Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Sets and Systems 158 (2007) 325–336
FUZZY sets and systems
www.elsevier.com/locate/fss

Fuzzy fusion for skin detection

Aureli Soria-Frisch[a,*], Rodrigo Verschae[b], Aitor Olano[c]

[a] Technology Department, Pompeu Fabra University, Pg. Circumval.lació 8, 08003 Barcelona, Spain
[b] Department of Electrical Engineering, Universidad de Chile, Av. Tupper 2007, 837-0451 Santiago, Chile
[c] Innovae Vision, Mikeletegi 54, 20009 Donostia

Available online 10 November 2006

Abstract

Complex image processing tasks rarely succeed through the application of just one methodology. The implementation of different methodologies, whose treatment of the input images is complementary, can help in the successful attainment of the system goal. The result of the complementary procedures has to be eventually fused in order for the system to improve the result of each methodology taken on its own. Computer vision systems mostly employ simple fusion strategies for this aim. This simplicity downplays the relevance of the fusion stage.

The paper presents a framework for skin detection, a pre-processing task very useful in application fields like video surveillance, human–machine interface, and cyber-crime prosecution. The framework is based on the employment of the fuzzy integral, which subsumes the performance of more simple fusion operators. As it is shown herein the framework manages therefore to cope with the complexity of skin detection under changing illumination conditions. The performance evaluation of the framework is undertaken on hand of a benchmark database in the paper.

Keywords: Fuzzy measure theory; Fuzzy integral; Multi-sensory fusion; Data fusion; Computer vision; Color image processing; Skin detection

1. Introduction

Skin detection attains the isolation of the pixels of an image (or video) that correspond to human skin. It is very useful for detecting human parts (faces, hands, etc.), what can be applied in numerous application fields, e.g. video surveillance, human–machine interface, cyber-crime prosecution. The main difficulties for detecting skin areas are the existence of pixels with skin-like colors in the background that do not correspond to human skin, and the fact that the skin distribution becomes multi-modal under uncontrolled illumination conditions.

One of the possibilities to cope with the problems described in the former paragraph is the combination of different detection methodologies. In this context, each procedure presents its particular properties and tackles the problem from a different perspective. Therefore, the results of these procedures operate as different pieces of evidence, whose fusion can deliver a successful solution [14]. The fulfillment of this goal depends both on the performance of the detection procedures and the features of the applied fusion operator.

The fuzzy integral presents very good features for the realization of the fusion operation in computer vision because of its mathematical properties [14]. Among these properties it is worth mentioning herein the consideration of the

* Corresponding author.
E-mail addresses: aureli.soria-frisch@ieee.org (A. Soria-Frisch), rodrigo@verschae.org (R. Verschae).

0165-0114/$ - see front matter
doi:10.1016/j.fss.2006.10.018

ranking among the input channels in the weighting scheme. Hence a different set of weights is defined for each canonical region of the input channels' space. This property results in the robustness of the operator w.r.t. a change in the luminance/saturation of color hues, which has been exploited in [15] for the detection of ink seals on document images. Taking all these facts into account a framework for the robust skin detection that is based on the application of the fuzzy integral is presented herein.

The fuzzy integral was introduced by Sugeno [16]. In computer vision the fuzzy integral outperforms other fusion operators, e.g. maximum, weighted sum, order weighted averaging [14]. Since the fuzzy integral mathematically generalizes these operators, its outperformance has to be understood as the possibility of the fuzzy integral of computing a larger diversity of results than these operators. For instance, if the optimal result is achieved by applying an average operator, you can attain this optimal result by applying a fuzzy integral. This outperformace is not for free, but results from the increment in the number of parameters that have to be tuned if the fuzzy integral is employed. As a consequence the relevance of methodologies for parameterizing the operator in an automated manner increases.

The fuzzy integral is used two-fold in the framework presented herein. First the methodology for the detection of particular color clusters presented in [15] is applied. This methodology, which was denoted as *Difference of Fuzzy Integrals* (*DoFI*), is applied in order to estimate the skin distribution in the input images. Second the fuzzy integral is used in the fusion of the *DoFI*'s estimate with these of two other well-known skin detection methodologies, namely a *mixture of Gaussians* (*MoG*) [11], and a fuzzified version of the Hsu-methodology [10]. To the best of our knowledge no fusion approach has been presented hitherto for skin detection.

The paper is organized as follows. Section 2 gives an overview on the state of the art for skin detection and particularly describes the two methods used in the fusion procedure. The fuzzy integral and its application in the framework presented herein are described in Section 3. It is worth pointing out that the elucidation of the relationship between the ranking property of this operator formerly mentioned and its robustness for computer vision can be found in the introduction of this section. The procedure for the performance evaluation is described in Section 4. The reader can find the analysis of the framework results in Section 5. Finally, some conclusions and an overview of the projective work is given (Section 6).

2. Background on skin detection

Skin detection is still an open research question. The aim of skin detection is to find out all skin pixels of an image or video. Skin detection is one of the first preprocessing steps of a larger system for e.g. video surveillance. Since its performance affects all following modules of the system, it is important for it to perform both with high accuracy and low processing time.

To face this problem many different color spaces have been used (*RGB*, *normalized-RGB*, *YCbCr*, *Lab*, *HSV*, etc.) [19]. [1] Nevertheless, [1] shows that the optimality of a skin detection algorithm does not depend on the employed color space, but on the design of the detection algorithm. Skin detection algorithms can be classified into two groups: pixel-based and context-based. Since context-based methods are built on top of pixel-based ones, an improvement on a pixel-based methodology supposes a general advancement in the resolution of skin detection.

