Robust skin detection based on the fuzzy integral

Aitor Olano
Innovae Vision
aitor@innovaevision.com

Aureli Soria-Frisch
Dept. Security Tech.,
Fraunhofer IPK
aureli.soria-frisch@ieee.org

Rodrigo Verschae
Dept. Electrical Engineering,
Universidad de Chile
rodrigo@verschae.org

Abstract

Skin detection is a low-level operation used in multiple application fields, e.g. video surveillance, human computer interface. The paper presents a multi-procedure methodology followed by a fusion stage that can be applied in the robust detection of skin. Two well-known methodologies are fused with a novel one, which is presented here as well, through the application of the fuzzy integral. The automated parameterization of the resulting procedure and its performance on a benchmark database are finally described.

1 Introduction

The research field of skin detection attains the recognition of the image (or video) pixels that correspond to human skin. Skin detection is applied as a pre-processing stage in applications like video surveillance, internet security or human computer interface. The main difficulties for detecting skin areas are: (1) the existence of skin-like color pixels that may appear in the background and (2) the multi-modality of the skin distribution in the color space when the illumination conditions are not controlled.

One of the possibilities to cope with the problems described in the former paragraph is the combination of different methodologies for skin detection. 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 feasible solution [1]. The fulfillment of this goal depends on the features of the applied fusion operator.

The fuzzy integral was introduced by Sugeno [2]. It presents a very good performance for the realization of the fusion operation in computer vision [1] because of its mathematical properties. Taking the facts formerly mentioned into account, a framework for the robust skin detection that is based on the twofold application of the fuzzy integral is presented herein.

A methodology based on the fuzzy integral was employed in [3] for the detection of ink seals on document images. This methodology, which will be denoted as *Difference of Fuzzy Integrals (DoFI)*, is first extended herein in order to successfully detect skin. Second the fuzzy integral is used in the fusion stage of a multi-procedure methodology, where the *DoFI* is combined with two procedures based on the mixture of gaussians (*MoG*) [4], and on a fuzzified version of the Hsu-methodology [5]. To the best of our knowledge, no fusion approach has been presented hitherto in the the resolution of skin detection.

The paper is organized as follows. Section 2 gives the theoretical background needed in order to understand the presented methodologies. After this, the extension of the *DoFI* and the fusion stage for skin detection are respectively described in Secs. 3 and 4 Finally, some results (Sec. 5) and conclusions (Sec. 6) are given.

2 Theoretical background

An overview on the research field of skin detection is described in the following paragraphs. Thence the application of the fuzzy integral in computer vision is presented.

2.1 Skin detection in computer vision

Skin detection algorithms can be classified in two groups: pixel based and segmentation based. Pixel-based algorithms classify each pixel individually without considering the others pixels of the image. Many different color spaces have been used within this approach, since most of the algorithms bound the skin pixel distribution to a particular color space. Normalized RGB [6], and HSV are the color spaces commonly used [7], although YCbCr seems the most successful one [5]. Other works within this category use statistical models in order to solve the skin/non-skin classification problem. They can be classified as parametric (Mixture of Gaussians –*MoG*) and non-parametric (Histograms) [4]. This classification is sometimes realized in different steps, which delivers much more complex systems [8].

Segmentation based algorithms have been newly presented [9]. They classify each pixel by taking the information in its neighborhood into account. In [9] a diffusion (region growing) algorithm is used to perform the classification.

The following subsections describe two pixel-based methods. These are fused through the fuzzy integral with the results of the *DoFI* methodology, which is presented in Sec. 3, within the fusion paradigm presented herein.

2.1.1 Ellipse

In [5], using the Cb' and Cr' (components a rotated YCbCr color space) an ellipse is *a priori* defined in the feature subspace, where skin-like pixels are supposed to be found. In the YCbCr color space the skin color distribution is more compact that in others color spaces, like RGB or HSV and this allows obtaining a a better segmentation. In this paper the parameters parameters used in [5] are employed. Furthermore the approach is fuzzified in order to obtain a real value, which can be processed by the fuzzy integral.

