Spatial-based skin detection using discriminative skin-presence features ☆

Michal Kawulok [*,1], Jolanta Kawulok [2], Jakub Nalepa [1]

Institute of Informatics, Silesian University of Technology, Akademicka 16, 44-100 Gliwice, Poland

ABSTRACT

In this paper we propose a new method for skin detection in color images which consists in spatial analysis using the introduced texture-based discriminative skin-presence features. Color-based skin detection has been widely explored and many skin color modeling techniques were developed so far. However, efficacy of the pixel-wise classification is limited due to an overlap between the skin and non-skin pixels reported in many color spaces. To increase the discriminating power of the skin classification schemes, textural and spatial features are often exploited for skin modeling. Our contribution lies in using the proposed discriminative feature space as a domain for spatial analysis of skin pixels. Contrary to existing approaches, we extract the textural features from the skin probability maps rather than from the luminance channel. Presented experimental study confirms that the proposed method outperforms alternative skin detection techniques, which also involve analysis of textural and spatial features.

© 2013 The Authors. Published by Elsevier B.V. All rights reserved.

1. Introduction

The goal of skin detection is to determine whether a given video sequence, an image, a region or a pixel presents the human skin. The applications are of a wide range and significance, including gesture recognition and human–computer interaction (Bilal et al., 2012; Radlak and Smolka, 2012; Nalepa et al., 2014), objectionable content filtering (Lee et al., 2007), image retrieval (Kruppa et al., 2002), image coding using regions of interest (Chen et al., 2003), and many more.

Skin detection is a challenging problem that has been extensively studied over the years, but no satisfactory solution has been proposed so far. The existing techniques are based on the premise that the skin color can be effectively modeled in various color spaces, which in turn allows for segmenting the skin regions. Although the color features constitute the primary source of information for skin detection, the effectiveness of color-based classification is limited due to a significant overlap between the skin and

non-skin pixels that appears in all of the popularly used color spaces. Skin color depends on a number of individual factors, and also intra-personal differences may be substantial, mainly because of variations in lighting conditions. Furthermore, the background objects often possess skin-like color, which results in observing false positive errors in the segmentation outcome.

It has been reported that the discriminating power of skin detectors can be improved by employing additional features extracted from the texture or relying on spatial analysis of pixels classified as skin. Basically, the skin regions are usually smooth and form consistent areas, in contrast to many background objects that contain pixels of color values similar or identical to the skin. Also, skin detection errors can be reduced if a color model is adapted to a particular scene or an individual, for example based on detected facial or hand regions.

In this paper we explore how to exploit the textural features to establish a suitable propagation domain for the spatial analysis of the skin pixels. The proposed method is based on the discriminative features extracted from skin probability maps, obtained using a conventional Bayesian classifier (Jones and Rehg, 2002). This is in contrast to alternative texture-based techniques which rely on the texture of grayscale images. During our research we identified that the discriminating power of the textural features is definitely higher if they are extracted from the skin probability maps rather than from the luminance channel. This approach was not investigated in alternative studies.

Here, we introduce the *discriminative skin-presence features* (DSPFs), derived from the *discriminative textural features* (DTFs) described in our earlier works on skin detection (Kawulok, 2012) and

image colorization (Kawulok et al., 2012). The differences between the DTFs and the DSPFs, discussed later in Section 4, allow for the computation time reduction and efficacy improvement. The most important contribution of this work lies in using the DSPF space as a domain for propagating the "skinness", which allows for a substantial increase of the detection rate.

First, we transform the skin probability map using the DSPF space, and then apply the spatial analysis in the transformed skin map. For the spatial analysis we adapted an algorithm based on the distance transform in a combined domain (DTCD) of hue, luminance and skin probability, developed during our earlier study (Kawulok, 2013). Originally, the DTCD algorithm operates in raw skin probability maps, but here it has been adapted to perform the propagation in the DSPF space. The proposed method was compared with another approach proposed by Jiang et al. (2007), who also exploit the textural features, which is followed by the spatial analysis. Not only do the results indicate that our method outperforms the alternative algorithms, but also the gain attributed to the spatial analysis is larger than in case of processing raw skin probability maps.

The paper is organized as follows. In Section 2 an overview of existing skin detection methods is presented with particular attention given to the texture-based methods. Section 3 describes the existing algorithms which form the background for the proposed method. Our contribution to the skin detection is described in Section 4. The results of experimental validation are reported and discussed in Section 5, and finally our research is concluded in Section 6.

2. Related work

Existing color-based skin segmentation techniques take advantage from the observation that skin-tone color has common properties which can be defined in various color spaces. In general, skin color detectors rely on rule-based or statistical skin modeling. A thorough survey comparing various color-based skin detection approaches was presented by Kakumanu et al. (2007).

