

Modélisation de la teinte du visage par fusion d'informations couleur dans le cadre de la théorie de Dempster-Shafer

Francis Faux, Franck Luthon

▶ **To cite this version:**

Francis Faux, Franck Luthon. Modélisation de la teinte du visage par fusion d'informations couleur dans le cadre de la théorie de Dempster-Shafer. CORESA'05, Nov 2005, Rennes, France. pp.;. hal-00408647

HAL Id: hal-00408647
https://hal.science/hal-00408647

Submitted on 19 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modélisation de la teinte du visage par fusion d'informations couleur dans le cadre de la théorie de Dempster-Shafer

F. Faux F. Luthon

Laboratoire LIUPPA (EA 3000)
IUT Informatique, Château Neuf, Place Paul Bert, 64100 Bayonne, France
faux@iutbayonne.univ-pau.fr

Résumé

Dans cet article, la modélisation de la teinte chair d'un visage sous la contrainte du temps réel est réalisée par un processus de fusion pixel dans le cadre de la théorie de l'évidence de Dempster-Shafer. Pour cela, une phase d'initialisation simple et rapide prend en compte au plus près la réalité terrain. Pour initialiser le traitement, l'utilisateur sélectionne manuellement sur une image, une zone ombrée (cou), une zone saturée (front, nez) et une zone d'intensité moyenne. Pour chaque zone, trois modèles a priori sont déterminés à partir des densités de probabilités de trois sources couleur indépendantes du point de vue cognitif. A partir de ces modèles des fonctions de masse sont élaborées puis fusionnées. La prise en compte de données contextuelles spécifiques au visage, la complémentarité des sources couleur et la fusion dans le cadre de la théorie de l'évidence améliorent la détection et diminuent la sensibilité face aux conditions d'éclairement.

Mots Clefs

Couleur de peau, détection de visage, Dempster-Shafer, fusion.

Abstract

In this paper, a real time skin hue model is built by a pixel fusion process within the framework of the Demster-Shafer theory. In a simple and fast initialising stage, the user selects successively on an image, a shadowy zone, an overexposed zone and a zone of mean intensity of the face. For each zone, three a priori models are determined from three cognitive independent colour sources densities of probability. From these models, mass sets are elaborated and then fused. The taking into account of contextual variables specific to the face under study, the colour source cognitive complementarity and the fusion process within the framework of the evidence theory improve the detection quality and decrease the sensitivity to lighting conditions.

Keywords

Skin hue, face detection, Dempster-Shafer, fusion.

1 Introduction

La détection d'un visage dans une image ou une séquence vidéo est nécessaire à de nombreuses applications telles que l'interaction homme machine (IHM), la reconnaissance, l'identification, la visioconférence, la robotique ou la télésurveillance.

Plus de 150 méthodes de détection [1, 2], allant des techniques dites de bas niveau utilisant des primitives telles que la texture, les contours, le mouvement, la couleur [3, 4] jusqu'aux approches de haut niveau telles que les modèles d'apparence, les réseaux de neurones ou les SVM (Support Vector Machines) [5, 6], sont proposées dans la littérature. Dans toutes ces approches, la couleur de peau est une primitive souvent utilisée en tant que première estimation de localisation et de segmentation pour réduire la zone de recherche. Les informations couleur sont incontestablement pertinentes pour l'analyse de visage en vidéo car elles sont très spécifiques et permettent des algorithmes rapides invariants aux changements d'orientation et d'échelle. Cependant une difficulté importante est de s'accommoder aux variations des conditions d'éclairement. Une solution est d'utiliser des espaces couleur invariants vis à vis de la luminance. Les plus employés sont les espaces RGB normalisés ou l'espace TLS. Mais ces transformations couleur demeurent sensibles aux conditions d'éclairement et très bruitées dans les zones ombrées. Afin de pallier ces inconvénients, augmenter la robustesse et le contraste, les informations couleur de peau sont représentées dans l'espace logarithmique couleur LUX [7] dans cet article. Les modélisations existantes au niveau du pixel de la couleur de peau peuvent être regroupées en méthodes explicites, qui utilisent des règles de décision adéquates de manière empirique, et statistiques [8]. Ces dernières se décomposent en modélisations paramétriques et non paramétriques. Les approches paramétriques représentent la distribution de couleur sous la forme d'un modèle gaussien, parfois elliptique ou utilisent des mélanges de gaussiennes pour traduire le caractère multimodal de la distribution de couleur de peau. Les approches non paramétriques, indépendantes vis à vis de la forme de la distribution de la teinte, utilisent des histogrammes pour représenter la densité de probabilité dans l'espace couleur. Cependant les in-