Pixel-based algorithms classify each pixel individually without taking the other pixels of the image into consideration. These methodologies realize the skin detection either by bounding the skin distribution or by using statistical models on a given color space. Among pixel-based methodologies those using statistical models estimate the skin distribution either by using parametric (e.g. *MoG*) or non-parametric functions (e.g. histograms) [11]. A parametric model is presented in [10] and described in Section 2.2. A more complete review of existing pixel-based methodologies can be found in [19].

Context-based algorithms have been recently presented [12,3]. They classify each pixel by taking contextual information into consideration. In [12] a diffusion (region growing) algorithm, called *SkinDiff*, is developed to perform the classification. The seeds of the diffusion algorithm are pixels with large probability of belonging to the skin class. On the other hand, the mean-shift algorithm have been used in order to first segment the input image, whose pixels are eventually classified through a *MoG* [3].

Two pixel-based methodologies, namely *MoG* and the Hsu-methodology, are employed in this paper in order to find two different estimates of the skin distribution in the input images. The resulting distributions are eventually fused

[1] The interested reader is referred to [19,5,6] for a summary on the mathematical expressions of color space transformations.

together with the *DoFI* (see Section 3.1) estimate through the fuzzy integral. The following subsections describe the two pixel-based methods that are used later on this work.

2.1. MoG

The performance of statistical models in skin detection does not depend much on the color space they are trained on. They have shown good performance for pixel-based classification [11]. We have used herein the *MoG* provided by [11], which consist of 16 Gaussians in the *RGB* color space and whose parameters were obtained using the Expectation–Maximization algorithm. Hence, the parameters of the *MoG* are estimated from a training set, which is described in Section 4. The *MoG* presents the following probability expression:

$$\mathcal{P}(x) = \sum_{i=1}^{N} \frac{w_i}{(2\pi)^{3/2}|\sum_i|^{1/2}} \mathrm{e}^{-(1/2)(x-\mu_i)^{\mathrm{T}} \sum_i^{-1}(x-\mu_i)}. \tag{1}$$

Therefore, *MoG* delivers the probability of each pixel of belonging to the skin class (in contrast to the non-skin class), i.e. a real value. Therefore, it can be fused with the result of other methodologies without further transformation. Usually, this methodology requires more processing time than a bounding model, like the one described in the following section, but it is more accurate.

2.2. Hsu-methodology

The Hsu-methodology [10] defines an ellipse in order to bound the skin distribution on a two-dimensional feature space. The feature space taken into account results from the rotation of the chromaticity components of the color space *YCrCb*, where the skin color distribution is more compact that in other color spaces, e.g. *RGB*, *HSV*. This fact allows successfully characterizing the skin distribution with a simple model. In this paper the same transformation employed in [10] is applied. This includes the parameters needed to transform the chroma and the parameters of the ellipse in the transformed chroma space. For the sake of compactness the reader is referred to [10] for the exact formulas and values of these parameters.

The Hsu-methodology delivers a binary output that represents if the pixels falls inside or outside of the ellipse. Since the fuzzy integral works in the real domain, the Hsu-methodology is modified herein by fuzzifying it. The fuzzification procedure is described in Section 3.2.

3. Two novel approaches for skin detection based on the fuzzy integral

The theoretical background on the fuzzy integral is described in the following section. This serves as introductory material to the description of the two pixel-based approaches for skin detection, which are newly presented in this paper (see Sections 3.1 and 3.2). Both approaches employ the fuzzy integral as fusion operator.

The fuzzy integral [16] constitutes a generalization of different fusion operators [14]. A fuzzy integral presents three elements in its mathematical expression. The values to be integrated are denoted by $h_i(x_i)$, where $i = 1, \ldots, n$ and n is the number of information sources. The coefficients of the fuzzy measures, which are the weighting functions used in the operator, are denoted by $\mu(A_i)$. The third element are the fuzzy connectives used in the operation of the two elements previously mentioned.

The used type of fuzzy connectives defines the type of fuzzy integral. Although there are several types of fuzzy integral [14,9], mainly two of them have been used in real applications. The first one uses a maximum ($\vee$) and a minimum ($\wedge$) operators as the fuzzy connectives. This integral is known as the Sugeno Fuzzy Integral and presents the following expression:

$$\mathcal{S}_\mu[x_1, \ldots, x_n] = \bigvee_{i=1}^{n} [h_{(i)}(x_i) \wedge \mu(A_{(i)})]. \tag{2}$$

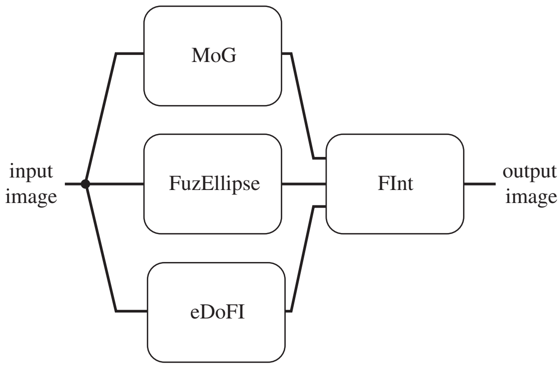

Fig. 1. Block diagram of the framework for the detection of skin based on the fusion paradigm. *MoG*, mixture of Gaussians; *FuzEllipse*, fuzzified Hsu-methodology; *eDoFI*, extended difference of fuzzy integrals (see explanations in text).