There are a number of methods which are based on fixed decision rules defined in various color spaces after analyzing skin-tone distribution. The rules are applied after color normalization to determine if a pixel color value belongs to the skin. Skin-tone color was modeled in the *HSV* color space by Tsekeridou and Pitas (1998). Kovac et al. (2003) proposed a model defined in the *RGB* color space. An approach introduced by Hsu et al. (2002) takes advantage of common skin color properties in nonlinearly transformed YC_bC_r color space using an elliptical skin color model. A technique operating in multiple color spaces to increase the stability was described by Kukharev and Nowosielski (2004). Cheddad et al. (2009) proposed reducing the *RGB* color space to a single dimension, in which the decision rules are defined.

Statistical modeling is based on analysis of skin pixel values distribution for a training set of images, in which skin and non-skin areas are identified and annotated. This creates a global model of skin color, which allows determining the probability that a given pixel value belongs to the skin class. Skin color can be modeled using a number of techniques, including the Gaussian mixture model (Greenspan et al., 2001) and the Bayesian classifier (Jones and Rehg, 2002). The latter method, given more attention in Section 3.1, was used in the research reported here to generate the skin probability maps.

There are a number of adaptive models that are designed to decrease the impact of overlaps between skin and non-skin pixels in a color space, which improves the segmentation accuracy for a presented scene. In general, adaptive approaches can be divided according to the information source used for the adaptation, i.e.

skin-like object tracking, whole-image analysis, face and hand detection, and the model type being adapted, i.e. threshold-based, histogram analysis, Gaussian mixture models.

Lee et al. (2007) proposed a method based on a multi-layer perceptron extracting lighting features from an analyzed image to adjust the skin detector. An approach for adapting the segmentation threshold in the probability map based on the assumption that a skin region is coherent and should have homogenous textural features was introduced by Phung et al. (2003). This method was further extended by Zhang et al. (2004) by involving the artificial neural network (ANN) for estimating an optimal acceptance threshold. The ANN was also used for adaptation by Yang et al. (2010). A method for a dynamic model adaptation based on observed changes in the histogram extracted from a tracked skin region was proposed by Soriano et al. (2000). Motion detectors for the skin color model adaptation were explored by Dadgostar and Sarrafzadeh (2006). Analysis of facial regions for effective adaptation of the skin model to local conditions was investigated by Fritsch et al. (2002); Stern and Efros, 2002; Kawulok, 2008; Kawulok et al., 2013 and Yogarajah et al., 2012.

Analysis of textural features extracted from an input image was applied to improve the performance of color-based methods. In the approach proposed by Wang et al. (1985) segmentation in the *RGB* and YC_gC_b color spaces is enhanced by analyzing various textural features, including contrast, entropy, homogeneity and more, extracted using the gray-level co-occurrence matrix (GLCM). The experimental results reported in the original work showed that the method is competitive for complex background detection, but the skin detection rate was significantly worse compared to color-based approaches. Additionally, the time complexity of calculating GLCM is proportional to $O(g^2)$ (Clausi and Jernigan, 1998), where g is the number of gray levels of the input grayscale image.

An interesting algorithm incorporating color, texture and space analysis was given by Jiang et al. (2007). Initially, skin probability map color filter with a low acceptance threshold is applied in the *RGB* color space. Then, textural features are extracted using the Gabor wavelets from an input color image converted to the grayscale. The obtained response is subject to a threshold Θ_T, which produces a binary texture mask. The aim of applying the texture mask is to reduce the false positive rate by filtering out the regions with large values of the texture feature, i.e. those that are not as smooth as skin, but were improperly classified as skin by the color filter. Finally, skin regions are grown using the watershed segmentation with well-defined region markers to exploit the spatial information. It was shown that the method reduced the false positive rate (from 20.1% to 4.2%) with simultaneous increase of the true positive rate (from 92.7% to 94.8%) compared to the color filtering for a data set containing 600 images. However, the authors did not provide any sensitivity analysis and it seems that a different threshold value is applied to every image, which makes it difficult to get satisfactory results for a larger number of images. Additionally, if human skin is not smooth (e.g. in case of the elders), then it may also be filtered out by the texture filter. On the contrary, if the skin-like background is smooth, then the misclassified pixels are not filtered out.

Simple textural features were used to boost the performance of a number of skin detection techniques and classifiers, including the ANN (Taqa and Jalab, 2010), non-parametric density estimation of skin and non-skin classes (Zafarifar et al., 2010), Gaussian mixture models (Ng and Pun, 2011), and many more (Forsyth and Fleck, 1999, Conci et al., 2008, Fotouhi et al., 2009, Abin et al., 2009). Generally, the analysis of texture in an input image helps reducing the number of pixels misclassified by pixel-wise color-based detectors. However, the roughness of skin and non-skin regions can vary among images, which in turn makes the response of a

texture-based segmentation algorithm difficult to generalize for real-life data sets. In none of the mentioned methods the skin probability maps were analyzed with regards to their textural features.

