

Un algorithme basé sur les meilleurs arbres de dépendances pour la détection de la peau dans les images couleurs.

S. El Fkihi^{1, 2}

M. Daoudi¹

D. Aboutajdine²

¹ GET/TELECOM Lille1, LIFL (UMR USTL-CNRS 8022)

Université des Sciences et Technologies de Lille
Cité Scientifique - Rue Guglielmo Marconi , Villeneuve d'Ascq , France

{elfkihi, daoudi}@telecom-lille1.eu

² Laboratoire GSCM-LRIT

Université Mohammed V-Agdal, faculté des sciences de Rabat
4 Avenue Ibn Battouta B.P. 1014 RP, Rabat, Maroc

aboutaj@fsr.ac.ma

Résumé

Dans cet article, nous présentons une approche originale pour la détection de la peau dans les images couleurs. Notre contribution est fondée sur les distributions des arbres qui trouvent leurs applications dans plusieurs domaines de recherche. L'originalité porte sur la combinaison des probabilités issues des Meilleurs-Arbres-Recouvrants pour l'approximation des lois de probabilités peau et non peau afin d'améliorer la prise de décision.

La base publique Compaq est utilisée pour évaluer les performances de la méthode proposée par rapport à deux autres méthodes : (1) celle de base supposant l'indépendance entre les pixels et (2) celle basée sur un Meilleur-Arbre-Recouvrant. Les mesures de la courbe ROC et de l'air sous celle-ci ont prouvé que par rapport aux méthodes étudiées, le nouveau modèle offre la meilleure détection de la peau .

Mots clefs

Arbre de dépendance, modèles probabilistes, détection de la peau, théorie des graphes.

1 Introduction

La détection de la peau consiste à détecter les pixels de peau dans une image couleur pour obtenir une image binaire. Elle constitue une étape de prétraitement fondamentale dans plusieurs applications telles que la détection de visage[1], la segmentation vidéo et le filtrage des images sur le Web [2][3][4]. Cependant, c'est une tâche difficile à réaliser à cause de la variation de la couleur de la peau (européen, africaine, ...) et la diversité des conditions de prise de vue (lumière, bruit, ...) qui créent une classe de la peau avec beaucoup de variations inter-classe. Ainsi, plusieurs recherches basées aussi bien sur des ap-

proches paramétriques que sur d'autres non paramétriques, ont été menées dans le cadre de cette application (les histogrammes de couleurs, le mélange Gaussien [5], ...).

Dans ce travail, nous considérons le modèle probabiliste qui représente la loi a posteriori $p(y_s = 1|x)$, où x représente un bloc de taille 3×3 et y_s représente l'état du pixel central s . y_s peut prendre deux valeurs selon l'état du pixel s , $y_s = 1$ si s est un pixel peau et $y_s = 0$ sinon. Nous proposons dans ce papier un modèle probabiliste porté par des arbres qui approximent en même temps les deux lois a posteriori $p(y_s = 1|x)$ et $p(y_s = 0|x)$.

Nicu et al. ([6] et [7]) proposent un réseau bayésien pour approximer une loi a posteriori. Les auteurs se sont basés sur les arbres de Chow et Liu [8] pour la détection de la peau. Il s'agit d'un arbre de recouvrement ou de dépendance optimal au sens de l'information mutuelle. Cet arbre passe par tous les noeuds du graphe et maximise le score défini pour tous les arcs possibles (MWST ou Maximum Weight Spanning Tree). Il est prouvé que la probabilité portée par cet arbre est une meilleure approximation de la vraie probabilité.

Dans [9] nous avons présenté une heuristique pour la détection de la peau, nommée Meilleur-Arbre-Recouvrant ou meilleur arbre de dépendance. Cet arbre est basé sur l'algorithme des MWSTs et modélise deux distributions de probabilité à la fois. Ensuite, dans [10], nous avons présenté un algorithme exact pour trouver un Meilleur-Arbre-Recouvrant (voir section 2), dans le cadre de la même application.

