N(c|a_k)$$

avec $a_k = (\mu_k, \Sigma_k)$, μ_k et Σ_k étant respectivement le centre et la matrice de variance de la $k$ième composante normale $N(\cdot|a_k)$.
Cette fois, non seulement les proportions p_k mais aussi les positions-dispersions a_k peuvent être estimées car le mélange est identifiable. Cette augmentation de la flexibilité du modèle permettra de mieux estimer la distribution des couleurs. L'estimation de tous ces paramètres peut se faire en maximisant la log-vraisemblance du paramètre global θ :

$$\ln f(C|\theta) = \sum_{i=1}^{n} \ln f(c_i|\theta),$$

avec $C = (c_1, \ldots, c_n)$. Cette maximisation peut être faite par l'algorithme EM (*expectation - maximisation*) de Dempster, et al. (1977).
De plus, le nombre K de composantes peut aussi être estimé en retenant la valeur entre $K = 1$ et K_{sup} (K_{sup} est à choisir a priori) qui minimise le critère bayésien BIC (Schwarz, 1978)

$$\text{BIC}(K) = -2\ln f(C|\hat{\theta}_K) + \nu_K \ln n,$$

avec $\hat{\theta}_K$ et ν_K respectivement l'estimateur du maximum de vraisemblance et le nombre de degrés de liberté du modèle. Un concurent au critère BIC, le critère ICL (Biernacki et al., 1998) peut aussi être employé. Ce critère correspond à une version classifiante du critère bayésien BIC. Il a l'avantage d'être plus robuste que BIC à la non normalité des classes et aussi à l'augmentation de la dimension de l'espace des données. Il s'écrit

$$\begin{aligned} \text{ICL}(K) &= -2\ln f(C|\tilde{z}_K, \hat{a}_K) + \nu_K \ln n \\ &\quad -2\ln\Gamma(\frac{K}{2}) - 2\sum_{k=1}^{K} \ln\Gamma(\tilde{n}_k + \frac{1}{2}) \\ &\quad +2K\ln\Gamma(\frac{1}{2}) + 2\ln\Gamma(n + \frac{K}{2}), \end{aligned}$$

avec $\tilde{z}_K$ la partition des n points en K classes obtenue à partir de l'estimateur $\hat{\theta}_K$, $\hat{a}_K$ l'estimateur du maximum de vraisemblance des centres et des matrices de variance obtenu en utilisant la partition $\tilde{z}_K$, et enfin $\tilde{n}_k$ le cardinal de la classe k de cette partition.
Les sous-figures 1(b) et (c) donnent le comportement des deux critères avec l'exemple du premier objet de la séquence violetta.
Finalement, un objet $\mathcal{O}$ est indexé par $K-1$ proportions si un histogramme est utilisé. Dans le cas du mélange gaussien, $K-1+Kd+K(d(d+1)/2)$ paramètres ($K-1$ pour les proportions, Kd pour les centres et $K(d(d+1)/2)$ pour les matrices de variance) sont nécessaires, avec d la dimension du vecteur couleur ($d=1$ en niveau de gris, $d=3$ en RGB).
Notons que la détermination du nombre de classes n'est pas propre au cas du mélange gaussien car les critères précédents peuvent aussi bien s'appliquer dans le cas des histogrammes. Néanmoins, la méthode d'appariement des histogrammes dont nous parlerons plus loin suppose un même nombre de classes pour tous les histogrammes. La liberté du choix de K indépendamment pour chaque histogramme est donc limitée. Il faudrait proposer une extension de ce type d'appariement, problème que nous n'avons pas approfondi ici.

1.2.3 Robustesse au changement de luminosité

Il est fort probable que le modèle de mélange gaussien décrive mieux, au sens de la vraisemblance, la distribution des couleurs que ne peut le faire l'histogramme avec un nombre «raisonnable» de classes. Pourtant, il est possible que l'indexation par l'histogramme soit plus robuste que notre approche.
En effet, un changement de luminosité d'un objet se traduit essentiellement par une translation de ses couleurs dominantes. L'influence sur les centres de nos gaussiennes est donc évidente (Gros et al., 1997). Au contraire, si l'histogramme n'a pas trop de classes, le centre de la couleur dominante, ainsi que l'essentiel de la masse de probabilité, avant et après le changement de luminosité ont de fortes de chances de rester dans la même classe de l'histogramme. Par conséquent, l'histogramme est très peu modifié en général.
Pour atténuer l'effet de la luminosité sur le mélange gaussien, nous proposons d'augmenter «artificiellement» la dispersion de chaque couleur dominante. Ainsi, la densité mélange sera moins sensible à une translation «raisonnable» des centres. L'augmentation de la dispersion des classes se fera en multipliant chaque matrice de variance $\hat{\Sigma}$ par un même coefficient λ ($\lambda > 1$). Le terme λ se doit d'être assez grand pour atténuer l'influence de la luminosité. Il se doit aussi de ne pas être trop grand au risque de perdre l'information discriminante dans l'objet.
Pour définir en pratique λ, nous utilisons la borne supérieure d'un intervalle de confiance de la variance généralisée. La variance généralisée $|\Sigma|$ est définie comme le déterminant de la matrice de variance Σ. Elle correspond donc à une extension de la variance unidimensionnelle. Asymptotiquement, on a la distribution asymptotique suivante de l'estimateur sans biais $\hat{\Sigma}$ de Σ (Anderson, 1971) :