formations à modéliser ne sont jamais parfaites en raison du principe même de l'acquisition d'images (passage de l'espace 3D au plan 2D), des limites du capteur, du bruit et ces imperfections prennent différentes formes incluant principalement l'ambiguïté, l'imprécision, l'incertitude ou l'incomplétude. Afin d'améliorer la modélisation il est important de prendre en compte ces imperfections. Dans le contexte d'éclairage non supervisé et de fond complexe, le visage est un objet non rigide dont la particularité est de présenter des zones ombrées (cou) ou surexposées (front, nez) dont la localisation est variable essentiellement en raison du mouvement dans une séquence vidéo. Ces effets d'ombre et effets de réflectance spéculaire rendent la modélisation complexe. Les modélisations probabilistes classiques (règle d'inférence bayésienne) présentent une perte de performance lorsque l'apprentissage au préalable n'est plus représentatif des mesures réelles. Dans cet article, la méthode de Dempster-Shafer [9, 10] a été introduite pour faire face à cet inconvénient en intégrant le contexte d'acquisition sous la forme de degrés de confiance. Aussi, cette approche permet de tenir compte à la fois de l'incertitude, de l'imprécion et de gérer le conflit. C'est pourquoi, une modélisation de la teinte du visage par un processus de fusion au niveau du pixel de données hétérogènes de teinte chair dans le cadre de la théorie de l'évidence est présentée dans cet article. Une phase d'initialisation simple et rapide prend en compte au plus près la réalité terrain. L'utilisateur sélectionne manuellement sur une image, une zone ombrée (cou), une zone saturée (front, nez) et une zone d'intensité moyenne. Pour chaque zone, trois modèles a priori sont déterminés à partir des densités de probabilités de trois sources couleur dans l'espace couleur LUX. A partir de ces modèles des fonctions de masse sont élaborées. Dès lors le processus de fusion combine ces masses dans le cadre de la théorie de l'évidence tout en intégrant des variables contextuelles et en considérant la fiabilité des capteurs par des degrés de confiance. Cette modélisation permet de segmenter l'image en régions de teinte chair ("blobs"). L'article est organisé comme suit. Le chapitre 2 est consacré à un rappel sur la théorie de l'évidence, le chapitre 3 développe le processus de modélisation, le chapitre 4 présente les résultats. Les évolutions et perspectives de la méthode font office de conclusion au chapitre 5.

2 Théorie de l'évidence

La théorie de l'évidence de Dempster-Shafer (DS) a été introduite par Dempster et formalisée par Shafer [9, 10]. Elle représente à la fois l'imprécision et l'incertitude à l'aide de fonctions de masse m, de plausibilité Pl et de croyance Bel. Cette théorie se décompose en trois étapes : la définition des fonctions de masse, la combinaison d'informations et la décision.

2.1 Définition des fonctions de masse

L'ensemble des hypothèses pour une source (typiquement une classe dans un problème de classification multisource) est défini sur l'espace $\Omega = \{\omega_1, \omega_2, ..., \omega_k, ...\omega_N\}$ appelé espace de discernement où ω_k désigne une hypothèse en faveur de laquelle une décision peut être prise.
Les fonctions de masse sont définies sur tous les sous-ensembles de l'espace Ω et non seulement sur les singletons comme dans les probabilités.
Une fonction de masse m est définie comme une fonction de 2^Ω dans [0,1]. En général on impose $m(\emptyset) = 0$ et une normalisation de la forme :

$$\sum_{A \subseteq \Omega} m(A) = 1$$

Une fonction de croyance Bel est une fonction totalement croissante de 2^Ω dans [0,1] définie par :

$$\forall A_1 \in 2^\Omega, ..., A_K \in 2^\Omega,$$
$$Bel(\cup_{i=1...K} A_i) \geq \sum_{I \subseteq \{1...K\}, I \neq 0} (-1)^{|I|+1} Bel(\cap_{i \in I} A_i)$$

où $|I|$ désigne le cardinal de I et $Bel(\emptyset) = 0$, $Bel(\Omega) = 1$.
Etant donné une fonction de masse m, la fonction Bel définie par :

$$\forall A \in 2^\Omega, Bel(A) = \sum_{B \subseteq A, B \neq 0} m(B)$$

est une fonction de croyance.
Inversement, à partir d'une fonction de croyance Bel, on peut définir une fonction de masse m par :

$$\forall A \in 2^\Omega, m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} Bel(B)$$

Une fonction de Plausibilité Pl est également une fonction de 2^Ω dans [0,1] définie par :

$$\forall A \in 2^\Omega, Pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

La plausibilité mesure la confiance maximum que l'on peut avoir en A.
La possibilité d'affecter des masses aux hypothèses composées et donc de travailler sur 2^Ω plutôt que sur Ω constitue un des avantages de cette théorie. Elle permet une modélisation très riche et très souple, en particulier de l'ambiguïté ou de l'hésitation entre classes.