Un système de détection de peau n'est jamais parfait et les utilisateurs utilisent différents critères pour l'évaluation des performances d'un algorithme de détection de la peau. L'aspect général des zones de peau détectées, ou d'autres critères globaux pourraient être importants pour une utili-

sation ultérieure. Pour l'évaluation quantitative, nous emploierons les taux des faux positifs (proportion des pixels non peaux interprétés comme peaux) et des vrais positifs (proportion des pixels peaux correctement interprétés). L'utilisateur pourrait souhaiter combiner ces deux indicateurs. Par conséquent nous proposons un système où le rendement n'est pas binaire mais un nombre flottant entre zéro et un, plus la valeur est grande, plus la croyance pour un pixel de peau est grande. L'utilisateur peut alors appliquer un seuil pour obtenir une image binaire. Des taux d'erreur pour tous les seuils sont récapitulés dans la courbe appelée la courbe ROC.

Dans la suite, nous présentons une formulation mathématique du problème étudié. Nous décrivons dans la section 3 la nouvelle approche que nous proposons dans cet article. Dans la section 4 nous fournissons les résultats de la comparaison de notre approche par rapport à d'autres modèles.

2 Formulation du problème

Considérons un vecteur d'observations $x = (x_1, \dots, x_n)$ issu d'un bloc de pixels et un système de voisinage qui détermine la structure des interactions locales dans la distribution des images. Notons respectivement C_1 et C_0 les deux classes peau et non peau représentées par la variable aléatoire y .

La probabilité jointe $P(x, y)$ est inconnue, en revanche nous possédons une base des images segmentées $\{(x^{(1)}, y^{(1)}), \dots, (x^{(N)}, y^{(N)})\}$; $x^{(N)}$ est l'image et $y^{(N)}$ son masque associé. Pour chaque pixel s , nous avons besoin de calculer $P(y_s = 1|x)$ qui est la probabilité d'avoir un pixel de la classe C_1 sachant les valeurs de pixels appartenant au bloc centré sur s . Ainsi, nous désirons définir un classifieur probabiliste représentant les probabilités a posteriori, $P(y_s = 0|x)$ et $P(y_s = 1|x)$, et basé sur les distributions d'arbres probabilistes.

Afin de simplifier les notations nous posons $P(y_s = 0|x) = q(x)$ et $P(y_s = 1|x) = p(x)$.

Désignons par $G(N, A)$ le graphe non orienté correspondant à x dont les noeuds (N) sont les éléments de ce vecteur et les arcs (A) sont déterminés par le système de voisinage choisi. Dans le cas où G est un arbre noté T , la probabilité approximée par cette structure s'écrit sous la forme [11] :

$$P(x|T) = \prod_{(u \sim v) \in T} \frac{P_{uv}(x_u, x_v)}{P_u(x_u)P_v(x_v)} \prod_{x_u \in N} P_u(x_u) \quad (1)$$

avec $P_u(x_u)$ et $P_{uv}(x_u, x_v)$ sont les marginaux de $P(x|T)$ l'approximation de $P(x)$ par l'arbre T . u et v sont deux noeuds du graphe. Ils sont notés $u \sim v$ s'ils sont voisins. Nous définissons d'abord un classifieur de Meilleur-Arbre-Recouvrant.

Définition 1 *Le classifieur de Meilleur-Arbre-Recouvrant est un système de classification basé sur un arbre performant pour deux classes complémentaires. Ici perfor-*

mant indique que les deux probabilités modélisant les deux classes sont optimales.