$$\sqrt{\frac{n}{2d}}\left(\frac{|\hat{\Sigma}|}{|\Sigma|}-1\right) \sim N(\cdot|0,1),$$

avec n la taille de l'échantillon et d la dimension. On en déduit que la borne supérieure b_α de $|\Sigma|$ de l'intervalle de confiance unilatéral à droite avec niveau $1-\alpha$ est

$$b_\alpha = \frac{|\hat{\Sigma}|}{\sqrt{\frac{2d}{n}}u_\alpha + 1},$$

avec u_α le fractile d'ordre α de la loi normale. Ainsi

$$\lambda = \frac{1}{\sqrt{\frac{2d}{n}}u_\alpha + 1}.$$

Le nombre de couleurs d'un objet intervenant dans l'estimation de la matrice de variance de la classe de couleur dominante correspondante est très grand. Ce nombre est donc inutilisable pour obtenir un intervalle de confiance assez large avec α classique, et donc obtenir un λ significativement plus grand que un. Nous proposons alors de prendre une taille fictive $n = 100d$, taille qui augmente donc avec la dimension de l'espace. Prenant aussi $\alpha = 10^{-3}$, nous obtenons

$$\lambda = 1.7763.$$

2 Appariement avec les objets indexés

Durant la phase d'apprentissage, m objets $\mathcal{O}^1, \ldots, \mathcal{O}^m$ ont été respectivement indexés par les paramètres de mélange $\hat{\theta}^1, \ldots, \hat{\theta}^m$. Le paramètre $\hat{\theta}$ correspond ici aussi bien au cas de l'histogramme qu'au cas du mélange gaussien avec adaptation ou non au changement de luminosité. L'objectif est alors d'apparier tout nouvel objet $\mathcal{O}$ avec un des m objets indexés. Nous distinguons les méthodes d'appariement en fonction des deux types d'indexation.

2.1 Histogramme

Dans le cas des histogrammes, Schiele (1997) a comparé plusieurs méthodes d'appariement : dissimilarité du χ^2, intersection d'histogramme, distances quadratiques. Il ressort qu'une version de la dissimilarité du χ^2 donne les meilleurs résultats. Elle est donnée par :

$$\chi^2(\mathcal{O}, \mathcal{O}^j) = \begin{cases} \sum_{k=1}^{K} (\hat{p}_k - \hat{p}_k^j)^2 / (\hat{p}_k + \hat{p}_k^j) & \text{si } \hat{p}_k + \hat{p}_k^j \neq 0 \\ 0 & \text{sinon,} \end{cases}$$

où $\hat{p}_k$ est la proportion estimée de la composante k dans l'histogramme de $\mathcal{O}$ et $j = 1, \ldots, m$.

2.2 Mélange gaussien

Dans le cas du mélange gaussien, des hypothèses probabilistes précises ont été faites. Nous proposons alors de les utiliser pour l'appariement. Nous apparions $\mathcal{O}$ à $\mathcal{O}^j$ qui est le plus probable conditionnellement à la connaissance de $\mathcal{O}$. Cette probabilité se calcul par

$$P(\mathcal{O}^j|\mathcal{O}) = \frac{P(\mathcal{O}|\mathcal{O}^j)P(\mathcal{O}^j)}{P(\mathcal{O})}.$$

Supposant chaque objet de la base équiprobable a priori, on obtient :

$$P(\mathcal{O}^j|\mathcal{O}) \propto P(\mathcal{O}|\mathcal{O}^j).$$

Notant $c_1, \ldots, c_n$ les couleurs de $\mathcal{O}$, on pose

$$P(\mathcal{O}|\mathcal{O}^j) = f(c_1, \ldots, c_n|\hat{\theta}^j).$$

Ainsi, on apparie simplement $\mathcal{O}$ avec l'objet $\mathcal{O}^j$ dont le paramètre $\hat{\theta}^j$ est le plus vraisemblable.