2.2 Combinaison évidentielle

En présence de plusieurs capteurs ou de plusieurs informations provenant d'un même capteur, il devient intéressant de combiner les connaissances de chaque source pour en extraire une connaissance globale afin d'améliorer la prise de décision. Dans la théorie de DS, les masses sont combinées par la somme orthogonale de Dempster.
Soit m_j la fonction de masse associée à la source j, pour un sous-ensemble A de Ω on obtient :

$$(m_1 \oplus ... \oplus m_l)(A) = \frac{\sum_{B_1 \cap ... \cap B_l = A} m_1(B_1)...m_l(B_l)}{1 - \sum_{B_1 \cap ... \cap B_l = \emptyset} m_1(B_1)...m_l(B_l)} \quad (1)$$

Ce type de combinaison qui n'est pas idempotente suppose l'indépendance cognitive des sources plutôt que l'indépendance statistique [11].
Le mode de combinaison disjonctif est aussi possible en remplaçant l'intersection dans la formule (1) par une opération ensembliste :

$$(m_1 \oplus_\cup ... \oplus_\cup m_l)(A) = \sum_{B_1 \cup ... \cup B_l = A} m_1(B_1)...m_l(B_l) \quad (2)$$

2.3 Processus de décision

Contrairement à la théorie Bayésienne où le critère de décision est très souvent le maximum de vraisemblance, la théorie de l'évidence propose de nombreuses règles. Les plus utilisées sont le maximum de crédibilité, le maximum de plausibilité, les règles basées sur l'intervalle de confiance, le maximum de probabilité pignistique [12] et la décision par maximum de vraisemblance.

3 Description de la méthode

3.1 Sources et champ de discernement

Les informations couleur de peau sont représentées dans l'espace logarithmique couleur LUX. Cet espace couleur non linéaire est basé sur une transformation logarithmique (modèle LIP de Jourlin et al.) [7].
Les expressions des composantes LUX à partir de l'espace couleur RGB codé sur 3×8 bits sont données par :

$$L = (R+1)^{0.3}(G+1)^{0.6}(B+1)^{0.1} - 1$$
$$U = \begin{cases} 128\left(\frac{L+1}{R+1}\right) \text{ pour } R > L \\ 256 - 128\left(\frac{R+1}{L+1}\right) \text{ sinon} \end{cases}$$
$$X = \begin{cases} 128\left(\frac{L+1}{B+1}\right) \text{ pour } B > L \\ 256 - 128\left(\frac{B+1}{L+1}\right) \text{ sinon} \end{cases}$$

La procédure de fusion utilise 3 sources ("capteurs" couleurs) notées S_j, $(j = 1, 2, 3)$ telles que :

$$S_1 = U,\ S_2 = X \text{ et } S_3 = 0.5(L+U) = W \quad (3)$$

Pour la source S_3, on ajoute à la composante de luminance L, très riche du point de vue sémantique, la source S_1 afin de mieux caractériser les variations de teinte chair dues aux conditions d'éclairement. Chaque source S_j fournit une mesure notée M_j.
Le champ de discernement $\Omega = \{\omega_1, \omega_2\}$ est défini par 2 hypothèses ω_1 et ω_2 telles que : ω_1 représente l'hypothèse visage et ω_2, complément de ω_1, symbolise le fond ($\omega_2 = \bar{\omega}_1$).

3.2 Modèle a priori

Afin de déterminer le modèle a priori, lors d'une phase d'initialisation, l'utilisateur sélectionne respectivement trois zones caractéristiques du visage : une zone ombrée, une zone d'éclairement d'intensité moyenne et une zone surexposée.