L'optimalité dans ce cas est calculée au sens de la divergence de Kullback-Leibler [12] entre les deux distributions de probabilités à approximer. Ainsi, plus la divergence entre ces deux probabilité est grande plus les résultats sont optimums comme décrit le théorème suivant :

Théorème 1 *Les probabilités $p(x|T)$ et $q(x|T)$ sont optimums si seulement si l'arbre T est un MWST dont le poids maximise la formule suivante :*

$$\sum_{(u \sim v) \in T} \{KL(p_{uv}, q_{uv}) - KL(p_u, q_u) - KL(p_v, q_v)\} \quad (2)$$

avec KL la divergence de Kullback-Leibler. La preuve de ce théorème est donnée dans [10].

Les MWSTs ont été appliqués dans plusieurs domaines de recherche sous l'onglet de l'optimisation. Toutefois, la solution offerte par cette approche n'est pas toujours unique. En effet, considérons le cas le plus simple d'un graphe dont les pondérations des arcs sont identiques. Tout arbre recouvrant est un MWST. Or dans la plupart des cas, ce qui est demandé c'est l'arbre 'idéal' qui approxime en mieux la probabilité réelle et non le MWST.

Par conséquent, un problème qui s'impose est la multitude de choix entre tous les Meilleurs-Arbres-Recouvrants possibles.

Dans un travail précédent, nous avons proposé le mélange des arbres de recouvrement optimaux[13] dans le but de déterminer un arbre de recouvrement 'idéal' qui modélise la probabilité du mélange :

$$P_{mix}(x|T_{mix}) = \sum_{T \in S(G)^k} \lambda(T)P(x|T) \quad (3)$$

où

- $P_{mix}(x|T_{mix})$: la probabilité du mélange approximée par l'arbre du mélange T_{mix} ;
- $\lambda(T)$: le coefficient du mélange associé à l'arbre T . Avec $\sum_{T \in S(G)^k} \lambda(T) = 1$ et $\lambda(T) \geq 0$
- $S(G)^k$: l'ensemble de tout les arbres recouvrant optimaux du graphe G ; dont le cardinal est égal à k .

La loi de probabilité résultante offre une meilleure approximation de la vraie loi de probabilité d'une classe donnée. Cela est justifié théoriquement par la formule : ($\forall j = 1, 2, \dots, k$)

$$KL(P(x), P_{mix}(x|T_{mix})) \leq KL(P(x), P(x|T_j)) \quad (4)$$

avec $P(x)$ la vraie loi de probabilité et $P(x|T_j)$ la probabilité approximée par l'arbre T_j .

En s'inspirant de ce travail, nous avons souhaité réaliser de la même manière le mélange des Meilleurs-Arbres-Recouvrants. Pour cela, soient :

- $j_0 \in \{1, 2, \dots, k\}$, avec k le nombre de Meilleurs-Arbres-Recouvrants trouvés ;
 - p_{mix} et q_{mix} les probabilités des mélanges des deux classes C_1 et C_0 respectivement, vérifiant l'équation (3) ;
 - p_{j_0} et q_{j_0} les probabilités de C_1 et C_0 portées par le Meilleur-Arbre-Recouvrant T_{j_0} .
- nous avons comparé la divergence de Kullback entre p_{mix} et q_{mix} à celle entre p_{j_0} et q_{j_0} , nous avons trouvé :

$$KL(p_{mix}, q_{mix}) \leq KL(p_{j_0}, q_{j_0}) \quad (5)$$

du moment que l'objectif des Meilleurs-Arbres-Recouvrants est de maximiser la divergence entre les deux probabilités approximées, nous déduisons qu'un tel mélange d'arbre n'offrira pas de meilleurs résultats par rapport au modèle basé sur un Meilleur-Arbre-Recouvrant. La preuve de l'inégalité (5) est présentée dans l'annexe. Dans la section suivante nous proposons une nouvelle méthode pour combiner les Meilleurs-Arbres-Recouvrants en conservant les informations portées par chaque arbre considéré.

3 Notre approche

La méthode que nous proposons est composée de deux étapes importantes : (1) sélection des Meilleurs Arbres Recouvrants, et (2) combinaison de ces arbres. La procédure (1) suivante permet de déterminer tous les meilleurs arbres de recouvrement possibles.