On the other hand, the Choquet Fuzzy Integral makes use of the sum ($\sum$) and the product ($\cdot$), as stated by

$$\mathcal{C}_\mu[x_1, \ldots, x_n] = \sum_{i=1}^{n} h_{(i)}(x_i) \cdot [\mu(A_{(i)}) - \mu(A_{(i-1)})], \tag{3}$$

where $\mu(A_{(0)}) = 0$. In the skin detection procedure described herein $x_i \forall i = 1, 2, 3$ denote the image grayvalues of each color channel in a particular color representation, h_i denote the fuzzification functions applied on these data, and $\mu(A_i)$ weight the importance of the color channels in the detection of skin.

A fuzzy measure $\boldsymbol{\mu}$ presents 2^n coefficients $\mu(A_j)\ \forall j = 1, \ldots, 2^n$, so many as subsets A_j can be formed among the input information sources. From all these coefficients just n are taken into account in each fuzzy integration $\mu(A_{(j)}) = \mu(\{x_{(1)}, \ldots, x_{(j)}\}) \forall j = 1, \ldots, n$. These are selected upon the sorting operation denoted by the enclosed subindices in expressions (2) and (3). If, for e.g., the three input channels fulfill $x_3 > x_2 > x_1$, then $x_{(1)} = x_3, x_{(2)} = x_2$, and $x_{(3)} = x_1$. This operation involves taking the coefficients $\mu(\{x_3\})$, $\mu(\{x_2, x_3\})$ and $\mu(\{x_1, x_2, x_3\})$ into account.

For the sake of comprehension a numerical example is given, where two input channels are fused through the Choquet fuzzy integral, i.e. $n = 2$ and $i = 1, 2$. Let us define the fuzzy measure $\boldsymbol{\mu}$ by its coefficients: $\mu(A_1) = \mu(\{x_1\}) = 0.5$, $\mu(A_2) = \mu(\{x_2\}) = 0.2$, and $\mu(A_3) = \mu(\{x_1, x_2\}) = 1$. If $x_1 = 4$ and $x_2 = 6$, i.e. $x_2 > x_1$, then $x_{(1)} = x_2$ and $x_{(2)} = x_1$. In this case the selected coefficients are $\mu(A_{(1)}) = 0.2$ and $\mu(A_{(2)}) = 1$. The Choquet fuzzy integral results therefore in

$$\mathcal{C}_\mu[x_1, x_2] = 6 \cdot (0.2 - 0) + 4 \cdot (1 - 0.2) = 4.4,$$

while for $x_1 = 6$ and $x_2 = 4$ would have resulted in 5.

If the fuzzy integral is used in the fusion of image channels, the expressions (2) and (3) are applied for each pixel. For deeper information on the application of the fuzzy integral in computer vision the reader is referred to [14,17]. However, it is worth commenting one of the more remarkable features of the fuzzy integrals, namely the definition of a set of weighting coefficients for each canonical region of the feature space [14,9]. A canonical region is an area of the feature space, where a particular ordinal relationship among the feature components is fulfilled. This kind of definition improves the robustness of the operator w.r.t. a change in the acquisition or the illumination conditions as it has been shown in [14]. The robustness of the fuzzy integral lays on the fact that a change in the illumination conditions modifies the grayvalue of the pixels in the color channels but not the ordinal relationship among these grayvalues.

As it can be observed in Fig. 1 the fuzzy integral is used for skin detection both within the *DoFI* methodology [15], and for the fusion of three estimates of the skin distribution on the input images. These two approaches based on the employment of the fuzzy integral are described in the following two sections.

3.1. Difference of fuzzy integrals (DoFI) and its extension for skin detection

As formerly mentioned a methodology based on the fuzzy integral was used for the segmentation of ink seals in [15] and denoted as *DoFI*. The main feature of this methodology is its ability to isolate a cluster characterized by a

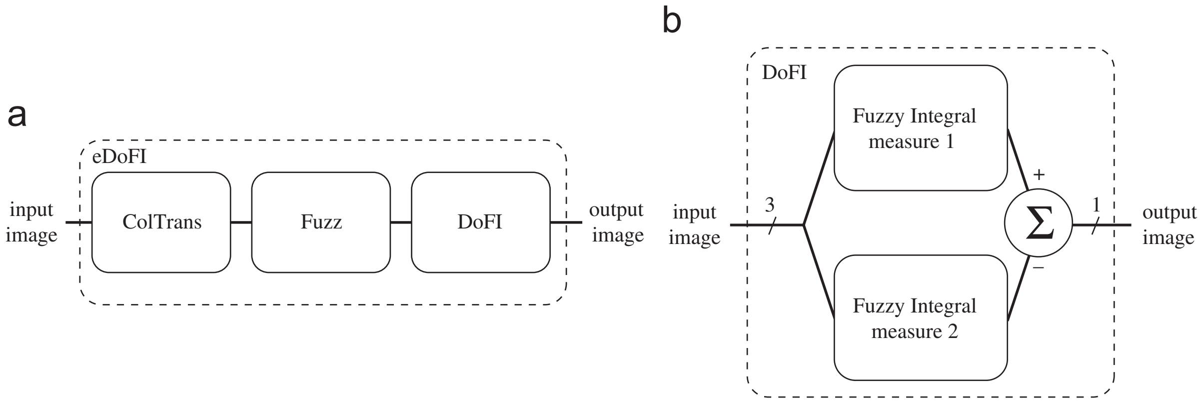

Fig. 2. Submodule for skin detection based on an extension of the *Difference of Fuzzy Integrals* [15]. (a) Block diagram of the extended *DoFI*. The basic methodology is extended by adding a color space transformation (*colTrans*) and a fuzzification (*Fuzz*) stages. (b) Block diagram of the basic *DoFI* [15]. A fuzzy integral is firstly computed with respect to two different fuzzy measures. The change of fuzzy measure mainly affects the skin color cluster, leaving the other components of the image unmodified.

particular color hue from an input image. The detection of these areas is attained by computing the difference of two fuzzy integrals. The two fuzzy integrals differ on the fuzzy measure used on them, $\boldsymbol{\mu}^1$ and $\boldsymbol{\mu}^2$. The two fuzzy measures change the value of these coefficients that control the canonical region occupied by the color cluster of interest, whereas the rest of them are left untouched. Therefore, just the pixels within this cluster are modified, or at least this is the purpose of this change. In contrast to the application of the *DoFI* for seal detection [15], where bluish ink seals were detected, the element to be detected herein is characterized by the hues that correspond to different skin types. The block diagram of the procedure is depicted in Fig. 2.