2.1.2 Mixture of Gaussians (*MoG*)

Statistical methods have shown to outperform other pixel based methodologies. In this work we used the *MoG* provided by [4], which consist of 16 Gaussians in the RGB color space. Thus the following equation is applied on the image pixels x [4]:

$$P(x) = \sum_{i=1}^{16} w_i \frac{e^{-\frac{1}{2}(x-\mu_i)^T \sum_i^{-1}(x-\mu_i)}}{(2\pi)^{3/2}|\sum_i|^{1/2}} \qquad (1)$$

2.2 The fuzzy integral for computer vision

The fuzzy integral [2] constitutes a generalization of different fusion operators [1]. A fuzzy integral presents three elements in its mathematical expression: the values to be integrated, the coefficients of the fuzzy measure $\boldsymbol{\mu}$, which is the weighting function used in the operator, and two fuzzy connectives. The fuzzy connectives are the characterizing factor among different types of fuzzy integral [10]. The so-called Choquet Fuzzy Integral makes use of the sum ($\sum$) and the product ($\cdot$):

$$\mathcal{C}_\mu[x_1,..,x_n] = \sum_{i=1}^{n} x_{(i)} \cdot [\mu(A_{(i)}) - \mu(A_{(i-1)})]. \quad (2)$$

The coefficients of the fuzzy measure are denoted as $\mu(A_{(i)}) = \mu(\{x_{(1)},...,x_{(i)}\})$. Fuzzy measures generalize classical measures, i.e. probability measures, by relaxing its additivity axiom. Being $\mathcal{X}$ the set of information sources each fuzzy measure coefficient, $\mu(A_i)$, characterizes the *a priori* importance of each subset $A_i \subset \mathcal{X}$. Hence, fuzzy measures are functions on fuzzy sets, $\mu : \mathcal{P}(\mathcal{X}) \to [0,1]$, satisfying in the discrete case the following conditions :

I. $\mu\{\emptyset\} = 0$; $\mu\{\mathcal{X}\} = 1$, and

II. $A_j \subset A_k \to \mu(A_j) \leq \mu(A_k)\ \forall A_j, A_k \in \mathcal{P}(\mathcal{X})$.

A fuzzy measure presents 2^n coefficients, so many as subsets can be formed among the input information sources, e.g. the color values in the channels R, G, and B for a color image represented in the RGB color model. From all the coefficients just n are taken into account in each integration. These are selected upon the sorting operation denoted by the enclosed subindices in the expression (2). If e.g. the three input channels fulfill $x_B > x_G > x_R$, then $x_{(1)} = x_B$, $x_{(2)} = x_G$, and $x_{(3)} = x_R$, and the coefficients $\mu(\{x_B\})$, $\mu(\{x_G, x_B\})$ and $\mu(\{x_R, x_G, x_B\})$ will be taken into account. This sorting is responsible of the definition of a set of weighting coefficients for each canonical region of the feature space. A canonical region is an area of the feature space, where a particular ordinal relationship among the feature components is fulfilled. In computer vision this property improves the robustness of the operator w.r.t. a change in the illumination conditions [1].

3 Difference of fuzzy integrals (*DoFI*) and its extension for skin detection

A recent work [3] presents a methodology used for the segmentation of ink seals, which are characterized by a particular color. Since the segmentation is based on the difference of two fuzzy integrals, the methodology is denoted as Difference of Fuzzy Integrals (*DoFI*). In the application of the *DoFI* for skin detection, which is presented herein, the element to be segmented is characterized by skin hues.

The subtracted fuzzy integrals differ on the fuzzy measure used on them, $\boldsymbol{\mu}^1$ and $\boldsymbol{\mu}^2$. Particularly the value of the coefficients that control the canonical region occupied by the skin color cluster changes. This change ideally attains that just the pixels within the skin cluster are modified, i.e. unfortunately other pixels are taken as skin pixels as well. The block diagram of the *DoFI* procedure is depicted in Fig. 1.

Taking the existent frameworks for skin detection based on the utilization of different color models (see Sec. 2.1) as a starting point, 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: $I = \{I_R, I_G, I_B\} \rightarrow \{C_1, C_2, C_3\}$, where C_i are the

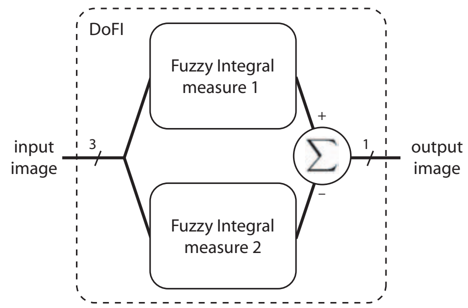

Figure 1: Block diagram of the framework for the segmentation of particular color clusters denoted as *Difference of Fuzzy Integrals*, *DoFI* [3]. A fuzzy integral is firstly computed with respect to two different fuzzy measures $\boldsymbol{\mu}^1$ and $\boldsymbol{\mu}^2$.