Procédure 1 *Les meilleurs arbres de recouvrement*

- Entrée : Donnée $\{(x^{(1)}, y^{(1)}), \dots, (x^{(n)}, y^{(n)})\}$.
- Étapes :
 1. Définir le système de voisinage à considérer.
 2. Construire le vecteur d'observation x issu de chaque bloc de voisinage.
 3. Construire $G(\mathcal{N}, \mathcal{A})$ le graphe non orienté correspondant à x ; avec
 - \mathcal{N} : l'ensemble des noeuds. Chaque élément de x est un noeud.
 - \mathcal{A} : l'ensemble des arcs déduit par le système de voisinage utilisé.
 4. pour chaque couple de noeuds u et v du graphe, utiliser un estimateur classique pour calculer les probabilités empiriques : p_u et p_{uv} (resp. q_u et q_{uv}) de p (resp. q) :

$$p_u(x_u = i) = f_i^1(x_u) \quad , \quad q_u(x_u = i) = f_i^0(x_u)$$

$$p_{uv}(x_u = i, x_v = j) = f_{ij}^1(x_u, x_v)$$

$$q_{uv}(x_u = i, x_v = j) = f_{ij}^0(x_u, x_v)$$

où pour $m = 1, 0$, $f_{ij}^m(x_u, x_v)$ est la fréquence jointe de $x_u = i$ et $x_v = j$ étiquetées m . Et $f_i^m(x_u)$ est la fréquence de $x_u = i$ étiquetée m .

5. Calculer le coût des arcs défini par :

$$KL(p_{uv}, q_{uv}) - KL(p_u, q_u) - KL(p_v, q_v) \quad \forall u \sim v$$

6. Utiliser l'algorithme de *K-Best-Spanning-Trees* [14].

- Sortie : Meilleurs-Arbres-Recouvrants $(T_j, j = 1, 2, \dots, k)$.

A l'issue de cette procédure, utilisée dans la phase d'apprentissage, les meilleurs arbres approximant les deux vraies distributions de probabilités des classes C_1 et C_0 sont déterminées. Afin de conserver les informations apportées par chaque arbre, lors du calcul des inférences, nous proposons de combiner ces diverses sources de probabilités afin d'améliorer la prise de décision.

Pour cela, nous faisons appel dans un premier temps à un algorithme de vote. Pour chaque arbre T_j , nous calculons $P_j(y|x)$ que nous comparons par rapport à un seuil s_1 . Ce dernier paramètre permet de déterminer la probabilité d'appartenance d'un vecteur x à la classe C_1 ; il est calculé dans la phase d'apprentissage, lors de la construction des histogrammes correspondants aux classes C_1 et C_0 . La comparaison permet de partager les différentes probabilités calculées en deux ensembles : (1) ensemble correspondant à la classe C_1 et (2) ensemble correspondant à la classe C_0 . Le plus grand cardinal des deux ensembles permet de déterminer la classe dominante, soit C_d (égale à C_0 ou C_1). Dans un deuxième temps, nous calculons la probabilité $P(y|x)$ qui sera affectée au pixel central du bloc représenté par le vecteur x . Nous choisirons cette probabilité égale au maximum des probabilités désignant la classe C_d dans le cas où $C_d = C_1$. Dans le cas où $C_d = C_0$, la probabilité affectée au pixel central sera égale au minimum des probabilités désignant la classe C_0 .

4 Résultats Expérimentaux

Nos expériences sont effectuées sur la base d'images publique Compaq [15]. C'est une collection d'environ 18696 images couleurs et de leurs masques segmentées manuellement. Nous décomposons aléatoirement cette base en deux parties. La première correspond à une base d'apprentissage constituée d'environ 2 milliards de pixels, et sert à calculer les lois empiriques. La seconde partie de la base correspond à la base test et permet d'évaluer les performances de notre approche.