Taking the existing frameworks for skin detection based on the utilization of different color models (see Section 2) as a reference, a module (*ColTrans*) for the transformation among different color spaces has been added to the basic methodology (see Fig. 2). Thus, this module takes the input color image $I(x, y)$ in the *RGB* color space, and realizes the mapping

$$I(x, y) = \{I_R(x, y), I_G(x, y), I_B(x, y)\} \rightarrow \{I_1(x, y), I_2(x, y), I_3(x, y)\}, \tag{4}$$

where I_i are the color channels of the new color representation, e.g. *rgb*, *YCbCr*. The application of another module (*Fuzz*) undertakes the fuzzification of this color representation delivering F_i (see Fig. 2a), which can be formally described as the mapping

$$h(x) : \{I_1(x, y), I_2(x, y), I_3(x, y)\} \rightarrow \{F_1(x, y), F_2(x, y), F_3(x, y)\}. \tag{5}$$

This module has been added to the original *DoFI* methodology [14] as well. *Fuzz* can be easily implemented by applying Parzen windows [4] on a training image set. Parzen windows is a non-parametric method for estimating probability density functions $h(x)$ up to n sample data points. The estimation is undertaken by placing a generic window function $\varphi(u)$ of the form [4]:

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-u^2/2} \tag{6}$$

on each data point x_i. The resulting estimation can be expressed as

$$h(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{v_n} \varphi\left(\frac{x - x_i}{v_n}\right), \tag{7}$$

where $v_n = v_1/\sqrt{n}$ and v_1 is a parameter at user's disposal [4]. The fuzzification function succeeds therefore in the form $F_j = h(I_j)$.

Thence a difference image $D(x, y)$ is obtained in the *DoFI* methodology by applying

$$D(x, y) = \|\mathcal{F}_{\mu^1}(x, y) - \mathcal{F}_{\mu^2}(x, y)\|, \tag{8}$$

where $\mathcal{F}_{\mu^i}$ states for the images resulting from the computation of the fuzzy integral w.r.t. the fuzzy measure $\boldsymbol{\mu}^i$ on each color pixel, as expressed by (2) and (3) by taking $x_1 = F_1(x_i, y_i)$, $x_2 = F_2(x_i, y_i)$, and $x_3 = F_3(x_i, y_i)$.

The application of the procedure results in a normalized real value image $D(x, y)$, whose pixel values indicate the membership degree of the pixels to a particular color class. In the application described herein this class corresponds to the skin class.

The automated parameterization of the *DoFI*, which attains the determination of the coefficients of the fuzzy measures $\boldsymbol{\mu}^1$ and $\boldsymbol{\mu}^2$, can be realized by applying genetic algorithms [7]. [2] This methodology has been successfully used in the parameterization of the fuzzy integral [18,2]. The process is described in the following paragraph.

The coefficients of the fuzzy measures, which are going to be determined, are first codified as the genes of the individuals in the population. This codification succeeds in a binary domain. Therefore, the real values are quantified and expressed as binary strings. Once the initial codification has been realized, the standard search procedure of genetic algorithms is applied. For this purpose the standard mutation, crossover, and selection operators [7] can be applied. The selection of the fitness function that drives the search deserves further attention (see Section 4.2).

3.2. Fusion through the fuzzy integral

The methodology presented herein is completed by a fusion module (see Fig. 1). This module is implemented through the fuzzy integral. The *MoG* (see Section 2.1) and a fuzzified version of the Hsu-methodology (see Section 2.2) are combined with the DoFI (see Section 3.1) through the application of a fuzzy integral.

The fuzzified version of the Hsu-methodology described in Section 2.2 is obtained by giving a value of 1.0 to the center of the ellipse and then linearly decreasing the value until reaching the border of the ellipse, to which a value of 0.5 is assigned. The pixels in the outer part of the ellipse receive a value lower than 0.5, which is proportional to their distance to the border. This modification is applied in order for the methodology to deliver a result defined in the real domain. The fulfillment of this constraint, which characterizes the *MoG* and the *DoFI* methodologies, allows the application of the fuzzy integral for the fusion without further transformations.

The fusion stage adds a new parameter to the general methodology $\boldsymbol{\mu}^{\mathrm{f}}$, which denotes the fuzzy measure used in the fusion. The coefficients of this fuzzy measure are automately assessed by applying again genetic algorithms. It is worth mentioning that the *DoFI* and the fusion stage are parameterized apart. The fitness functions used in the construction of all parameters is given in Section 4.2.

4. Performance evaluation

Two data sets have been used in this work, [3] one for training/adapting the fuzzy measures $\boldsymbol{\mu}^1$, $\boldsymbol{\mu}^2$, $\boldsymbol{\mu}^{\mathrm{f}}$ of the fuzzy integrals involved, and a different one for the performance evaluation. For the training, 44 human skin photos were downloaded from internet and manually labeled, obtaining 350,000 different skin pixels. A similar amount of non-skin pixels was randomly selected from a database of more than 2,000,000 non-skin pixels obtained for images containing no skin.