Trois variables contextuelles z_i , $(i = 1, 2, 3)$ sont prises en considération : zone ombrée z_1 ; zone moyenne z_2 ; zone surexposée z_3.
Les histogrammes calculés sur chaque zone sélectionnée et pour chaque mesure M_j permettent de déterminer les densités de probabilité conditionnelles $p(M_j/\omega_1, z_i)$. Ces dernières sont ensuite approximées par des fonctions gaussiennes $N_{ij}(\mu_{ij}, \sigma_{ij})$ (Fig. 3).
μ_{ij} et σ_{ij} sont respectivement la moyenne et l'écart type de la mesure M_j sur la zone z_i.
En plus des informations statistiques sur la zone, la répartition spatiale des données (M_1, M_2, M_3) dans l'espace couleur (U, X, W), c'est à dire le domaine couleur noté D_i, contribue de manière implicite à l'élaboration du modèle.
Pour chaque contexte z_i, 3 couples de coefficients α_{ij} et β_{ij} génèrent 6 segments de droite $(\mu_{ij} + \beta_{ij} - \alpha_{ij}\sigma_{ij}; \mu_{ij} + \beta_{ij} + \alpha_{ij}\sigma_{ij})$ qui permettent de synthétiser un parallélépipède.
Ces coefficients sont calibrés tels que l'enveloppe du parallélépipède englobe au mieux le domaine couleur D_i (Fig. 4).
Dès lors le modèle a priori est défini par une fonction notée $skin_{ij}$ telle que :

$$skin_{ij} = \begin{cases} \frac{N_{ij}(\mu_{ij},\sigma_{ij})}{\max(N_{ij}(\mu_{ij},\sigma_{ij}))} \\ \text{si } \mu_{ij} + \beta_{ij} - \alpha_{ij}\sigma_{ij} \leq M_j \leq \mu_{ij} + \beta_{ij} + \alpha_{ij}\sigma_{ij} \\ 0 \text{ sinon.} \end{cases} \quad (4)$$

3.3 Fonctions de masse

Appriou [13] a suggéré une approche qui consiste à introduire chaque densité de probabilité a priori $p(M_j/\omega_1, z_i)$ et son degré de confiance d_{ij} correspondant dans une fonction de masse $m_{ij}(.)$. Cet ensemble est défini par une approche axiomatique dans l'ensemble de discernement Ω.
Dans la démarche développée ici les densités de probabilités sont remplacées par les fonctions $skin_{ij}$ qui prennent en considération implicitement la répartition volumique des données dans l'espace couleur (U, X, W).
Les éléments focaux associés à $m_{ij}(.)$ sont ω_1, ω_2 et Ω.
Les fonctions de masses sont définies par :

$$\begin{aligned} m_{ij}(\omega_1) &= \frac{d_{ij} R_i skin_{ij}}{1 + R_i skin_{ij}} \\ m_{ij}(\omega_2) &= \frac{d_{ij}}{1 + R_i skin_{ij}} \\ m_{ij}(\Omega) &= 1 - d_{ij} \end{aligned} \quad (5)$$

Les coefficients de pondération R_i traduisent la prise en compte des données $skin_{ij}$ pour chaque zone i.

3.4 Degrés de confiance

Pour les sources $S_1 = U$ et $S_2 = X$ l'ambiguïté entre classes est faible. Dès lors une modélisation probabiliste est utilisée d'où $d_{i1} = d_{i2} \approx 1$.

Par contre la source S_3 dépend de la luminance. La fiabilité de cette source pour la classe visage s'avère maximale ($d_{i3} = 1$) uniquement pour le niveau de gris moyen (μ_{i3}) de la zone modélisée.
L'ambiguïté entre classes croît ($m(\Omega) > 0$) lorsque M_3 s'écarte de μ_{i3}.
C'est pourquoi une fonction d'appartenance floue (Eq. 6) est suffisante pour caractériser la fiabilité de la source S_3 pour le contexte z_i (Fig. 1).
Sachant que $\mu_{13} \leq \mu_{23} \leq \mu_{33}$ on obtient :

$$\text{pour } i = 1 : \quad d_{13} = \begin{cases} \frac{M_3 + 2\mu_{23} - 3\mu_{13}}{2\mu_{23} - 2\mu_{13}} \text{ si } 3\mu_{13} - 2\mu_{23} \leq M_3 \leq \mu_{13} \\ \frac{-(M_3 - \mu_{23})}{\mu_{23} - \mu_{13}} \text{ si } \mu_{13} \leq M_3 \leq \mu_{23} \\ 0 \text{ sinon} \end{cases}$$

$$\text{pour } i = 2 : \quad d_{23} = \begin{cases} \frac{M_3 - \mu_{13}}{\mu_{23} - \mu_{13}} \text{ si } \mu_{13} \leq M_3 \leq \mu_{23} \\ \frac{-(M_3 - \mu_{33})}{\mu_{33} - \mu_{23}} \text{ si } \mu_{23} \leq M_3 \leq \mu_{33} \\ 0 \text{ sinon} \end{cases}$$