3 Expérimentation

3.1 La séquence «violetta»

La séquence d'images «violetta» est représentée sur les figures 2 et 3. Elle comporte 97 objets répartis en 67 images, chaque image étant donc composée de un à deux objets. La séquence est divisée en 11 plans et les objets similaires dans le même plan ont le même label. Par exemple le personnage à la casquette dans le plan 5 a le label 7. Le label complet d'un objet sera noté $x.y$ avec x le numéro de l'image où se produit l'objet y. Par exemple 31.7 pointe sur l'objet de label 7 dans l'image 31. L'indexation se fait sur l'(es) objet(s) de la première image de chaque plan, soit au total 16 objets dans la base d'indexation : 2.1, 4.2, 17.3, 17.4, 26.6, 25.6, 31.7, 31.8, 43.9, 51.10, 51.11, 69.12, 73.13, 78.14, 78.15, 83.16.
Des objets de plans différents peuvent aussi être similaires comme le personnage à la casquette de label 3 et 7 dans respectivement les plans 3 et 5. Voici la liste des 9 ensembles d'objets similaires, qu'ils soient ou non dans le même plan (le signe «$*$» indiquant l'indifférence sur le numéro de l'image) : $\{*.1,*.13\}$, $\{*.2,*.6,*.9,*.10,*.16,*.14\}$, $\{*.3,*.7\}$, $\{*.4,*.8\}$, $\{*.5\}$, $\{*.11\}$, $\{*.12,*.15\}$. Cette relation de similarité entre les différents objets sera utile pour mesurer la performance des méthodes d'appariement que nous allons étudier.

3.2 Méthodes comparées

Nous proposons de comparer les approches suivantes pour les niveaux de gris, les couleurs RGB et les couleurs RGB après normalisation interne :

- mélanges gaussiens d'origine avec critère BIC ou ICL ;
- mélanges gaussiens «flous», c'est-à-dire utilisant la technique de robustification au changement de luminosité, avec critère BIC ou ICL ;
- histogrammes avec différents nombres de classes par axe : 4, 8, 32.

FIG. 2 – *Première partie de la séquence d'images «violetta». Les objets de la base sont notés «LEARN».*

La détection du nombre de classes dans les méthodes gaussiennes est automatique grâce aux critères BIC et ICL. Dans le cas des niveaux de gris, BIC et ICL sélectionnent ce nombre entre 1 et 6. Dans le cas RGB avec ou sans normalisation, ce nombre est choisi entre 1 et 8. L'estimation des paramètres du mélange se fait en initialisant l'algorithme CEM (Celeux et Govaert, 1992) 10 fois au hasard puis en lançant l'algorithme EM avec les paramètres initiaux ainsi obtenus. Nous limitons ainsi la dépendance de EM au choix des paramètres initiaux. D'autre part, l'estimation ne se fait qu'avec 10% des pixels choisis au hasard dans le cas RGB avec ou sans normalisation, ceci pour limiter le temps de calcul.
L'indexation par ces approches se fait sur les 16 objets décrits précédemment. Le nombre moyen de classes retenu pour chacun de ces 16 objets et son écart-type sont donnés dans le tableau 1.
Notons qu'un problème se pose pour la calcul des histogrammes dans le cas de RGB après normalisation interne. La normalisation interne consiste à s'affranchir d'un changement interne de la source de lumière (changement de luminosité). Elle remplace chaque (r_i, g_i, b_i) par (r'_i, g'_i, b'_i) définis par

$$r'_i = n\frac{r_i}{\sum_{j=1}^n r_j}, \quad g'_i = n\frac{g_i}{\sum_{j=1}^n g_j}, \quad b'_i = n\frac{b_i}{\sum_{j=1}^n b_j}.$$