The resulting adaptation of the methodologies presented herein have been evaluated on a different database as the training one. Hence, a subset of the data set used in [13] has been used in the evaluation of the procedures. This data set contains 427 labeled images taken from popular movie sequences. The video streams present imagery in the RGB color space of 8 bits/channel. The label streams are three-valued. From the three labels employed there, just the skin and the non-skin ones have been taken into consideration. It is worth mentioning that the results obtained with the two methodologies based on the fuzzy integral are compared with those obtained from the application of the *MoG* and the Hsu-methodology. These two methodologies are being extensively used in applications of skin detection [11,10]. Therefore, the Receiver Operation Curve analysis is not only presented as a performance measure in absolute terms, but for the sake of comparison among different skin detection methodologies as well.

[2] A very useful implementation of genetic algorithms can be found in `http://lancet.mit.edu/ga/`. The framework presented herein employs this software library.

[3] The well-known Compaq database [11] for training and testing skin detection algorithms is not available anymore.

4.1. Measure of performance through a ROC curve

The performance of a detection system can be evaluated on hand of a receiver operating characteristic (ROC) curve. [4] ROC curves summarize the performance of a system by presenting the detection rate as a function of the false positive rate (FPR) of the system. Hence, the two measures used to build a ROC curve in the detection of skin are defined as following:

- *FPR*: Relative number of non-skin pixels the system has wrongly classify as positive skin detection.
- *True positive rate* (*TPR*): Relative number of skin pixels the system did detect as skin.

The *FPR* and *TPR* are estimated by means of a labeled training set $S = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, with $x_i \in \Re^3$ being a color pixel and $y_i \in \{0, 1\}$ its label. When y_i value is equal 1, x_i is a skin pixel and when y_i is 0, x_i is an non-skin pixel. For a given classification function $f(x) : \Re^3 \to \{0, 1\}$, the FPR and TPR are given by

$$FPR = \frac{|\{(x, y) \in S : y = 0 \wedge f(x) = 1\}|}{|\{(x, y) \in S : y = 0\}|}, \tag{9}$$

$$TPR = \frac{|\{(x, y) \in S : y = 1 \wedge f(x) = 1\}|}{|\{(x, y) \in S : y = 1\}|}. \tag{10}$$

These two measures can be used in order to numerically assess the system performance. The *FPR* represent the failures of the system in the classification, whereas the *TPR* its success.

The numerical relationship between *TPR* and *FPR* represent the implicit contradiction between the ideal goals of a classification system. Therefore, the average number of succeeds cannot be increased, without having to pay the price of an associated increment in the average number of mistakes. The performance of the system is fixed up by establishing a system operation point, i.e. a pair (*TPR*, *FPR*), that corresponds to the best trade-off between these system two goals. The operation points of the ROC curves are obtained by using different thresholds in the computation of the algorithm output being evaluated. Herein the *TPR* and the *FPR* are used within one of the fitness functions that drives the parameterization of the methodologies (see Section 4.2).

4.2. Parameterization of the methodologies

The *DoFI* approach was evaluated on three color spaces: *RGB*, *normalized-RGB*, and *YCbCr*. Two types of fuzzy integral were evaluated: the Sugeno and Choquet fuzzy integrals.

A *steady-state* genetic algorithm have proven to give the best performance. The genome of the individuals codifies the $3 \cdot (2n - 1)$ coefficients of $\boldsymbol{\mu}^1$, $\boldsymbol{\mu}^2$, and $\boldsymbol{\mu}^{\mathrm{f}}$ (see Section 3).

The genetic algorithm is parameterized as following. The probability of replacement is set in the interval [0.8, 0.9]. The population presents a number of individuals in the interval [50, 500], where this parameter just influences the convergence time. Probabilities is set as 0.9 for crossover and 0.2 for mutation. The applied GA makes use of the well-known *roulette wheel* selection scheme [7] and a *2-point* crossover [7] until a maximum of 200 generations is achieved. Different fitness functions are tested. Only two of them show some significance. They are described in the following paragraphs.

In the first one the set of training pixels, i.e. the pixels of the images that are part of the training databases, is split in two subsets. The first one, which is denoted as S, includes the pixels that belong to skin areas. The second one, which is denoted as $\bar{S}$, includes the remaining pixels. This first fitness function attains the maximization of the result for those pixels included in S and the minimization of the result for those included in $\bar{S}$. This fact can be formulated

[4] This kind of performance analysis was initially undertaken in Signal Theory (see e.g. J.P. Egan, Signal Detection Theory and ROC Analysis, Academic Press, New York, 1975) and is nowadays used in several disciplines from medical diagnosis to pattern recognition.

as follows:

$$\phi_1 = \frac{\sum_{i \in S} R_i}{\sum_{i \in (S \cup \bar{S})} R_i} + \left(1 - \frac{\sum_{i \in \bar{S}} R_i}{\sum_{i \in (S \cup \bar{S})} R_i}\right), \tag{11}$$

where the result R_i states for $D(x_i, y_i)$ as expressed by (8) in the parameterization of the *eDoFI* and for $\mathcal{F}_{\mu^{\mathrm{f}}}$ in the parameterization of the fusion stage.

The second one presents the following expression:

$$\phi_2 = TPR + (1 - FPR). \tag{12}$$

This fitness function takes into consideration the *TPR* and the *FPR* as defined by Eqs. (9) and (10), which are explained in Section 4.1. It measures the overall rate of correctly classified pixels, giving the same importance to both classes. For this purpose the classification function $f(x)$ given in Eqs. (9) and (10) is defined as

$$f(x) = \begin{cases} 1, & R_i > \theta, \\ 0, & R_i \leqslant \theta, \end{cases} \tag{13}$$

where R_i denotes the real-valued result as explained in the former paragraph, and θ the threshold between the skin and non-skin crisp classes.