Pour un pixel s , nous considérons le système de voisinage suivant :

$$V_s = \{(i, j) / |i - i_s| < 2, |j - j_s| < 2\} \setminus \{(i_s, j_s)\} \quad (6)$$

avec i et j les coordonnées d'un pixel dans une image.

La base de données Compaq est assez grande, ainsi nous considérons l'espace de couleur RGB, et nous utilisons dans la phase d'apprentissage des histogrammes à 32 couleurs. Chaque pixel sera décomposé en ses trois composantes pour construire un vecteur d'observation de taille égale à 27.

La procédure (1) a permis de détecter trois Meilleurs-Arbres-Recouvrants. Les résultats obtenus par le modèle basé sur la combinaison des arbres résultant, sont présentés

dans la figure 1. L'entrée du modèle est une image couleur et la sortie est une image en niveaux de gris.



Figure 1 – Les entrées et les sorties du modèle proposé.

Nous comparons ce modèle avec deux autres modèles : (1) le modèle de base ultra simplifié qui suppose l'indépendance entre les pixels [16] et (2) le modèle basé sur un Meilleur-Arbre-Recouvrant. Notons que l'arbre utilisé dans ce dernier modèle est déterminé à partir de l'algorithme donné dans [10]; c'est le premier Meilleur-Arbre-Recouvrant trouvé.

Dans le but d'une mesure quantitative de notre approche par rapport aux deux autres modèles étudiés, nous proposons d'étudier les courbes ROC (Receiver Operating Characteristic) des différents modèles. Nous combinons ainsi les faux positifs et les vrais positifs.

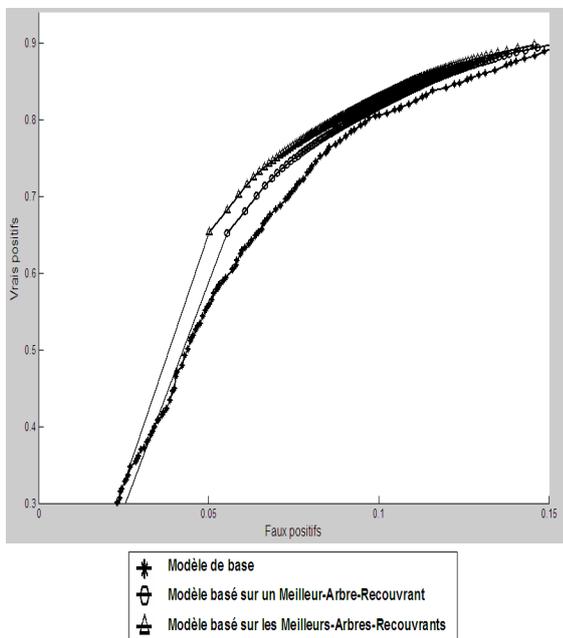


Figure 2 – Les courbes ROC des trois modèles étudiés.

La figure 2 représente les différentes courbes. Nous remarquons que la courbe en triangle qui représente le modèle

basé sur la combinaison des Meilleurs-Arbres-Recouvrants est toujours au dessus de la courbe du modèle basé sur le Meilleur-Arbre-Recouvrant (courbe en cercle) quelque soit le taux des faux positifs. A partir d'un taux de faux positifs égal à 2.5% la courbe du modèle proposé reste toujours supérieure à celle du modèle de base (courbe en croix). Nous déduisons que le nouveau modèle donne la meilleure détection de la peau.

Une manière d'estimer l'efficacité globale des différents modèles étudiés consiste à calculer l'aire normalisée sous les courbes ROC (Area Under Curve ou AUC). Ainsi, dans l'intervalle $[0 \quad 0.15]$, la valeur du AUC est égale à 0.0493 pour le modèle de base, à 0.1177 pour le modèle basé sur le Meilleur-Arbre-Recouvrant, et à 0.1227 pour le nouveau modèle. Ainsi, une fois de plus, il est prouvé que la combinaison des Meilleurs-Arbres-Recouvrants présente de meilleurs résultats par rapport aux deux autres méthodes. Les objets ayant une couleur qui ressemble à celle de la peau présentent quelques anomalies du nouveau modèle (figure 3).