Although the color-based skin models can be efficiently adapted to a given image, it was proved by Zhu et al. (2004) that it is hardly possible to separate skin from non-skin pixels using such approaches. It is easy to see that skin pixels are usually grouped into blobs whereas the non-skin false positives are scattered around the spatial domain. A number of skin segmentation techniques emerged based on this observation: Kruppa et al. (2002) assumed that the skin blobs are of an elliptical shape, a threshold hysteresis was applied by Argyros and Lourakis (2004) and recently by Baltzakis et al. (2012). Conditional random fields were used by Chenaoua and Bouridane (2006) to exploit spatial properties of skin regions. An approach based on the cellular automata for determining skin regions was proposed by Abin et al. (2009).

The analysis of skin probability map domain for skin segmentation using a controlled diffusion was proposed by del Solar and Verschae (2004). Here, the diffusion seeds are extracted at first. They are formed by those pixels, whose skin probability, extracted from the pixel-wise skin probability maps, exceeds the *seed threshold* (P_α). Then, the skin regions are built according to the criteria of the diffusion process. A neighboring pixel p_j is adjoined to the source pixel p_i if (1) a distance between p_i and p_j in the diffusion domain is smaller than a given *diffusion threshold* (Δ) and (2) the skin probability of p_j is larger than the *propagation threshold* (P_β). The main drawback of this method is its performance in case of blurry region boundaries, since the diffusion process does not stop if the transitions between skin and non-skin pixels are smooth.

In our earlier research (Kawulok, 2010) we introduced an energy-based technique for skin blobs analysis. The pixels are adjoined to the skin regions depending on the amount of the energy which is spread over the image according to the local skin probability. Recently, we proposed to use the distance transform in a combined domain (DTCD) of hue, luminance and skin probability (Kawulok, 2013). The algorithm was proved to be very competitive and outperformed our energy-based method and the method proposed by del Solar and Verschae (2004). We overcame the most significant shortcoming of the latter approach, i.e. misbehaving in case of smooth transitions between skin and non-skin regions, by taking advantage of the cumulative character of the distance transform. This method is exploited in the research reported here, and therefore it is given more attention in Section 3.3.

3. Theoretical background

In this section, three methods, which form the basis for the proposed technique, are given attention, namely: (1) the Bayesian skin classifier, (2) linear discriminant analysis (LDA), and (3) spatial analysis using the distance transform.

3.1. Bayesian skin classifier

In the presented study, the skin probability maps were obtained using Bayesian skin modeling introduced by Jones and Rehg (2002). The method consists in analyzing the color histograms of the skin and non-skin pixels, and the skin probability, given a certain color value, is determined using the Bayes rule. During training, the probability is computed for every possible color value and a look-up table is generated, which allows for converting an input color image into a skin probability map.

At first, based on a training set, histograms for the skin (C_s) and non-skin (C_{ns}) classes are built. The probability of observing a given color value (v) in the C_x class can be computed from the histogram:

$$P(v|C_x) = C_x(v)/N_x, \tag{1}$$

where $C_x(v)$ is the number of v-colored pixels in the class x and N_x is the total number of pixels in that class. Maximal number of histogram bins depends on the pixel bit-depth and for most color spaces it equals $256 \times 256 \times 256$. However, it was reported beneficial (Phung et al., 2005; Kawulok et al., 2014) to reduce the number of bins per channel, thus, in our research we used 64 bins per each channel in the *RGB* color space.

It may be expected that a pixel presents the skin, if its color value has a high density in the skin histogram. Moreover, the chances for that are larger, if the pixel color is not very frequent among the non-skin pixels. Taking this into account, the probability that a given pixel value belongs to the skin class is computed using the Bayes rule:

$$P(C_s|v) = \frac{P(v|C_s)P(C_s)}{P(v|C_s)P(C_s) + P(v|C_{ns})P(C_{ns})}, \tag{2}$$

where *a priori* probabilities $P(C_s)$ and $P(C_{ns})$ may be estimated based on the number of pixels in both classes, but very often it is assumed that they both equal $P(C_s) = P(C_{ns}) = 0.5$. The learning phase consists in creating a skin color probability look-up table, which maps every color value in the color space domain into the skin probability. After training, using the look-up table, an input image is converted into a skin probability map, in which skin regions may be segmented based on an *acceptance threshold* (P_{acc}). The threshold value should be set to provide the best balance between the false positives and false negatives, which may depend on a specific application.