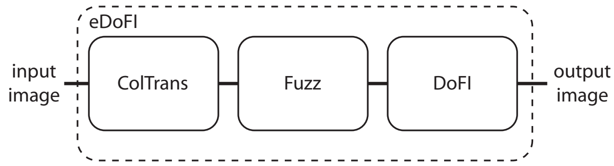

Figure 2: Block diagram of the here presented framework for skin detection, which is denoted as *Extended Difference of Fuzzy Integrals*, *eDoFI.* The framework extends an already presented framework [3] (see Fig. 1) through the application of a color space transformation (*ColTrans*) and an automated fuzzification process (*Fuzz*).

color channels of the new color representation, e.g. *rgb*, *YCbCr*. Thence the module *Fuzz* undertakes the fuzzification of this color representation and delivers F_i (see Fig. 2) as stated by $\{C_1, C_2, C_3\} \rightarrow \{F_1, F_2, F_3\}$. This module has been added to the original *DoFI* methodology [3] as well. *Fuzz* can be easily implemented by applying Parzen windows on a training image set [1].

Being $\{F_1(x, y), F_2(x, y), F_3(x, y)\}$ the fuzzified color image, a difference image $D(x, y)$ is obtained in the *DoFI* methodology by applying:

$$D(x, y) = \|\mathcal{F}_{\mu^1}(x, y) - \mathcal{F}_{\mu^2}(x, y)\|, \qquad (3)$$

where $\mathcal{F}_{\mu^i}$ states for the images resulting from the computation of the fuzzy integral with respect to the fuzzy measure $\boldsymbol{\mu}^i$ on each color pixel, as expressed by (2) for $x_1 = F_1(x_i, y_i)$, $x_2 = F_2(x_i, y_i)$, and $x_3 = F_3(x_i, y_i)$. The application of the pro-

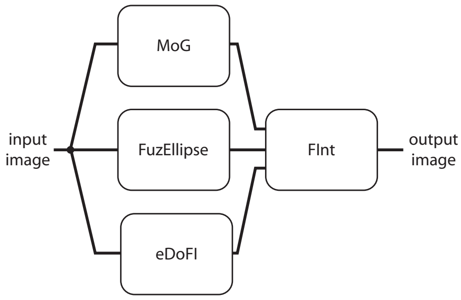

Figure 3: Block diagram of the fusion framework based on the fuzzy integral (*FInt*) and used for skin detection. Acronyms in text.

cedure 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 skin.

The automated parameterization of the methodology, which takes the determination of the coefficients of the fuzzy measures $\boldsymbol{\mu}^1$ and $\boldsymbol{\mu}^2$ into consideration, can be undertaken by applying genetic algorithms [11]. This methodology has been successfully used in the parameterization of the fuzzy integral [3][12]. The process is described in Sec. 5.1.

4 Robust skin detection through fusion

A fusion methodology is applied herein in order to improve the performance of the skin detection. The presence of this stage is motivated by the different nature of the simple skin detection procedures formerly presented (see Secs. 2.1.1, 2.1.2, and 3). Each methodology processes the image pixels in a different manner. Therefore they deliver complementary results (see Fig. 6). Taking this fact into consideration the skin detection is finally realized by applying a fuzzy integral on the results delivered by these methodologies (see Fig. 3).

Since the fusion through the fuzzy integral is computed w.r.t. a fuzzy measure $\boldsymbol{\mu}^f$, this stage adds a new parameter to the general methodology. The coefficients of $\boldsymbol{\mu}^f$ are automatically assessed by applying again genetic algorithms (see Sec. 5.1). It is worth mentioning that the fusion stage is parameterized once the parameters of the *eDoFI* methodology have been set up.

5 Performance evaluation

Different databases have been used in the analysis of results, two in the training phase, and one for performance evaluation. For training 44 human skin photos were downloaded from internet and manually segmented, obtaining 350,000 different skin color values. This colors correspond to caucasian human beings of different ages. This database is denoted in the following paragraphs as *Sk*. A similar amount of non-skin values were selected from a database of more than 2,000,000 non-skin color values. These pixels form the non-skin database *NSk*. The test database, which has been used as benchmark in other works, consisted on 427 ground truth labeled images taken from popular movie sequences [7].

5.1 Parameterization of the methodologies

The best results have been obtained by using the Choquet fuzzy integral w.r.t. general fuzzy measures. The optimal threshold, which is needed in order to defuzzify the result of the fuzzy integral in the skin/non-skin classification, is codified in the genomes together with the fuzzy measures ($\boldsymbol{\mu}^1$, $\boldsymbol{\mu}^2$, and $\boldsymbol{\mu}^f$).