$$\text{pour } i = 3 : \quad d_{33} = \begin{cases} \frac{M_3 - \mu_{23}}{\mu_{33} - \mu_{23}} \text{ si } \mu_{23} \leq M_3 \leq \mu_{33} \\ \frac{-(M_3 + 2\mu_{23} - 3\mu_{33})}{2\mu_{33} - 2\mu_{23}} \text{ si } \mu_{33} \leq M_3 \leq 3\mu_{33} - 2\mu_{23} \\ 0 \text{ sinon} \end{cases} \quad (6)$$

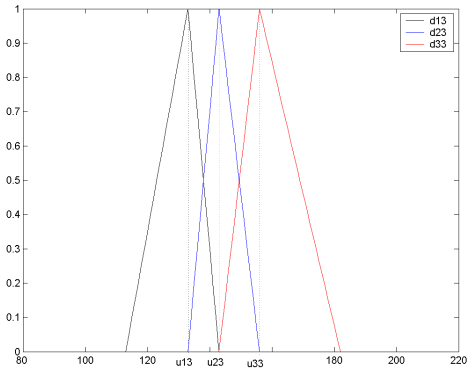


FIG. 1 – Degrés de confiance d_{i3} en fonction de M_3.

3.5 Décision

A chaque pixel en le site $s(x, y)$ de composantes couleur (M_1 ; M_2 ; M_3) (Eq. 3) sont associées trois masses contextuelles $m_i(s)$ par la règle de combinaison orthogonale normalisée de Dempster-Shafer (Eq. 1) :

$$m_i(s) = \oplus m_{ij}(s) \quad (7)$$

La règle de fusion disjonctive (Eq. 2) combine les masses contextuelles $m_i(s)$ (Eq. 5) afin d'associer une masse unique $m(s)$ à chaque pixel.
Cependant, les résultats expérimentaux montrent qu'une pondération des masses $m_i(s)$ par les degrés de confiance $d_{i3}(M_3(s))$ améliore la qualité de la segmentation. On obtient dès lors :

$$\begin{aligned} m(s) &= d_{13}(M_3(s)).m_1(s) \oplus_\cup d_{23}(M_3(s)).m_2(s) \\ &\quad \oplus_\cup d_{33}(M_3(s)).m_3(s) \end{aligned} \quad (8)$$

4 Résultats de la modélisation

4.1 Modèle a priori zone ombrée

La phase d'initialisation étant identique pour chaque zone, seule la modélisation de la zone ombrée z_1 est présentée en détail ci-après.
Dans un premier temps, l'utilisateur enregistre son visage dans une image puis sélectionne manuellement sur cette dernière une zone ombrée du visage (Fig. 2).

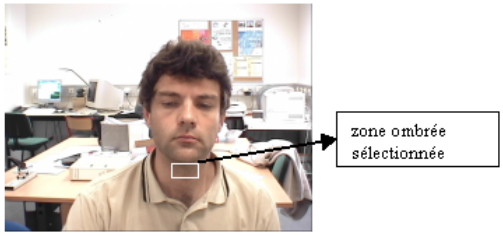
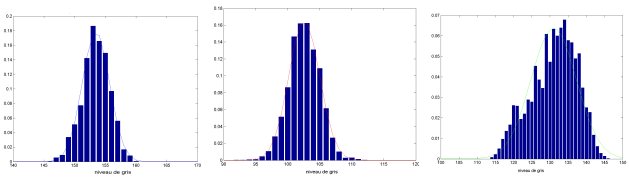


FIG. 2 – Phase d'apprentissage zone visage ombrée

Les densités de probabilité sont approximées pour chaque mesure M_j par une fonction gaussienne $N_{1j}(\mu_{1j}, \sigma_{1j})$ (Fig. 3).



a) b) c)

FIG. 3 – Densités de probabilité pour zone ombrée z_1 : a) $p(M_1/\omega_1, z_1)$; b) $p(M_2/\omega_1, z_1)$; c) $p(M_3/\omega_1, z_1)$.