3.2. Linear discriminant analysis

Linear discriminant analysis (Seber, 1984) is a supervised statistical feature extraction method frequently used in machine learning, and here it was applied to extract the discriminative textural features. It finds a subspace defined by the most discriminative directions within a given training set of M-dimensional vectors classified into K classes. The analysis is performed first by computing two covariance matrices: intra-class scatter matrix:

$$S_W = \sum_{i=1}^{K} \sum_{\mathbf{u}_k \in K_i} (\mathbf{u}_k - \boldsymbol{\mu}_i)(\mathbf{u}_k - \boldsymbol{\mu}_i)^T \tag{3}$$

and inter-class scatter matrix:

$$S_B = \sum_{i=1}^{K} (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T, \tag{4}$$

where $\boldsymbol{\mu}$ is a mean vector of the training set and $\boldsymbol{\mu}_i$ is a mean vector of the ith class (termed K_i). Subsequently, the matrix $S = S_W^{-1} S_B$ is subjected to the eigen decomposition $S = \Phi\Lambda\Phi^T$, where $\Lambda = diag(\lambda_1, \ldots, \lambda_M)$ is the matrix with the ordered eigenvalues along the matrix diagonal and $\Phi = [\mathbf{v}_1 | \ldots | \mathbf{v}_M]$ is the matrix with the correspondingly ordered eigenvectors as columns. The eigenvectors form the orthogonal basis of the feature space.

Originally, the feature space has M dimensions, but only those associated with the highest eigenvalues have strong discriminative power, while the remaining can be rejected. In this way the dimensionality is reduced from M to m, where $m < M$.

After having built the m-dimensional feature space, the feature vectors are obtained by projecting the original vectors $\mathbf{u}$ onto the feature space: $\mathbf{v} = \Phi^T \mathbf{u}$. The similarity between the feature vectors is computed based on their Euclidean distance in the feature space.

3.3. Spatial analysis of skin pixels using the distance transform

In the research reported in this paper we used a spatial analysis method which we developed during our earlier works (Kawulok, 2013). It consists of two general phases, namely: (1) seed extraction and (2) propagation using the distance transform.

The seeds are extracted taking advantage of the observation that if the image is binarized using a high-probability threshold, then the false positives are rather small, because usually only real skin regions contain pixels with very high skin probability values. If the skin probability of an individual pixel is over a high threshold P_α, then the pixel is added to the seed. After that, the seed pixels are grouped into blobs, and those blobs whose area is smaller than 10% of the largest blob are rejected.

In order to propagate the "skinness" from the seeds, first the shortest routes from the seed to every pixel are determined. This is achieved by minimizing *total path costs* from the set of seed pixels to every pixel in the image. The total path cost for a pixel x is defined as:

$$C(x) = \sum_{i=0}^{l-1} \rho(p_i \rightarrow p_{i+1}), \tag{5}$$

where ρ is a *local skin dissimilarity measure* between two neighboring pixels, p_0 is a pixel that lies at the seed boundary, $p_l = x$, and l is the total path length. The minimization is performed using the Dijkstra's algorithm as proposed by Ikonen and Toivanen (2007). In addition, P_β threshold is used as proposed by del Solar and Verschae (2004), which prevents the propagation to the regions of very low skin probability.

The route optimization outcome heavily depends on how the local costs ρ are computed. For skin detection, the local cost from pixel x to y, i.e. $\rho(x \rightarrow y)$ is obtained using both the image (ρ_I) and the probability (ρ_P) costs:

$$\rho(x \rightarrow y) = \rho_I(x, y) \cdot [1 + \rho_P(x \rightarrow y)], \tag{6}$$

$$\rho_I(x, y) = \alpha_{diag} \cdot (|Y(x) - Y(y)| + |H(x) - H(y)|), \tag{7}$$

$$\rho_P(x \rightarrow y) = \begin{cases} 1 - P(y) & \text{for } P(y) > P_\beta, \\ \infty & \text{for } P(y) \leqslant P_\beta, \end{cases} \tag{8}$$

where $Y(\cdot)$ is the pixel luminance, $H(\cdot)$ is the hue in the *HSV* color model, and $\alpha_{diag} \in \{1, \sqrt{2}\}$ is the penalty for propagation in the diagonal direction. The total path cost obtained after the optimization is inversely proportional to the "skinness", hence the final skin probability map is obtained by scaling the costs from 0 for the maximal cost to 1 for a zero cost (i.e. the seed pixels). The pixels which are not adjoined during the propagation process are assigned with zeroes. Finally, the skin regions are extracted using a fixed threshold in the distance domain.

4. Proposed texture-based skin detection method

Distribution of skin color has been effectively modeled in a number of color spaces and chrominance combined with the luminance is considered as the most distinctive skin feature that underpins virtually all of the existing skin detection methods. However, the discriminating capability of the color is limited because of its high variance within the skin class and also high similarity of many background objects to the skin. Hence, this imposes a certain upper bound on the effectiveness of color-based skin models. In general, they produce the higher false positive rates, the more universal a skin model is expected to be. In order to decrease the false positives, the skin model should take advantage of other distinctive features in addition to the pixel-wise color-based classification.

In our research we explored how to exploit the textural features to extend the conventional color models. Although the human skin is characterized by a distinctive pattern, it can be observed only in high-quality images of very high resolution. Otherwise, the skin regions usually appear plain, without any strong textural features, which sometimes can be seen in case of background objects having the skin-like color. This forms a basis for the existing texture-based methods that in general decrease the skin probability, if the textural features are apparent.