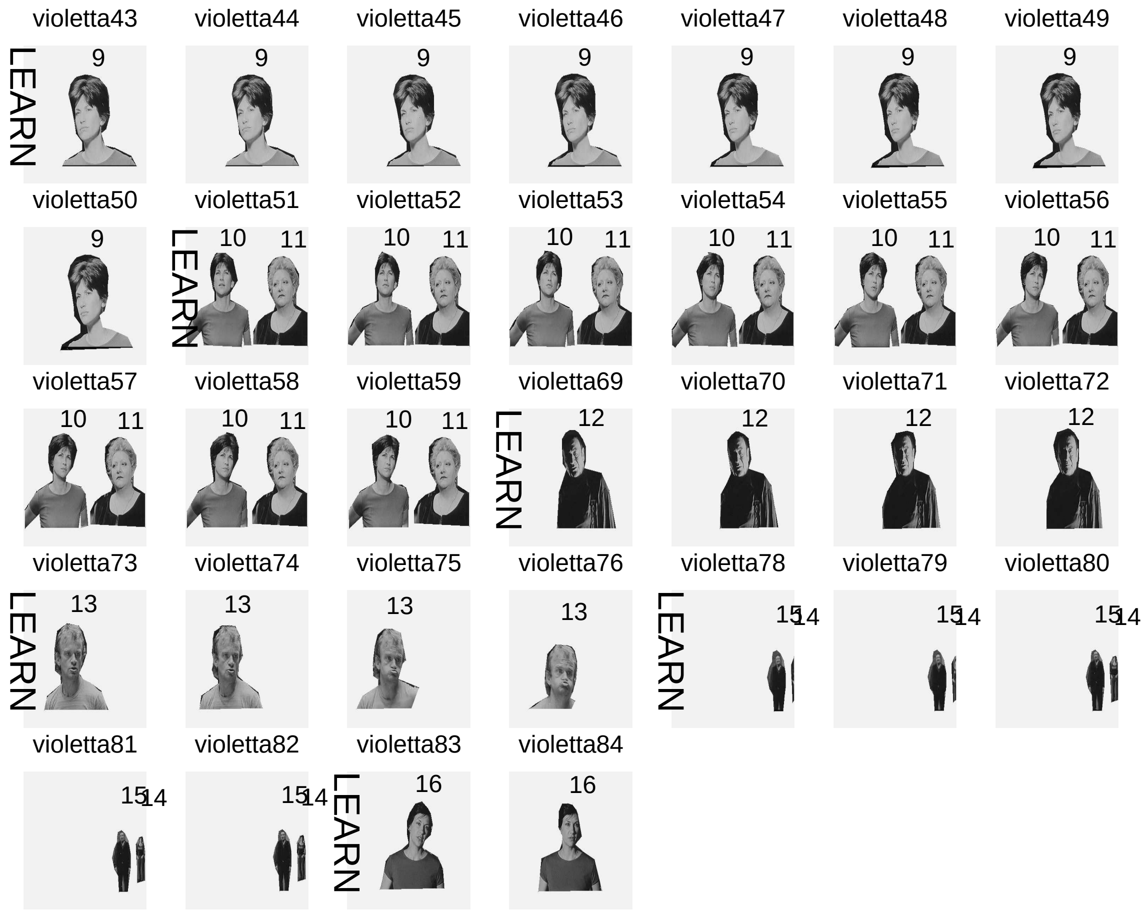


FIG. 3 – *Seconde et dernière partie de la séquence d'images «violetta». Les objets de la base sont notés «LEARN».*

	niveaux de gris	RGB	RGB norm.
BIC	5.37 (0.88)	5.87 (1.99)	5.93 (2.48)
ICL	3.56 (1.67)	5.18 (1.79)	5.31 (2.21)

TAB. 1 – *Moyenne du nombre de classes (entre parenthèses l'ecart-type) sélectionné par les critères BIC et ICL.*

Ainsi on a $\sum_{i=1}^{n} r'_i = \sum_{i=1}^{n} g'_i = \sum_{i=1}^{n} b'_i = n$. Au départ on a $0 \leq r_i \leq 255$ et maintenant $0 \leq r'_i \leq n$. La borne supérieure pour r'_i, g'_i et b'_i est donc n. Malheureusement, si l'on découpe l'intervalle $[0, n]$ en huit classes par exemple, en général tous les r'_i, g'_i et b'_i se situeront dans la première d'entre elle. Cela est du au fait que la moyenne de ces couleurs normalisées est égale à un, et par conséquent les valeurs sont souvent très groupées autour de cette valeur. Par exemple, dans l'image 2.1, on a $n = 21892$ et $\max(r'_i) = 2.16 << 21892$. Pour résoudre ce problème, nous remplacerons l'intervalle $[0, n]$ par $[0, \max(r'_i)]$ pour r'_i, et ainsi de suite pour les deux autres composantes.

3.3 Résultats

Pour chacune des sept méthodes (mélange gaussien avec BIC ou ICL, mélange gaussien «flou» avec BIC ou ICL, les trois types d'histogrammes), pour chacun des trois types de couleurs (niveaux de gris, RGB, RGB norm.) nous avons un tableau, soit au total 21 tableaux. Chaque tableau classe pour chacun des 97 objets les 16 objets de la base par ordre d'appariement. Par exemple, le tableau 2 indique les cinq premiers objets les mieux appariés, ainsi que leur classement, pour chacun des 97 objets de «violetta» dans le cas des niveaux de gris avec modèle de mélange utilisant le critère BIC.
L'exploitation fonctionnelle de ces 21 tableaux va être facilité en définissant un indice global de performance. Nous proposons comme indice le taux de bon appariement jusqu'au rang r, $\tau(r)$ $(r = 1, \ldots, 16)$. Il est donné par le rapport suivant :

$$\tau(r) = \frac{\text{nombre de bons appariements jusqu'au rang } r \text{ pour les 97 objets}}{\text{nombre d'appariements possibles jusqu'au rang } r \text{ pour les 97 objets}}.$$