Les segments de droite ($\mu_{1j} + \beta_{1j} - \alpha_{1j}\sigma_{1j}$; $\mu_{1j} + \beta_{1j} + \alpha_{1j}\sigma_{1ij}$) déterminés en faisant varier α_{1j} et β_{1j} sur chaque voie, génèrent un parallélépipède qui englobe le domaine couleur D_1.
Ici $\alpha_{11} = \alpha_{12} = 2$; $\alpha_{13} = 1.7$; $\beta_{11} = \beta_{12} = \beta_{13} = 0$; et $\sigma_{11} = 2.3$; $\sigma_{12} = 2.55$; $\sigma_{13} = 6.5$.
Ces coefficients α_{1j} et β_{1j} bornent dès lors les densités de probabilités et déterminent les fonctions $skin_{1j}$ (cf. Eq. 4 et Fig. 5).

4.2 Détection de la zone ombrée

Les fonctions de masses $m_{1j}(\omega_1) = \frac{d_{1j} R_1 skin_{1j}}{1 + R_1 skin_{1j}}$ (Eq. 5) dépendent du paramètre R_1 qui pondère l'importance des données $skin_{1j}$ caractérisant la classe visage. La fusion pixel conjonctive normalisée $m_1(s) = \oplus m_{1j}(s)$ (Eq. 7)

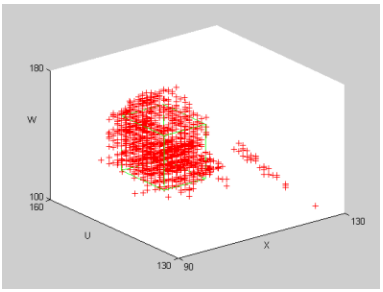
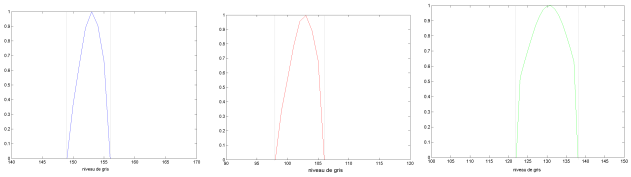


FIG. 4 – Représentation dans l'espace couleur (U, X, W), des composantes couleur des pixels (domaine D_1) et enveloppe parallélépipédique associée.



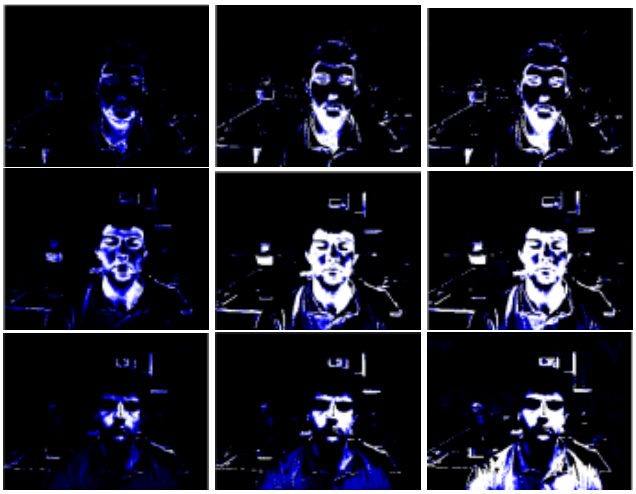
a) b) c)

FIG. 5 – Fonctions $skin_{1j}$: a) $skin_{11}$; b) $skin_{12}$; c) $skin_{13}$

affecte par indexation une masse $m_1(s)$ à chaque pixel en le site $s(x, y)$ de composantes $(M_1 ; M_2 ; M_3)$ (Eq. 3).
Les 3 images du haut de la Fig. 6 présentent les résultats expérimentaux de modélisation pour différentes valeurs de R_1. Pour $R_1 = 1$ la modélisation est satisfaisante car seules les parties teinte chair ombrées (en blanc) présentent sous le menton, le contour des yeux et sous les cheveux sont détectées. Pour $R_1 = 5$ ou $R_1 = 10$ la zone de modélisation s'élargit et des fausses détections apparaissent.