Our contribution is based on the observation that in many cases the false positives present relatively smooth texture in the color and luminance domains, however it is amplified in the skin color probability map. This is illustrated in Fig. 1. Here, an input color image (a) is converted to the grayscale (b), and to the skin probability map (c) using the Bayesian classifier (the darker shade in the map indicates the higher probability). The magnified non-skin region is characterized by a smooth texture with small variations in both the color and grayscale domains. However, it can be noticed that the variations of the skin probability in Fig. 1(c) are quite high in the region, which appears as a kind of a texture pattern that cannot be observed for the real skin areas. Also, the values in the probability map are generally high within the region, many of them exceeding the probability of the real skin areas (e.g. the boy's legs or neck), which would result in false positives after applying the acceptance threshold. A similar observation can be made in case of many images, which leads to a general conclusion that the textural features extracted from the probability map may be helpful for skin detection. During our experiments we found it much more effective to analyze the texture of the probability maps compared to the texture of color or grayscale images.

In order to determine which textural features offer the best discrimination between the skin and non-skin regions, we adapted a framework which we developed for the image colorization purposes (Kawulok et al., 2012). First, we compute simple image statistics in several kernels of different dimensions. This forms the *basic image features* (BIFs) which are subsequently subject to LDA, producing the *discriminative textural feature* (DTF) space. Using the DTF space, an input skin probability map can be transformed into the DTF skin map. The latter allows for better separation
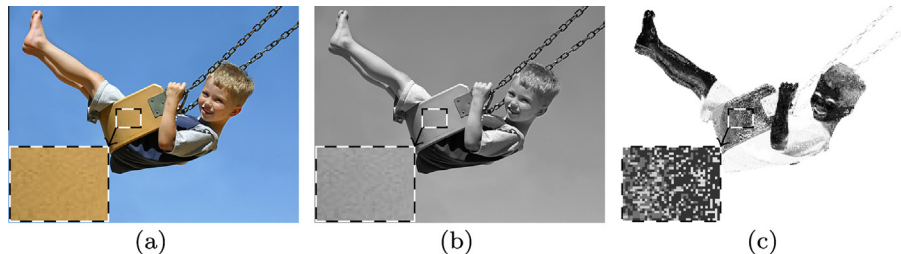
Fig. 1. An input color image (a), its appearance in the grayscale (b) and the obtained skin probability map (c). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

between the skin and non-skin pixels, resulting in lower detection errors, which was demonstrated in (Kawulok, 2012). In the research reported here, the basic image features were refined so as to decrease the computation time, and also raw skin probability value was appended to the BIF vector. After applying the LDA projection, the *discriminative skin-presence features* (DSPFs) are obtained. Furthermore, we observed that defining the local costs (ρ_p) in the DSPF space constitutes a very effective domain for the spatial analysis outlined earlier in Section 3.3.

4.1. Discriminative skin-presence features

In order to determine the discriminative skin-presence features, at first simple statistics are computed from every pixel's neighborhood in the probability map using kernels of different sizes. Here, we determined on the experimental basis that the following four features are the most relevant for skin detection purposes: (a) the *median* and (b) *minimal* values, (c) *standard deviation*, and (d) the *difference between the maximum and minimum* computed in each kernel. In addition, the raw skin probability value is appended to the feature vector. These features are supposed to indicate the skin probability value, as well as its variance in the neighborhood of the processed pixel. The BIFs are obtained at three different scales, namely: 5×5, 9×9 and 13×13 pixels. Hence, every pixel x is transformed into an M-dimensional basic feature vector u_x, where $M = 13$ in the presented case. The quoted basic features, as well as the dimensions of the kernels, were selected carefully on the basis that they delivered the best results for the reported case, however a different combination may be optimal for other applications. We considered including other features like entropy or local binary patterns (Ojala et al., 2002) into the BIFs, as they are often used for texture segmentation, however it did not improve the results, whereas the computational costs were increased. Overall, the simple statistics are sufficient to effectively extract the roughness of skin probability maps, which increases the discrimination power of the skin model.

Skin and non-skin pixels are transformed into two classes of feature vectors which are subsequently subject to LDA in order to increase their discriminating power. This training process, illustrated in Fig. 2(a), creates a discriminative feature space that is later used to generate the DSPF skin maps. First, for a given training set of images and associated skin masks, the probability look-up table is obtained using the Bayesian skin modeling (Jones and Rehg, 2002). After that, all the images from the training set are transformed into skin probability maps, from which the BIFs are extracted. Finally, based on the basic feature vectors labeled using the ground-truth skin masks, the LDA projection matrix is computed.

In Fig. 3, an example of an image decomposed into the basic features (a), as well as its projection onto the DSPF space (b), are presented. Here, we obtained the DSPF space using 2000 images as explained later in Section 5. The relation between the eigenvalues is illustrated in Fig. 3(c). It can be seen that two leading components inherit the majority of the discriminating power, which was also reflected in the results reported later in Section 5. It is worth to note that a pixel's BIFs projection onto the DSPF space does not indicate explicitly whether the pixel presents skin, and some reference points need to be given in the DSPF space to perform the classification. Basically, the DSPF space is expected to minimize the distances within the skin and non-skin classes, while maximizing the distances between them.