Cet indice est compris entre 0 et 1. La méthode est d'autant meilleure que son indice s'approche de 1.
Nous résumons la performance de chaque méthode par la moyenne de $\tau(r)$ sur l'ensemble des 16 rangs dans le tableau 3. Ce tableau donne également l'ecart-type de l'indice. En complément de ces statistiques sur l'indice, le tableau 4 donne la taille moyenne et l'écart-type de l'index. Cette taille correspond au nombre de degrés de liberté de l'histogramme ou du mélange gaussien, en fonction de la méthode considérée.
Il ressort de ces deux derniers tableaux que :

- Que ce soit avec le critère BIC ou ICL, la méthode gaussienne résistant au changement de luminosité (gauss. flou) améliore toujours les résultats de la méthode gaussienne de base (gauss.).
- En faible dimension (niveaux de gris), le critère BIC donne de meilleurs résultats que le critère ICL pour les méthodes gauss. et gauss. flou. Par contre en dimension supérieure (RGB et RGB norm.) le critère ICL est préférable. Le comportement meilleur de ICL en plus grande dimension était un fait connu. Nous ne considérerons plus que les résultats obtenus avec ICL dans la suite.
- Le nombre de classes dans l'histogramme a une grande influence sur les performances. Il ressort que l'histogramme avec huit classes (histo. 8) par axe donne globalement les meilleurs résultats, ce pour les trois types de couleurs. Nous ne considérerons plus que cet histogramme dans la suite.
- Les résultats de histo. 8 se dégradent bizarrement entre RGB et RGB norm. tandis qu'ils s'améliorent avec ICL. Cela est du à notre façon empirique de choisir le support de l'histogramme dans le cas RGB norm.
- La méthode gauss. avec ICL est moins bonne que histo. 8 en niveaux de gris. C'est le contraire en RGB et RGB norm.
- La méthode gauss. flou avec ICL est meilleure que histo. 8, ce pour les trois types de couleur.

Objet requête	Classement des objets de la base				
	1	2	3	4	5
2.1	**2.1**	17.4	31.8	4.2	25.5
3.1	**2.1**	31.8	17.4	4.2	25.5
4.2	**4.2**	25.5	**25.6**	51.11	2.1
5.2	**4.2**	25.5	**25.6**	51.11	**51.10**
6.2	**4.2**	25.5	**25.6**	51.11	**51.10**
7.2	**4.2**	25.5	**25.6**	51.11	**51.10**
8.2	**4.2**	25.5	**25.6**	51.11	**51.10**
9.2	**4.2**	25.5	**25.6**	51.11	2.1
10.2	**4.2**	25.5	**25.6**	51.11	2.1
11.2	**4.2**	25.5	**25.6**	51.11	2.1
12.2	**4.2**	25.5	**25.6**	51.11	2.1
13.2	**4.2**	**25.6**	25.5	2.1	17.4
14.2	**4.2**	**25.6**	25.5	**78.14**	69.12
15.2	**4.2**	**25.6**	25.5	**78.14**	69.12
16.2	78.15	**78.14**	**4.2**	69.12	**25.6**
17.3	**17.3**	**31.7**	73.13	17.4	25.5
17.4	**17.4**	**31.8**	2.1	73.13	25.5
18.3	**17.3**	**31.7**	73.13	17.4	51.10
18.4	**17.4**	**31.8**	2.1	73.13	25.5
19.3	**17.3**	**31.7**	73.13	17.4	51.10
19.4	**17.4**	**31.8**	2.1	73.13	17.3
20.3	**17.3**	**31.7**	73.13	25.5	17.4
20.4	**17.4**	**31.8**	2.1	73.13	17.3
21.5	**25.5**	17.4	17.3	43.9	31.8
22.5	**25.5**	17.4	17.3	43.9	31.8
23.5	**25.5**	4.2	51.11	17.4	31.8
24.5	**25.5**	4.2	51.11	17.4	43.9
25.5	**25.5**	4.2	51.11	17.4	43.9
25.6	**25.6**	**4.2**	25.5	17.4	**51.10**
26.5	**25.5**	4.2	51.11	43.9	17.4
26.6	**25.6**	**4.2**	25.5	**51.10**	17.4
27.5	**25.5**	51.11	4.2	43.9	17.4
27.6	**25.6**	**4.2**	**51.10**	25.5	17.4
28.5	**25.5**	51.11	4.2	43.9	17.4
28.6	**25.6**	**4.2**	**51.10**	25.5	17.4
29.5	**25.5**	51.11	4.2	43.9	17.4
29.6	**25.6**	**4.2**	**51.10**	25.5	17.4
30.5	**25.5**	4.2	51.11	43.9	2.1
30.6	**25.6**	**4.2**	**51.10**	17.4	25.5
31.7	**31.7**	**17.3**	73.13	51.10	2.1
31.8	**31.8**	**17.4**	2.1	83.16	4.2
32.7	**31.7**	**17.3**	73.13	51.10	2.1
32.8	**31.8**	**17.4**	2.1	83.16	4.2
33.7	**31.7**	**17.3**	73.13	51.10	2.1
33.8	**31.8**	**17.4**	2.1	83.16	4.2
34.7	**31.7**	**17.3**	73.13	51.10	17.4
34.8	**31.8**	**17.4**	2.1	83.16	4.2
35.7	**31.7**	**17.3**	73.13	51.10	17.4
35.8	**31.8**	**17.4**	2.1	83.16	4.2