The main difference between the fitness functions ϕ_1 and ϕ_2 is that this last one makes use of the result after the binarization through θ. Therefore, the threshold θ is applied on the classifier's output R_i in order to evaluate ϕ_2. It is worth mentioning that this threshold θ is part of the genome used during training. Summarizing, the fitness function ϕ_2 employs the binary classifier's output given by θ, whereas ϕ_1 employs the real-valued classifier's output R_i.

5. Performance analysis

In this section different results obtained by applying the described methodologies are analyzed. The Choquet Fuzzy Integral has performed much better than the Sugeno Fuzzy Integral in all realized tests. Therefore, just the results for this type of fuzzy integral are shown. The outperformance of the Choquet fuzzy integral with respect to Sugeno's one in classification, which confirms once again what is stated in the literature [14,8], may be due to the larger smoothness of the operator's level curves [9].

On the one hand, the performance of different color transformations and the convenience of using a fuzzification stage in the extension of the *DoFI* is analyzed. On the other, the performance obtained by applying the formerly described fitness functions ϕ_1 and ϕ_2 in the parameterization of both the *eDoFI* and the fusion stages are thence analyzed. This analysis, which can be found in the following sections, is undertaken on hand of the evaluation database.

5.1. DoFI for skin detection

First a comparison among the results obtained with an *eDoFI* (see Section 3.1) on different color spaces is depicted in Fig. 3. The results obtained by employing both fitness functions are shown in the form of the binarized resulting image. The depicted results compare the performance in terms of employed color space, whereas the comparison between the two fitness functions' application is shown later in the text.

The best results for ϕ_1 are obtained by applying a fuzzification stage based on Parzen windows as given by Eq. (5), which are priorly trained. Among all the color spaces the *DoFI* methodology shows the best performance on the *YCrCb* color space for this fitness function (first row of Fig. 3d). On the other hand, the fuzzification stage was not very useful when training the methodology w.r.t. the fitness function ϕ_2. Therefore, the shown results (second row of Fig. 3) are computed without taking this stage into consideration. In this case the best result was obtained on the RGB color space.

The comparison of the best performing configurations of the extended *DoFI* methodology (see Fig. 2a), namely:

- RGB to YCrCb for *ColTrans* plus Parzen windows for *Fuzz* plus the basic *DoFI* parameterized with ϕ_1,
- no transformation for *ColTrans* plus no *Fuzz* stage plus the basic *DoFI* parameterized with ϕ_2,

is shown in Fig. 4. The reader can observe the result for one of the images both in the grayvalue and the binary domains together with the ROC curve computed all over the database.

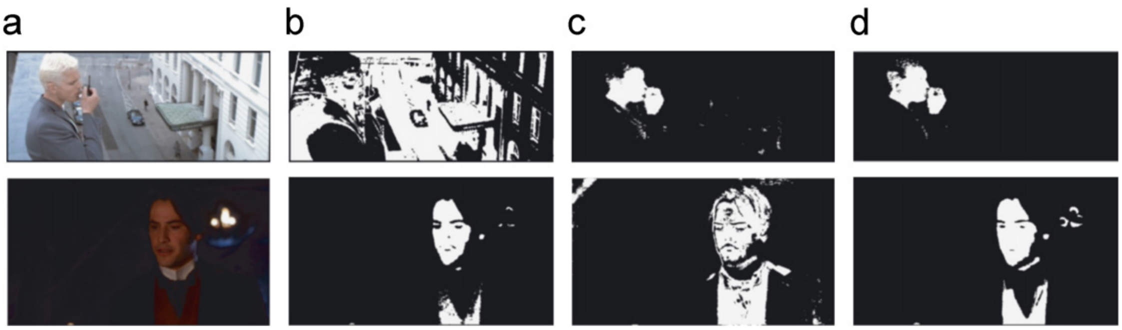

Fig. 3. Results of the difference of fuzzy integrals (*DoFI*) obtained on different color spaces for the fitness functions ϕ_1 (first row) and ϕ_2 (second row). (a) Input images [13]. (b) RGB-color space. (c) Normalized RGB-color space. (d) YCrCb-color space.

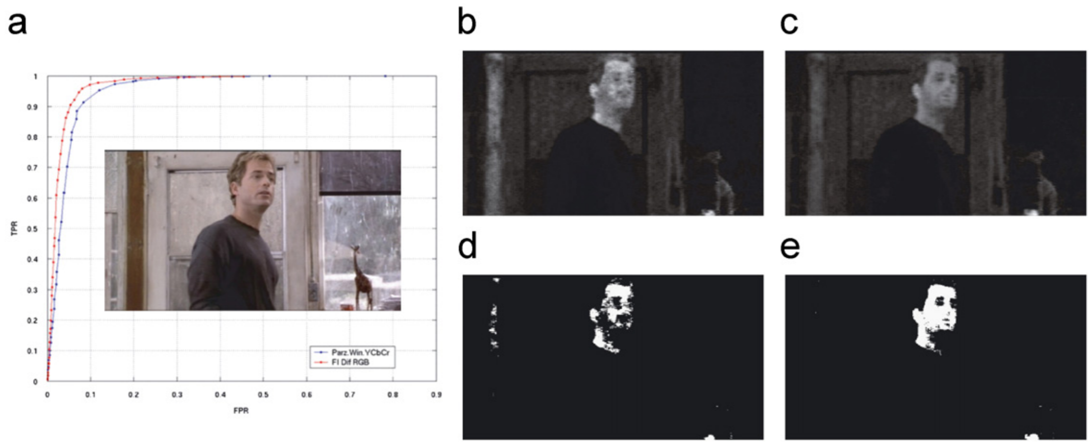

Fig. 4. Comparison of the best performing configurations of *eDoFI* for the fitness functions ϕ_1 and ϕ_2. (a) Input image together with ROC curve for the evaluation database [13]. (b) Fuzzy result for ϕ_1 on the YCrCb color space with Parzen windows. (c) Fuzzy result for ϕ_2 on the RGB color space without Parzen windows. (d) Automated binarization on (b). (e) Automated binarization on (c).