4.2. Skin detection using the DSPF space

Skin detection operates following three general steps, namely: (1) skin probability map computation using the Bayesian model, (2) generation of the skin map using the DSPF space, and (3) spatial analysis based on the DSPF skin map. A flowchart for the DSPF skin map generation is presented in Fig. 2(b) and it is also illustrated using an example in Fig. 4.

First of all, an original image (Fig. 4(a)) is converted into the skin probability map (Fig. 4(b)) using the look-up table obtained after training the Bayesian model. Subsequently, every pixel from the probability map is projected onto the DSPF space using the LDA projection matrix (an example of the projection outcome is shown in Fig. 3(b)).

The DSPF skin map is obtained based on the distance, computed in the DSPF space, of every pixel x from a *reference pixel r*:

$$D(x) = \left[\sum_{i=1}^{m} \left(v_i^{(x)} - v_i^{(r)} \right)^2 \right]^{1/2}, \tag{9}$$

where $v_i^{(x)}$ is the ith dimension of the DSPF vector obtained for the pixel x. The reference pixel is determined as a pixel of the maximal probability value in the skin probability map subjected to the erosion using a large 15×15 kernel (Fig. 4(c)). The erosion is necessary, because the reference pixel should not be an isolated skin spot, as the surrounding non-skin pixels could disturb the textural features. This operation assures that the reference pixel is determined so as its DSPFs are extracted exclusively from the pixels of high skin probability value. Therefore, the erosion kernel dimension is greater than the largest kernel used for extracting the BIFs (i.e. 13×13). Furthermore, the probability value of the reference pixel must be above the *reference-point threshold* (P_{ref}) in order to properly deal with the images which do not present the human skin at all.
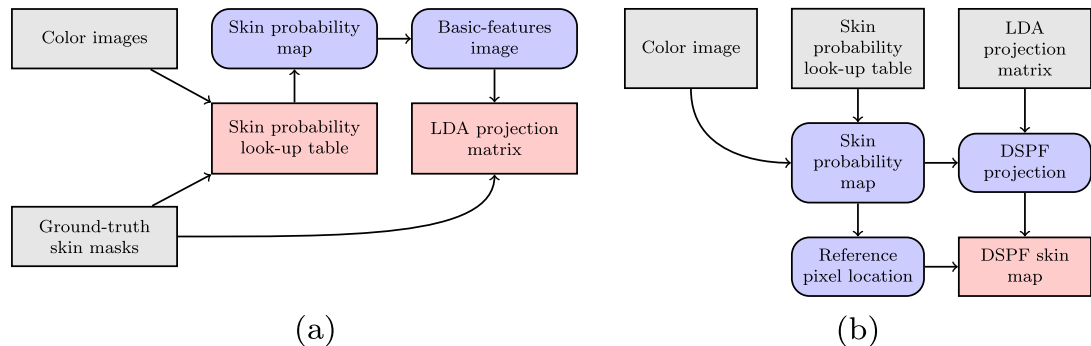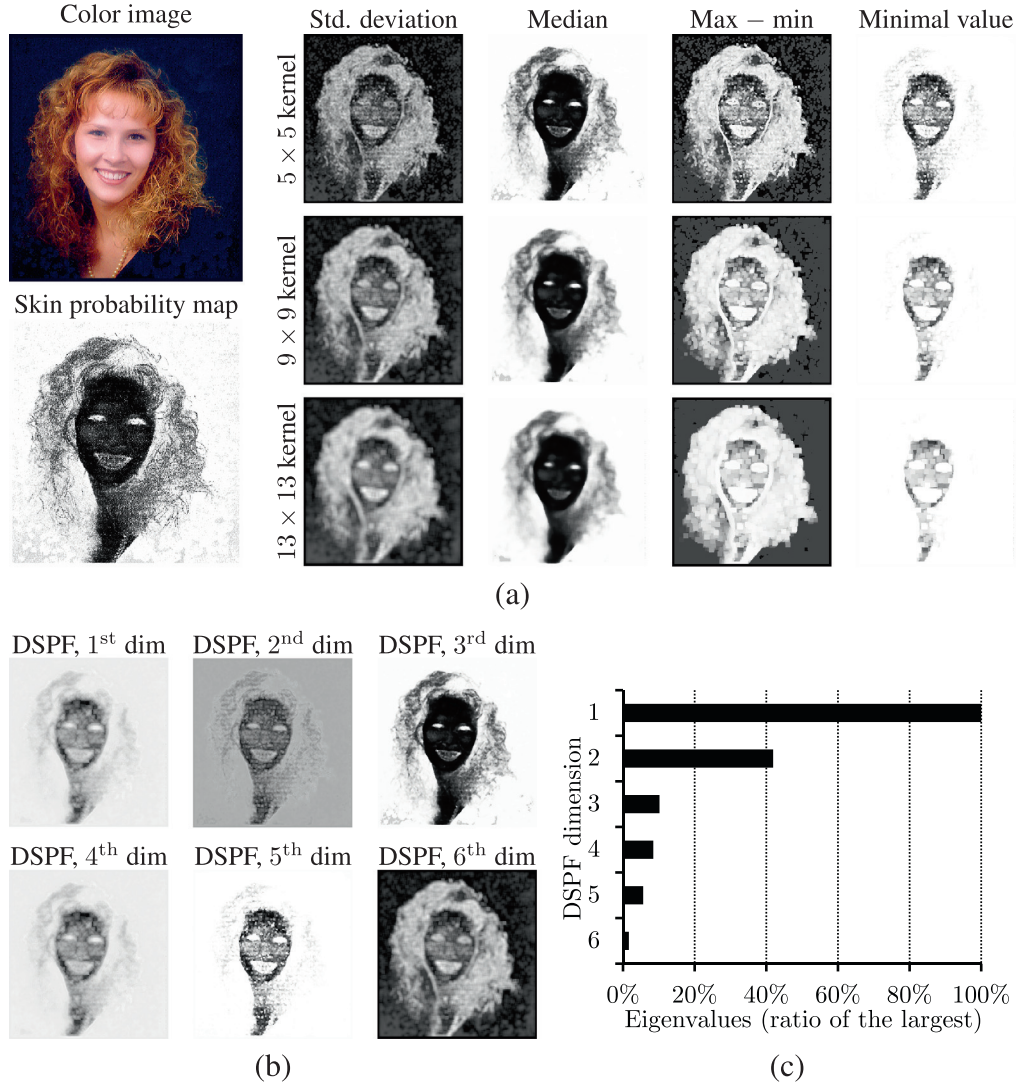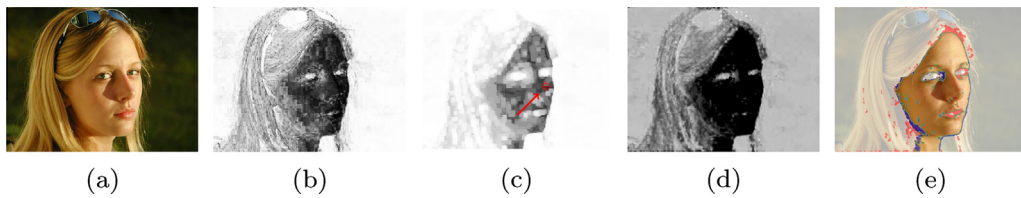