Objet requête	Classement des objets de la base				
	1	2	3	4	5
36.7	**17.3**	**31.7**	73.13	17.4	25.5
36.8	**31.8**	**17.4**	2.1	25.5	4.2
43.9	**43.9**	51.11	25.5	17.3	31.7
44.9	**43.9**	51.11	17.3	25.5	73.13
45.9	**43.9**	51.11	25.5	17.3	31.7
46.9	**43.9**	51.11	25.5	17.3	31.7
47.9	**43.9**	51.11	25.5	17.3	31.7
48.9	**43.9**	51.11	25.5	17.3	**4.2**
49.9	**43.9**	51.11	25.5	17.3	**4.2**
50.9	**43.9**	51.11	25.5	17.3	**4.2**
51.10	**51.10**	31.7	73.13	17.3	17.4
51.11	**51.11**	25.5	43.9	4.2	69.12
52.10	**51.10**	17.4	73.13	31.7	17.3
52.11	**51.11**	25.5	43.9	4.2	69.12
53.10	**51.10**	17.4	73.13	31.7	17.3
53.11	**51.11**	25.5	43.9	4.2	69.12
54.10	**51.10**	17.4	73.13	31.7	2.1
54.11	**51.11**	25.5	43.9	4.2	69.12
55.10	**51.10**	17.4	73.13	2.1	31.7
55.11	**51.11**	25.5	43.9	4.2	69.12
56.10	**51.10**	17.4	73.13	2.1	31.7
56.11	**51.11**	25.5	43.9	4.2	69.12
57.10	**51.10**	17.4	73.13	2.1	31.7
57.11	**51.11**	25.5	43.9	4.2	69.12
58.10	**51.10**	17.4	2.1	73.13	**25.6**
58.11	**51.11**	25.5	43.9	4.2	69.12
59.10	**51.10**	17.4	2.1	73.13	**25.6**
59.11	**51.11**	25.5	43.9	4.2	69.12
69.12	**69.12**	4.2	**78.15**	51.11	25.6
70.12	**69.12**	4.2	**78.15**	25.6	51.11
71.12	**69.12**	4.2	**78.15**	25.6	51.11
72.12	**69.12**	4.2	**78.15**	25.6	51.11
73.13	**73.13**	31.7	17.3	51.10	17.4
74.13	**73.13**	31.7	17.3	51.10	**2.1**
75.13	**73.13**	31.7	17.3	51.10	**2.1**
76.13	**73.13**	31.7	17.3	**2.1**	51.10
78.14	**78.14**	25.5	78.15	**25.6**	**4.2**
78.15	**78.15**	**69.12**	4.2	25.6	78.14
79.14	25.5	**78.14**	**4.2**	2.1	**25.6**
79.15	**78.15**	**69.12**	25.6	4.2	78.14
80.14	**78.14**	25.5	2.1	**25.6**	**4.2**
80.15	**78.15**	**69.12**	25.6	4.2	78.14
81.14	25.5	**78.14**	**25.6**	**4.2**	2.1
81.15	**78.15**	**69.12**	25.6	78.14	4.2
82.14	25.5	2.1	**25.6**	**4.2**	17.4
82.15	**78.15**	**69.12**	25.6	78.14	4.2
83.16	**83.16**	31.8	2.1	17.4	**4.2**
84.16	**83.16**	31.8	2.1	17.4	**4.2**

TAB. 2 – *Niveaux de gris avec modèle de mélange gaussien et critère BIC : les cinq objets les mieux appariés, ainsi que leur classement pour chacun des 97 objets de «violetta» . En gras sont indiqués les objets similaires à l'objet requête.*

- Quelque soit la méthode, la performance en niveau de gris est toujours supérieure à celle en RGB norm. (ou en RGB). Cela est étonnant puisque l'information en RGB norm. est plus importante qu'en niveaux de gris et doit donc permettre une meilleure reconnaissance. Cela est du au fait que certains objets similaires étaient bien appariés en niveaux de gris et que cet appariement est absent en RGB norm. Par exemple, la fille 4.2 à pull gris est parfois appariée avec la fille 25.6 similaire mais avec pull rouge, ce dans le cas des niveaux de gris. En RGB norm., le gris et le rouge sont bien distingués et l'appariement n'a plus lieu.