A *Steady State* genetic algorithm showed the best performance. The probability of replacement has been set in the interval $[0.8, 0.9]\%$ of the population, which presents a number of individuals in the interval [50,500]. Probabilities were set as 0.9 for crossover and 0.2 for mutation. The genomes present 8 bits for each value to be determined. The GA were run with a *Roulette Wheel* selection scheme [11] and *2-point* crossover [11] until a maximum of 2000 generations.

Different fitness functions have been tested. The first function tries to maximize the output value O of the algorithm being parameterized for the pixels p_i within the skin database *Sk* and to minimize it for those pixels within the non-skin database *NSk*. This is motivated by the fact that the algo-

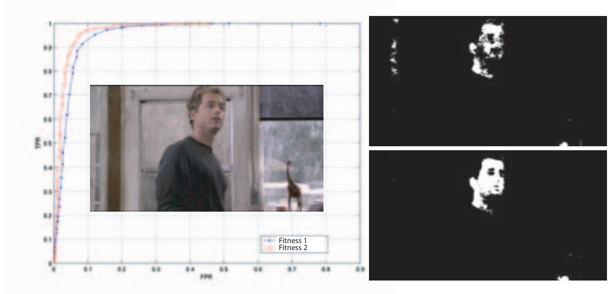

Figure 4: Comparison of the performance of the two presented fitness functions all over the database (ROC graphic) and for a particular image of the database (in graphic). (top) Result for ϕ_1. (bottom) Result for ϕ_2.

rithm should deliver a maximal membership function for the skin class in the skin pixels. The output value is normalized by the result for all pixels. This function can be summarized as:

$$\phi_1 = \frac{\sum_{p_i \in Sk} O(p_i)}{\sum_{p_i \in DB} O(p_i)} + (1 - \frac{\sum_{p_i \in NSk} O(p_i)}{\sum_{p_i \in DB} O(p_i)}), \quad (4)$$

where $DB = Sk \cup NSk$.

The second one attains the computation of the maximal difference between the *True Positive Rate* (*TPR*) and *False Positive Rate* (*FPR*) after binarizing the fused data with the threshold, which can be expressed as:

$$\phi_2 = 1 + (TPR - FPR). \quad (5)$$

5.2 Analysis of results

The selection of a suitable fitness function was first attained. A comparison between the performance of the *eDoFI* parameterized with the two fitness functions over the test database (see Fig. 4) shows the convenience of using the second one ϕ_2. The same conclusion was experimentally set for the fusion procedure.

A comparison among the results obtained within a *eDoFI* (see Sec. 3) on different color spaces for a parameterization based on ϕ_2 are depicted in Fig. 5. As it can be observed the best results are attained on the color space YCrCb.

Further results (see Fig. 6) compare the performance of the individual methodologies and of the

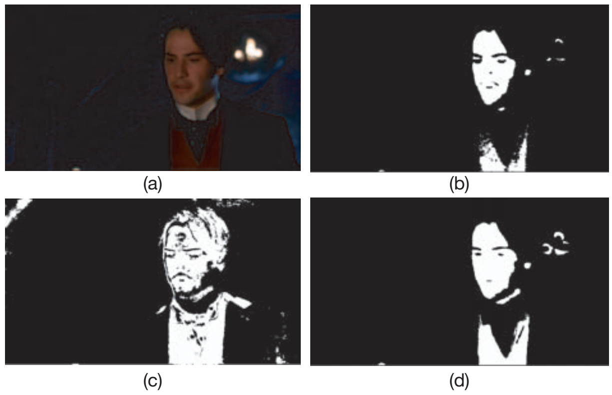

Figure 5: Results of the difference of fuzzy integrals extended herein (*eDoFI*) on different color spaces. (a) Input image [7]. (b) RGB-color space. (b) Normalized RGB-color space. (c) YCrCb-color space.

fusion procedure. The compensatory effect of the fusion can be observed in Fig. 6e. Some of the background parts segmented as skin by the fuzzy-ellipse (see Fig. 6c) are suppressed in the fusion result by taking the result of the *eDoFI* methodology into account (see Fig. 6d), 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 face in spite of the presence of a light reflection, whose detection is not achieved neither by the *MoG* nor by the fuzzy-ellipse. The detection of this reflected skin part by the *eDoFI* methodology is due to its robustness with respect to the luminance level mentioned in the introduction.

The compensatory effect of the fusion methodology can be observed in Fig. 7 as well, where a ROC of the results of each methodology all over the test database [7] is depicted. The different points of the ROC are obtained for different thresholds.