Fig. 2. A flowchart of the training (a) and detection (b) processes.

(a)

(b)　　　　　　　　　　　　　　　　(c)

Fig. 3. An example of an image decomposed into the BIFs (a), and obtained projections (b) onto the six leading dimensions of the DSPF space, whose eigenvalues are shown in (c).

(a)　　　　　　(b)　　　　　　(c)　　　　　　(d)　　　　　　(e)

Fig. 4. An example of generating the DSPF skin maps: (a) an input color image, (b) the skin probability map, (c) the reference pixel in the eroded probability map, (d) the DSPF skin map, and (e) obtained segmentation result using a fixed threshold (red tone: false positives; blue tone: false negatives; faded color: true negatives; boundary of the skin region: green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Finally, the DSPF skin map (Fig. 4(d)) is obtained based on the distances from the reference pixel in the DSPF space, scaled from 1 for the reference pixel to 0 for the maximal distance (D_{max}):

$$D_N(x) = (D_{max} - D(x))/D_{max}, \tag{10}$$

The scaling assumes that the image contains some non-skin pixels, but this condition is met virtually in all of the real-life images.

The obtained DSPF skin map offers higher separation between the skin and non-skin pixels, which overall decreases the detection errors when the acceptance threshold is applied (Fig. 4(e)). However, the detection errors can be further reduced if the spatial analysis is performed as explained in Section 3.3. In this case, the probability cost is computed in the normalized distance map:

$$\rho_P(x \to y) = \begin{cases} 1 - D_N(y) & \text{for } D_N(y) > P_\beta, \\ \infty & \text{for } D_N(y) \leqslant P_\beta. \end{cases} \tag{11}$$

As the DSPF skin map is obtained on the basis of the features that discriminate well between the skin and non-skin pixels, it constitutes a definitely better propagation domain than the conventional skin probability map. Also, the DSPF maps are normalized, which means that it is easier to determine an optimal value of the seed

threshold P_α. In this way, the seeds are formed by the pixels of a very high similarity in the DSPF domain to the reference pixel. An example of the spatial analysis using the DSPF skin map is presented in Fig. 5. The distance transform is performed from the seeds (Fig. 5(b)), and the "skinness" is obtained by scaling the distance map (Fig. 5(c)) from 1 for the seeds to 0 for the largest distance.