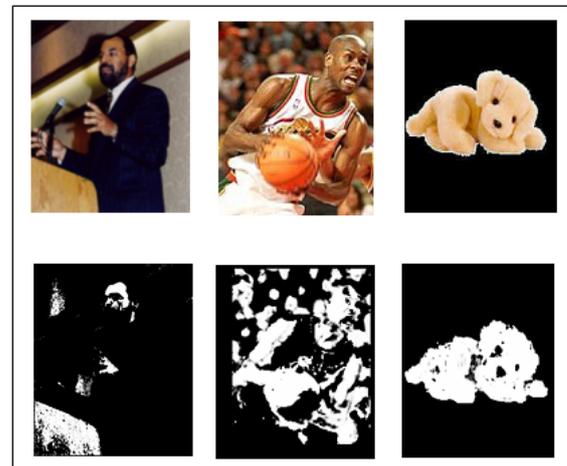


Figure 3 – Quelques anomalies du nouveau modèle.

5 Conclusion

Nous avons proposé une méthode originale pour l'approximation des lois jointes dans le cas de la détection de la peau dans les images couleurs. Cette approximation est fondée sur les Meilleurs-Arbres-Recouvrants dédiés à la modélisation des classes peau et non peau. Nous avons proposé une combinaison de ces arbres, permettant d'exploiter les informations portées par chacun d'eux.

Le modèle de détection de la peau proposé dans cet article a présenté de meilleurs résultats, dans la base Compaq, par rapport au modèle de base et celui fondé sur un Meilleur-Arbre-Recouvrant. En effet, la courbe ROC du nouveau modèle reste toujours supérieure à celle du modèle basé sur un seul arbre quelque soit le taux des faux positifs considéré. Aussi, par rapport au modèle de base, le nouveau modèle présente une amélioration à partir d'un taux

de faux positifs égal à 2.5%. Les mesures de l'AUC permettent d'approuver ce constat.

Notre approche peut être généralisée pour traiter le problème d'approximation des vraies lois de probabilités de deux classes complémentaires.

Annexe

Nous montrons l'inégalité (5). Nous avons :

$$KL(p_{mix}, q_{mix}) = \sum_{i=1}^k p_{mix}(x_i) \log \left(\frac{p_{mix}(x_i)}{q_{mix}(x_i)} \right) \quad (7)$$

$$= \sum_{i=1}^N \left(\sum_{j=1}^k \lambda_j p_j(x_i) \right) \log \left(\frac{\sum_{j=1}^k \lambda_j p_j(x_i)}{\sum_{j=1}^k \lambda_j q_j(x_i)} \right) \quad (8)$$

Puisque les termes $\lambda_j p_j(x_i)$ et $\lambda_j q_j(x_i)$ sont non-négatifs ($i = 1, \dots, N; j = 1, \dots, k$), selon le théorème de Log-sum-inequality, nous obtenons :

$$\begin{aligned} \sum_{i=1}^N \left(\left(\sum_{j=1}^k \lambda_j p_j(x_i) \right) \log \left(\frac{\left(\sum_{j=1}^k \lambda_j p_j(x_i) \right)}{\left(\sum_{j=1}^k \lambda_j q_j(x_i) \right)} \right) \right) \\ \leq \sum_{i=1}^N \left(\sum_{j=1}^k \lambda_j p_j(x_i) \log \frac{\lambda_j p_j(x_i)}{\lambda_j q_j(x_i)} \right) \quad (9) \end{aligned}$$

$$= \sum_{j=1}^k \left(\lambda_j \sum_{i=1}^N p_j(x_i) \log \frac{p_j(x_i)}{q_j(x_i)} \right) \quad (10)$$