5.2. Fusion performance

Further results (see Figs. 5–7) compare the performance of the individual methodologies, i.e. *MoG*, *FuzEllipse*, and *DoFI*, with this of the fusion procedure. In this case the best configuration of the *eDoFI* is selected. Hence no color space transformation and no fuzzification stages were taken into consideration. Therefore, the comparison is based on the *DoFI* already presented in [15].

The fitness function ϕ_2 drove again the framework to perform at its best. This can be observed in Fig. 5, where the ROC curves of the framework all over the database are shown. For the sake of clarity just the most significant part of the curves is shown. The performance of the fitness functions can be observed in the image domain on hand of Fig. 6.

The compensatory effect of the fusion stage can be observed in Fig. 7, where the fitness function ϕ_2 was used in its parameterization. Hence, some of the background parts segmented as skin by the FuzEllipse (see Fig. 7c) are suppressed in the fusion result, shown in Fig. 7e, by taking the result of the DoFI methodology into account (see Fig. 7d), whereas the segmentation of the lamp highlight of this result is reduced by the former one. The same compensatory effect manage to complete the skin detection on the actor face in spite of the presence of a light reflection, whose detection is not achieved neither by the *MoG* nor by the FuzEllipse. The detection of this reflected skin part by the DoFI methodology is due to its robustness w.r.t. the luminance level mentioned in Section 3. The compensatory effect of the fusion methodology can be observed in Fig. 5b as well. This ROC statistically shows the convenience of applying the fusion methodology.

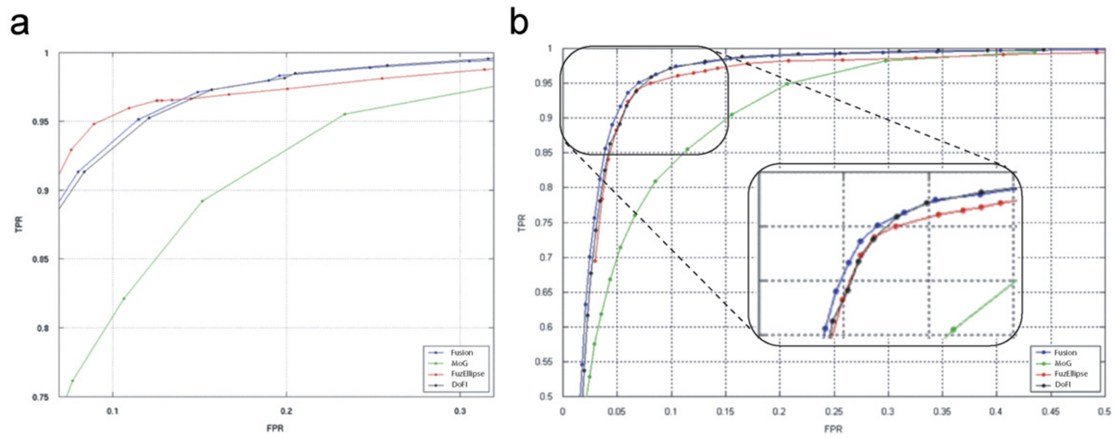

Fig. 5. ROC curves obtained on the evaluation databases [13] for all methodologies used herein (see legend). (a) ROC curve obtained by parameterizing the framework with ϕ_1. (b) ROC curve obtained by parameterizing the framework with ϕ_2. A detail of the most significant part of the curve is zoomed.

Fig. 6. Comparison of the fitness functions performance for the fusion stage in the binary image domain on hand of one of the images of the evaluation database. (a) Input image. (b) Binarized result of the fusion stage after parameterizing it through ϕ_1. (c) Binarized result of the fusion stage after parameterizing it through ϕ_2.

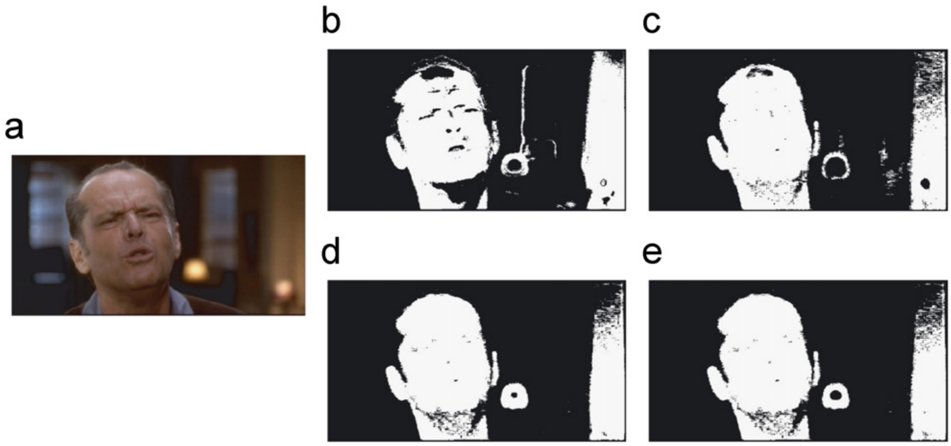

Fig. 7. Comparison of the results obtained for different methodologies in the binary domain by taking the fitness function ϕ_2 into account. (a) Input image [13]. (b) Mixture of Gaussians. (c) Fuzzy-ellipses. (d) Difference of Fuzzy Integrals. (e) Fusion of the former ones.