6 Conclusions and projective work

The proposed methodology outperforms well-know ones [4][5] in the resolution of skin detection. It shows robustness w.r.t. the illumination. This property results from the weighting scheme of the fuzzy integral formerly mentioned. It is worth mentioning as well the compensatory prop-

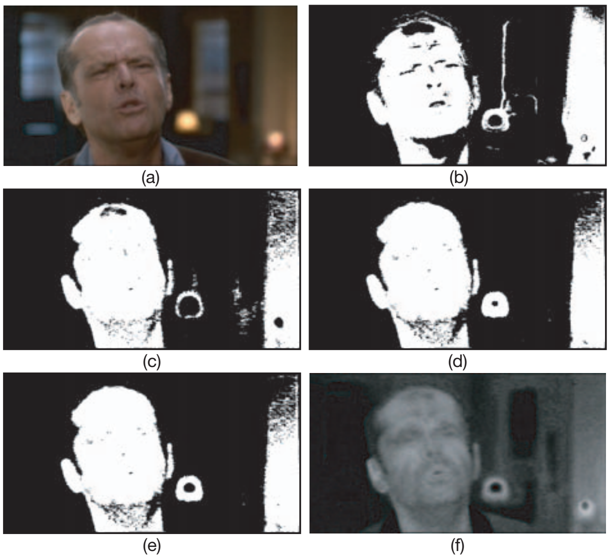

Figure 6: Comparison of the results obtained for different methodologies. (a) Input image [7]. (b) *MoG*. (c) FuzEllipse. (d) *eDoFI*. (e) Fusion of the former ones, which results from the binarization of the fuzzy integral result (f)).

erty of the fuzzy integral as fusion operator, which was described on hand of the results. Being the fuzzy integral parameterized through genetic algorithms, especial attention have to be devoted to the selection of the fitness function. The improvement of the parameterization procedure will be treated in the future.

References

[1] A. Soria-Frisch, *Soft Data Fusion for Computer Vision*, Ph.D. thesis, Technical University Berlin, 2004.

[2] M. Sugeno, *The Theory of Fuzzy Integrals and Its Applications*, Ph.D. thesis, Tokyo Institute of Technology, 1974.

[3] A. Soria-Frisch, "Color seal extraction from documents: robustness through soft data fusion," *J. of Advanced Signal Proc.*, in press.

[4] M.J. Jones and J.M. Rehg, "Statistical color models with application to skin detection," *Int. Journal of Computer Vision*, vol. 46(1): 81-96, 2002.

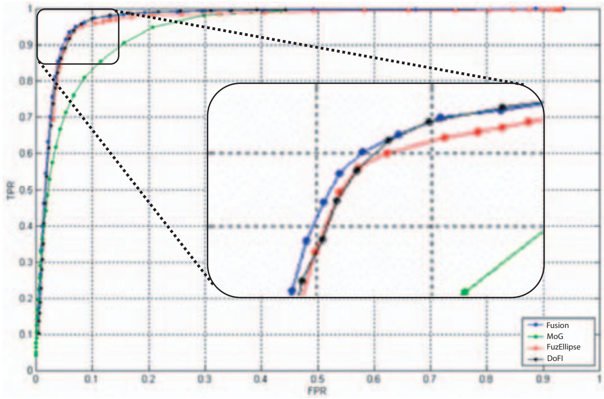

Figure 7: ROC of the database [7] for all methodologies used herein (see legend) with detail.

[5] R. L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face detection in color images," *PAMI*, vol. 24(5), pp. 696–706, 2002.

[6] A. Albiol, L. Torres, and Ed. Delp, "Optimum color spaces for skin detection," in *IEEE Int. Conf. Image Proc.*, 2001.

[7] L. Sigal, S. Sclaroff, and V. Athisos, "Skin color-based video segmentation under time-varying illumination," *PAMI*, vol. 26 (7), pp. 862–877, 2004.

[8] Q. Zhu, K.-T. Cheng, C.-T. Wu, and Y.-L. Wu, "Adaptive learning of an accurate skin-color model," in *Proc. 6th Int. Conf. Autom. Face & Gesture Recog.*, 2004, pp. 37–42.

[9] J. Ruiz del Solar and R. Verschae, "Robust skin segmentation using neighborhood information," in *11th Int. Conf. Im. Proc.*, 2004.

[10] M. Grabisch, H.T. Nguyen, and A. Walker, *Fundamentals of Uncertainty Calculi with Applic. to Fuzzy Inference*, Kluwer, 1995.

[11] D. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, 1989.

[12] T. Chen, J. Wang, and G. Tzeng, "Identification of General Fuzzy Measures by Genetic Algorithms Based on Partial Information," *IEEE Trans. SMC - B*, vol. 30 (4), pp. 517–528, 2000.