- En niveaux de gris, histo. 8 et gauss. (flou ou non) ICL ont une taille d'index du même ordre. Par contre, en RGB et RGB norm., gauss. (flou) ICL nécessite dix fois moins d'index que histo.8. Ce gain est considérable.

Ne retenant, pour les trois types de couleur, que le critère ICL (en gauss. et gauss. flou) et histo. 8, la figure 4 donne les valeurs de τ en fonction du rang r pour ces neuf situations. Cette figure appelle plusieurs remarques :

- La valeurs de $\tau(r)$ pour $r \leq 4$ est généralement plus grande pour histo. 8 que pour gauss. et gauss. flou ICL. Cela signifie que histo. 8 a un taux d'appariement meilleur dans les tous premiers rangs.

- Par contre, les valeurs de $\tau(r)$ pour $r > 4$ sont en faveur de gauss. et gauss. flou ICL. Cela signifie que gauss. et gauss. flou ICL ont un taux d'appariement meilleur lorsque l'on élargie le nombre de rangs. C'est pour cela que la moyenne des $\tau(r)$ était meilleure.

Concernant les vitesses d'indexation puis d'appariement, nous pouvons dire que :

- La vitesse d'indexation est beaucoup plus lente pour les méthodes gaussiennes que pour les histogrammes. L'écart est d'un facteur entre 60 et 100 dans nos expérimentations.

- La vitesse d'appariement est par contre du même ordre de grandeur. Les mélanges nécessitent juste le calcul d'une vraisemblance donc d'exponentielles. Les histogrammes nécessitent la transformation préalable de chaque objet requête en histogramme avant d'appliquer la distance du chi-deux.

		niv. gris		RGB		RGB norm.	
gauss.	BIC	0.7475	(0.1578)	0.7069	(0.1720)	0.7352	(0.1606)
	ICL	0.7473	(0.1674)	0.7117	(0.1725)	0.7451	(0.1548)
gauss. flou	BIC	0.7717	(0.1482)	0.7221	(0.1606)	0.7439	(0.1536)
	ICL	0.7589	(0.1474)	0.7295	(0.1634)	0.7515	(0.1478)
histo.	2	0.7148	(0.1503)	0.6769	(0.1628)	0.6796	(0.1591)
histo.	8	0.7564	(0.1604)	0.7101	(0.1693)	0.6693	(0.1706)
histo.	32	0.7403	(0.1600)	0.7098	(0.1605)	0.6695	(0.1705)

TAB. 3 – *Moyenne de la performance globale de chaque méthode (et entre parenthèses son écart-type).*

4 Conclusions

L'indexation de la distribution des couleurs par un mélange gaussien a le net avantage de choisir automatiquement le nombre de classes, par exemple par le critère ICL, et