$$= \sum_{j=1}^k (\lambda_j KL(p_j, q_j)) \quad (11)$$

Soit $j_0 \in \{1, 2, \dots, k\}$, par définition nous avons :

$$\forall j \in \{1, 2, \dots, k\} \quad KL(p_{j_0}, q_{j_0}) = KL(p_j, q_j) \quad (12)$$

donc

$$KL(p_{mix}, q_{mix}) \leq \sum_{j=1}^k (\lambda_j KL(p_{j_0}, q_{j_0})) \quad (13)$$

d'où inégalité (5).

Références

- [1] Jean-Christophe Terrillon, M. N. Shirazi, H. Fukumachi, et S. Akamatsu. Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. In *Fourth International Conference On Automatic Face and gesture Recognition*, pages 54–61, 2000.
- [2] James Z. Wang, Jia Li, Gio Wiederhold, et Oscar Firschein. Classifying objectionable websites based on image content. *Notes in Computer Science, Special issue on interactive distributed multimedia systems and telecommunication services*, 21/15 :113–124, 1998.
- [3] M.M. Fleck, D.A. Forsyth, et C. Bregler. Finding naked people. In *Proc. European Conf. on Computer Vision*, pages 593–602. B. Buxton, R. Cipolla, Springer-Verlag, Berlin, Germany, 1996.
- [4] H. Zheng, M.Daoudi, et B. Jedynek. Blocking adult images based on statistical skin detection. *Electronic Letters on Computer Vision and Image Analysis*, 4(2) :1–14, Décembre 2004.
- [5] L.M. Bergasa, M.Mazo., A. Gardel, M.A. Sotelo, et L. Boquete. Unsupervised and adaptive gaussian skin-color model. *Image and Vision Computing*, 18 :987–1003, 2000.
- [6] N. Sebe, I. Cohen, Thomas S. Huang, et T. Gevers. Skin detection : A bayesian network approach. *International Conference On Pattern Recognition*, Août 2004.
- [7] L.S Chen A. Garg I. Cohen, N. Sebe et T.S. Huang. Facial expression recognition from video sequences : Temporal and static modeling. In *CVIU special issue on face recognition*, 2003.
- [8] C. K. Chow et C. N. Liu. Approximating discrete probability distributions with dependence trees. *IEEE Transactions on Information Theory*, (IT-14 (3)) :462–467, 1968.
- [9] S. Elfkihi, M. Daoudi, et D. Aboutajdine. A tree distribution for skin detection. *The Second International Symposium on Communications, Control and Signal Processing (ISCCSP'06)*, Mars 2006.
- [10] S. Elfkihi, M. Daoudi, et D. Aboutajdine. *Probability Approximation Using Best-Tree Distribution for Skin Detection*, volume 4179, pp767-775 of *Lecture Notes in Computer Science*. Advanced Concepts for Intelligent Vision Systems (ACIVS 2006).
- [11] J. Pearl. *Probabilistic Reasoning in intelligent systems : networks of plausible inference*. Morgan Kaufmann, 1988.
- [12] T. M. Cover et J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, Août 1991.
- [13] S. Elfkihi, M. Daoudi, et D. Aboutajdine. Optimal spanning trees mixture based probability approximation for skin detection. *2nd International Conference on Computer Vision Theory and Applications*, Mars 2007.
- [14] T. Ibaraki N. Katoh et H. Mines. An algorithm for finding k minimum spanning trees. *Society for Industrial and Applied Mathematics*, 10/2, 2000.
- [15] M.J. Jones et J. M. Rehg. Statistical color models with application to skin detection. In *Computer Vision and Pattern Recognition*, pages 274–280, 1999.
- [16] B. Jedynek, H. Zheng, et M.Daoudi. Skin detection using pairwise models. *Image and Vision Computing*, 23(13) :1122–1130, Novembre 2005.