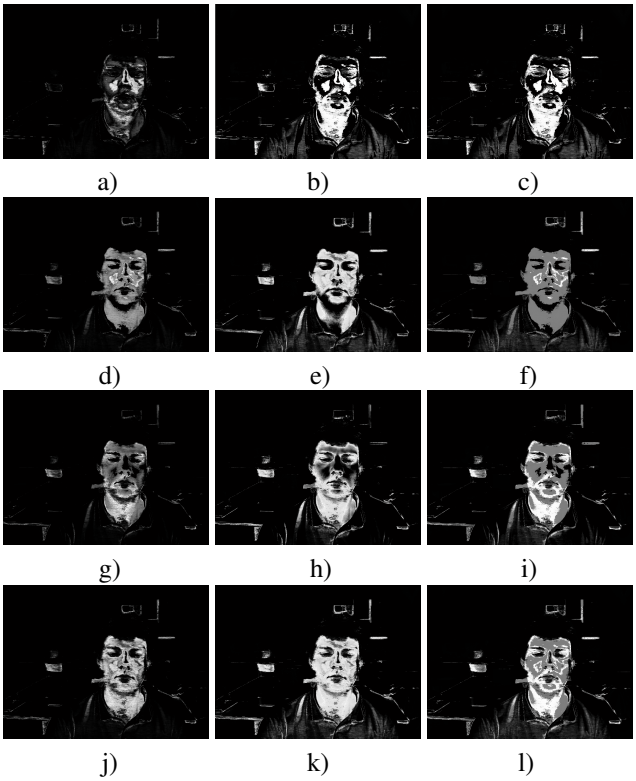
4.3 Synthèse de la modélisation

La Fig. 6 (milieu et bas) présente également les résultats de l'étape de la fusion conjonctive pour les zones moyenne et saturée, dont la démarche est similaire à celle présentée aux paragraphes 4.1 et 4.2.
La modélisation détecte assez fidèlement les parties ombrées et surexposées du visage (Fig. 6 haut et bas). Le modèle moyen (Fig. 6 milieu), bien que plus représentatif du visage, présente des défauts dans les zones ombrées ou surexposées. C'est pourquoi la fusion disjonctive (Eq. 8) est mise en œuvre pour mettre en concordance les différents modèles et optimiser le résultat de la détection (Fig. 7 gauche et milieu).
Lorsqu'uniquement deux masses sont combinées, la modélisation est insuffisante et présente des défauts soit :
– sur les parties du visage d'intensité moyenne (Fig. 7 a, b).
– sur les parties du visage surexposées (Fig. 7 d, e).
– sur les parties ombrées (Fig. 7 g, h).
La modélisation est optimale lorsque les 3 masses sont combinées ($R_1 = 1$, $R_2 = 1$, $R_3 = 1$) (Fig. 7 j). Cependant une pondération trop importante du paramètre R_i dé-



a) b) c)

FIG. 6 – Détection réalisée à partir des fonctions de masse m_i, $(i = 1, 2, 3)$ pour différentes valeurs de R_i : a) $R_i = 1$; b) $R_i = 5$; c) $R_i = 10$. En haut : zone ombrée $(i = 1)$; Au milieu : zone moyenne $(i = 2)$; En bas : zone surexposée $(i = 3)$.



a) b) c)
d) e) f)
g) h) i)
j) k) l)

FIG. 7 – Image issue de fusion pour différentes combinaison de masses dans le cadre de la théorie de DS (gauche et milieu) et probabiliste (droite) : a) $R_1 = 1$, $R_2 = 0$, $R_3 = 1$; b)c) $R_1 = 5$, $R_2 = 0$, $R_3 = 5$; d) $R_1 = 1$, $R_2 = 1$, $R_3 = 0$; e)f) $R_1 = 5$, $R_2 = 5$, $R_3 = 0$; g) $R_1 = 0$, $R_2 = 1$, $R_3 = 1$; h)i) $R_1 = 0$, $R_2 = 5$, $R_3 = 5$; j) $R_1 = 1$, $R_2 = 1$, $R_3 = 1$; k)l) $R_1 = 5$, $R_2 = 5$, $R_3 = 5$;

tériore légèrement la qualité de la modélisation (Fig. 7 k). Dans le cadre de la théorie de Dempster-Shafer, malgré un fond complexe composé d'éléments de couleur proche de celle de la chair (table, pull, affiche), le processus de fusion assure une bonne détection et s'avère aussi robuste aux occultations et aux variations de pose (Fig. 8 bas). Par contre, les Fig. 7 droite et Fig. 8 milieu représentent le résultat de la modélisation lorsque ($d_{ij} = 1$) c'est à dire lorsque la fiabilité des capteurs n'est plus prise en compte. Dès lors, $m_{ij}(\Omega) = 0$ (cf. Eq. 5) et seules les classes visage (ω_1) et fond (ω_2) sont considérées. Dans ce cadre probabiliste on constate que le résultat de la détection est de moins bonne qualité surtout dans les zones où l'imprécision des données n'est plus négligeable en raison de l'ambiguïté entre classes.