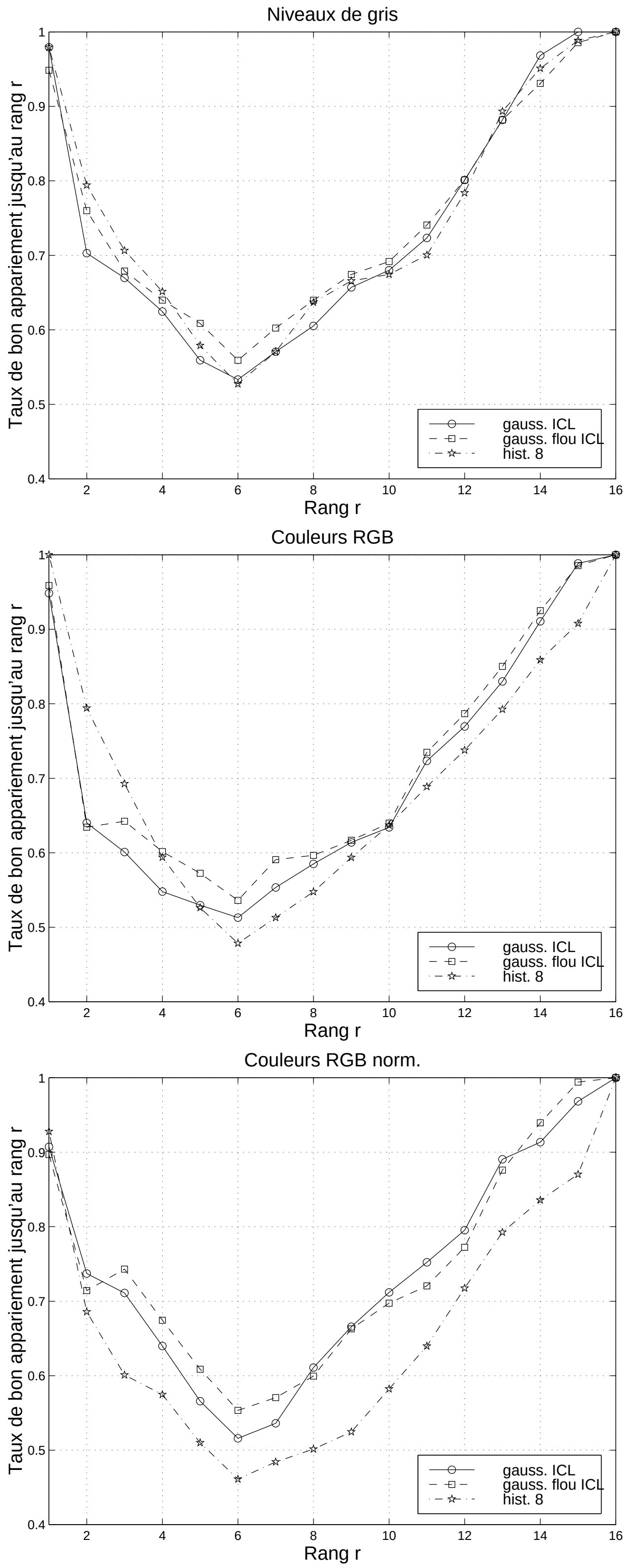


FIG. 4 – *Les valeurs du taux de bon appariement jusqu'au rang r pour les trois types de couleur et les méthodes gauss. ICL, gauss. flou ICL et histo. 8.*

de diminuer très fortement la taille de l'index dans le cas des couleurs RGB (la taille

		niv. gris		RGB		RGB norm.	
gauss.	BIC	15.12	(2.65)	57.75	(19.95)	58.37	(24.89)
	ICL	9.68	(5.01)	50.87	(17.96)	52.12	(22.12)
histo.	2	1	(0)	7	(0)	7	(0)
histo.	8	7	(0)	511	(0)	511	(0)
histo.	32	31	(0)	32767	(0)	32767	(0)

TAB. 4 – *Moyenne du nombre de degrés de liberté de chaque méthode (et entre parenthèses son écart-type).*

est du même ordre en niveaux de gris). L'inconvénient majeur est de nécessiter un temps d'indexation assez long. Le temps d'appariement, souvent plus crucial, est par contre similaire à celui des histogrammes.
Les qualités d'appariement sont légèrement supérieures pour les histogrammes en niveaux de gris, et plutôt en faveur des mélanges gaussiens dans le cas de la couleur RGB (normalisée ou non). En règle générale, les histogrammes donnent un meilleur appariement dans les tous premiers rangs, les mélanges gaussiens donnant un bon taux d'appariement meilleur jusqu'à des rangs plus élevés.
Une nette dégradation de la qualité d'appariement à été constatée dans le cas des histogrammes des couleurs RGB normalisées interne à cause d'une distribution des couleurs peu adaptée aux histogrammes. Les mélanges gaussiens étant beaucoup plus souples, ils ne rencontrent pas ce problème.
Une méthode de robustification des mélanges gaussiens au changement de luminosité a été proposée. Malgré son coté empirique bien qu'intuitif, cette technique a montré un très bon comportement sur nos exemples.
Une extension possible de ce travail est de considérer des signatures qui intègrent, en plus des couleurs, une information spatiale. On pourrait, par exemple, penser à approximer la distribution jointe de couleur et de gradient. Cela devrait améliorer le caractère discriminant des index et par conséquent leur pouvoir de bon appariement. Au regard des résultats précédents, il n'est pas exclu un meilleur comportement des mélanges gaussiens. En effet, ils ont montré de meilleurs résultats que les histogrammes quand la dimension augmente (ce qui est bien le cas ici) et ceci avec des tailles d'index de loin inférieurs.

Références

Anderson, T.W. (1971). An Introduction to Multivatiate Statistical Analysis. *Wiley: New York, Second Edition*

Biernacki, C., Celeux, G. and Govaert, G. (1998). Assessing a Mixture Model for Clustering with the Integrated Classification Likelihood. *Rapport de recherche INRIA 3521.*

Celeux, G. and Govaert, G. (1992). A Classification EM Algorithm for Clustering and two Clustering Versions. *Computational Statistics & Data Analysis*, **14**, 315-332.

Dempster, A. P., Laird N. M. and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data with the EM Algorithm (with discussion). *Journal of the Royal Statistical Society, Series B*, **39**, 1-38.

Gros, P., Mclean, G., Delon, R., Mohr, R., Schmid, C. and Mistler, G. (1997). Utilisation de la couleur pour l'appariement et l'indexation d'images. *Rapport de recherche INRIA 3269.*

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, **6**, 461-464.

Schiele, B. (1997). Reconnaissance d'Objets utilisant des Histogrammes Multidimensionnels de Champs Réceptifs. *Ph.D. Thesis, Institut National Polytechnique de Grenoble.*

Swain, M.J. and Ballard, D.H. (1991). Color Indexing. *International Journal of Computer Vision.* **7**, 11, 11-32.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399