Finally, the results of the framework in the grayvalue domain are shown (see Fig. 8). Due to the usage of a fuzzy methodology for the fusion stage, the result can be expressed in this image domain. This adds some degree of flexibility to the stage for skin detection in case of its inclusion in a more general processing system.

Fig. 8. Some of the frame images of the evaluation database [13] (odd columns) and the corresponding fuzzy membership function to the skin class, which result from the application of the framework presented herein (even columns).

6. Conclusions

Section 5.1 shows the applicability of the already presented DoFI methodology in skin detection. Therefore, the *DoFI* proves to be a useful tool in the detection of color clusters in the image domain. The extension of the basic methodology through a color transformation and a fuzzification stage show to be useless. Genetic algorithms can be successfully applied in order to parameterize all the fuzzy integrals involved in the fusion methodology. Special attention have to be devoted to the selection of the fitness function. In this context it is particularly relevant the outperformance of the fitness function related to ROC evaluation. This trend is successfully being used in classification parameterization.

The performance comparison (see Section 5.2) shows that the methodology proposed herein outperforms well-known methods [11,10]. It shows robustness particularly in the case of a non-homogeneous illumination. In this context it is worth mentioning the compensatory capability of the herein presented framework. Furthermore, methods like the one described here give a real value (not a crisp one) as an output, so they may be used as a base for other skin detection methods like [12], for human tracking on videos, or for face detection.

The projective work takes a more systematic analysis of the process for the fitness function selection, which demonstrated to be crucial for the successful performance of the framework. What the training concerns, it would be interesting to find out the minimal number of training samples necessary for successfully determining the value of the fuzzy measure coefficients. Furthermore, the methodology will be tested with other databases in order to attain an even more exact characterization of the framework.

References

[1] A. Albiol, L. Torres, E. Delp, Optimum color spaces for skin detection, in: IEEE Internat. Conf. on Image Processing, vol. 1, 2001, pp. 122–124.

[2] T.-Y. Chen, J.-C. Wang, G.-H. Tzeng, Identification of general fuzzy measures by genetic algorithms based on partial information, IEEE Trans. Systems Man Cybernet. Part B 30 (4) (2000) 517–528.

[3] T.-Y. Cow, K.-M. Lam, Mean-shift based mixture model for face detection on color image, in: IEEE Internat. Conf. on Image Processing, vol. 1, 2004, pp. 601–604.

[4] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, Willey, New York, 2001.

[5] J.-M. Geusebroek, R. van den Boomgaard, A.W. Smeulders, H. Geerts, Color invariance, IEEE Trans. Pattern Anal. Mach. Intelligence 23 (12) (2001) 1338–1350.

[6] T. Gevers, Color based image retrieval, in: M. Lew (Ed.), Principles of Visual Information Retrieval, Springer, Berlin, 2001, pp. 11–49.

[7] D.E. Goldberg, Genetic algorithms in search, Optimization and Machine Learning, Addison-Wesley Pub. Co., Inc., Reading, Massachussets, 1989.

[8] M. Grabisch, Fuzzy measures and integrals for decision making and pattern recognition, in: R. Mesiar, B. Rieca (Eds.), Fuzzy Structures: Current Trends, Mathematical Institute, Slovak Academy of Sciences, Bratislava, 1997, pp. 7–35.

[9] M. Grabisch, H.T. Nguyen, A.A. Walker, Fundamentals of Uncertainty Calculi with Applications to Fuzzy Inference, Kluwer, Dordrecht, Holland, 1995.

[10] R.L. Hsu, M. Abdel-Mottaleb, A. Jain, Face detection in color images, IEEE Trans. Pattern Anal. Mach. Intelligence 24 (5) (2002) 696–706.

[11] M. Jones, J. Rehg, Statistical color models with application to skin detection, Int. J. Comput. Vision 46 (1) (2002) 81–96.

[12] J. Ruiz-del Solar, R. Verschae, Robust skin segmentation using neighborhood information, in: IEEE Internat. Conf. on Image Processing, vol. 1, 2004, pp. 207–210.

[13] L. Sigal, S. Sclaroff, V. Athisos, Skin color-based video segmentation under time-varying illumination, IEEE Trans. Pattern Anal. Mach. Intelligence 26 (7) (2004) 862–877.

[14] A. Soria-Frisch, Soft data fusion for computer vision, Ph.D. Thesis, Technical University Berlin, 2004.

[15] A. Soria-Frisch, Color seal extraction from documents: robustness through soft data fusion, J. Adv. Signal Process. 2005 (13) (2005) 2146–2152.

[16] M. Sugeno, The theory of fuzzy integrals and its applications, Ph.D. Thesis, Tokyo Institute of Technology, 1974.

[17] H. Tahani, J.M. Keller, Information fusion in computer vision using the fuzzy integral, IEEE Trans. Systems Man Cybernet. 20 (3) (1990) 733–741.

[18] Z. Wang, K.-S. Leung, J. Wang, A genetic algorithm for determining nonadditive set functions in information fusion, Fuzzy Sets and Systems 102 (1999) 463–469.

[19] J. Yang, W. Lu, A. Waibel, Detecting human faces in color images, in: IEEE Internat. Conf. on Image Processing, vol. 1, 1998, pp. 127–130.