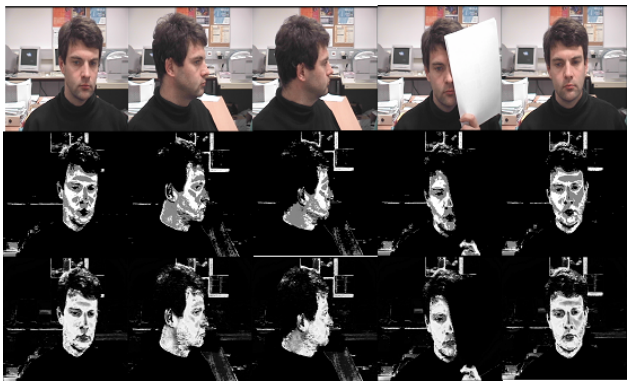


FIG. 8 – a) image 10 ; b) image 14 ; c) image 18 ; d) image 24 ; e) image 30

5 Conclusion et perspectives

L'originalité de la méthode réside en la modélisation de la teinte chair par la fusion de trois sources indépendantes du point de vue cognitif. De plus, les fonctions de masse (théorie de Dempster-Shafer) sont déterminées à partir de modèles a priori intégrant des données contextuelles spécifiques au visage et la fiabilité des capteurs. Ainsi, la particularité du visage de présenter des zones ombrées (cou) ou surexposées (nez, front) est prise en compte et la sensibilité face aux conditions d'éclairement est diminuée. Cependant la fusion d'informations couleur uniquement est parfois insuffisante et l'utilisation d'autres données indépendantes telles que le mouvement, la texture ou les contours doit améliorer la détection. La modélisation présentée est la première brique d'un système de vision active qui a pour finalité de détecter et de suivre un visage dans une séquence vidéo. Le modèle de teinte présenté trouvera aussi sa justification dans son intégration dans une boucle de rétroaction adaptative qui a pour finalité de gérer l'incomplétude mais aussi l'imprécision et le conflit dans le but d'optimiser les performances de la segmentation. Ceci fait l'objet de nos recherche en cours.

Références

[1] E. Hjelmås and B.K. Low, Face detection : A survey, *Computer Vision and Image Understanding*, Vol. 83, No. 3, pp. 236-274, Sept. 2001.

[2] Ming-Hsuan Yang, David Kriegman and Narendra Ahuja, Detecting faces in images : A survey, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 1, pp. 34-58, 2002.

[3] Prem Kuchi, Prasad Gabbur, P. Subbanna Bhat and Sumam David S., Human Face Detection and Tracking using Skin Color Modeling and Connected Component Operators, *IETE Journal of Research*, Vol. 38, No. 3&4, pp. 289-293, May-Aug 2002.

[4] S. Lu, G. Tsechpenakis, D. N. Metaxas, M. L. Jensen, and J. Kruse, Blob Analysis of the Head and Hands : A Method for Deception Detection, *International Conference on System Science* (HICSS'05), Hawaii, 2005.

[5] C.Garcia and M.Delakis. A neural architecture for fast and robust face detection, *Proc. of the IEEE-IAPR International Conference on Pattern Recognition* (ICPR'02), 2002.

[6] P. Viola and M. Jones, Fast and Robust Classification Asymmetric AdaBoost and a Detector Cascade, *Advances in Neural Information Processing System* 14, MIT Press, Cambridge, MA, 2002.

[7] M. Liévin, F. Luthon, Nonlinear color space and spatiotemporal MRF for hierarchical segmentation of faces in video, *IEEE Trans. on Image Processing*, Vol. 13, No. 1, Jan. 2004, pp.

[8] Vezhnevets V., Sazonov V., Andreeva A., A Survey on Pixel-Based Skin Color Dection Techniques, *Proc. Graphicon-2003* pp. 85-92, Moscow, Russia, September 2003.

[9] Dempster, A. P., A generalisation of Bayesian inference, *Journal of the Royal Statistical Society*, pp. 205-247, 1968.

[10] Shafer, G. A Mathematical Theory of Evidence, *Princeton University Press*,1976.

[11] I. Bloch, H. Maître, Data fusion in 2D and 3D image processing : An overview, http ://mirror.impa.br/sibgrapi97/anais/pdf/bloch.pdf.

[12] Ph. Smets. Constructing the Pignistic Probability Function in a Context of Uncertainty, *Uncertainty in Artificial Intelligence*, 5 :29-39, 1990.

[13] A. Appriou, DTIM Multisensor signal processing in the framework of the theory of evidence, *NATO/RTO Lecture Series 216 on Application of Mathematical Signal Processing Techniques to Mission Systems,* Nov 1999.