5. Experimental validation

The experiments were carried out for two data sets, namely: 4000 images from the ECU database (Phung et al., 2003), and 899 hand images (termed HGR) which we registered for the gesture recognition purposes (available at http://sun.aei.polsl.pl/~mkawulok/gestures). The images from both data sets are associated with ground-truth skin binary masks indicating skin regions. The ECU images were acquired in uncontrolled lighting conditions, and skin-color objects often appear in the background, which makes the skin regions difficult for segmentation. The HGR data set contains images of gestures presented by 12 different people. The data were acquired in controlled conditions, but in some cases the subjects were dressed in clothes of a skin-like color.

The ECU data set was split into two equinumerous subsets, each containing 2000 images. The first subset was used for training the Bayesian classifier and to determine the discriminative space from the BIFs, while the second one was used for validation. The file lists are available at http://sun.aei.polsl.pl/~mkawulok/dtf_skin. In order to decrease the computational requirements for training the DSPF space, we sampled every 15th pixel from every 15th row in each image. In this way we obtained a representative training set, whereas the time and memory costs remained at acceptable levels.

Skin detection performance was assessed based on two errors, namely: (1) *false positive rate* (δ_{fp}), i.e. a percentage of background pixels misclassified as skin, and (2) *false negative rate* (δ_{fn}), i.e. a percentage of undetected skin pixels. These errors depend on the value of the acceptance threshold (P_{acc}), and their mutual relation can be presented using a *receiver operating characteristic* (ROC). Also, we quote the *minimal detection error*:

$$\delta_{\min} = \delta_{fp} + \delta_{fn}, \tag{12}$$

which is obtained by setting the acceptance threshold (P_{acc}) to a value, for which this sum is minimal. Obviously, a uniform value of P_{acc} was applied to every image in the test set within each experiment. Furthermore, we verified whether and how the skin presence affects the false positive rate. This was proceeded by excluding the skin regions from processing for the ECU and HGR data sets, and the skin-excluded false positive rate (δ_{fp}^{SE}) was measured. Here, the same P_{acc} was applied, with which $\delta_{\min}$ was obtained for the whole set, including the skin regions. The experiments were conducted using a computer equipped with an Intel Core i7-3740QM 2.7 GHz (16 GB RAM) processor.

The proposed algorithm was compared with the Gabor wavelet-based texture analysis method proposed by Jiang et al. (2007). As it was mentioned in Section 2, the texture map threshold Θ_T is set

individually for every image presented in the original work. We tried to determine an optimal value of the threshold investigating the range $5 \leqslant \Theta_T \leqslant 100$, which covers all of the values quoted by the authors, but the overall results were quite poor. We overcame this problem by multiplying the skin probability map by the normalized textural filter response (0 for the maximal value, and 1 for the minimal), instead of applying the threshold Θ_T. This delivered sensible results which are quoted here. We used this method without and with the spatial analysis, however in the latter case we applied our distance transform-based method (i.e. DTCD) instead of the watershed segmentation, in order to provide a fair comparison regarding the texture analysis. Furthermore, we compared the method with the Bayesian classifier (Jones and Rehg, 2002), diffusion-based spatial analysis (DSA) introduced by del Solar and Verschae (2004) and our DTCD method performed in the skin probability maps obtained using the Bayesian classifier (Kawulok, 2013). Moreover, we extracted the BIFs from the luminance (we termed this option DSPF-L), as well as both from the luminance and skin probability maps (termed DSPF-L+). For DSPF-L, the dimensionality of the basic feature vectors remained the same ($M = 13$), while for DSPF-L+ it increased to $M = 25$, because the BIFs extracted from the luminance were appended to the features extracted from the skin probability maps. In both cases, the raw skin probability was included into the BIFs as specified earlier in Section 4.1. In order to provide a thorough comparison, we also investigated how the wavelet-based method behaves when the textural features are extracted from the skin probability maps rather than from the luminance as proposed in the original work of Jiang et al. (2007).

In Table 1 we present the minimal detection error obtained for different dimensionality (m) of the DSPF space without and with the spatial analysis applied. It can be seen that the best results are obtained using $m = 1$ or $m = 2$ dimensions, which is coherent with the eigenvalues presented earlier in Fig. 3(c). As the spatial analysis delivered the lowest detection error for $m = 2$ dimensions, all the results presented further in this Section were obtained using this setting. The thresholds for the proposed method (i.e. DSPF), as well as its variants (i.e. DSPF-L and DSPF-L+) were set as follows: $P_\alpha = 0.96$, $P_\beta = 0.3$, $P_{ref} = 0.6$, and they were constant in all of the experiments, whose results are reported here. In the case of the DTCD performed in the raw skin probability map, the threshold P_α was set to 0.6 for the ECU data set and to 0.4 for the HGR set as suggested in Kawulok (2013).

ROC curves for the ECU and HGR data sets are shown in Fig. 6, and the error rates are quoted in Table 2. In the case of the DSA (del Solar and Verschae, 2004), the ROC curves were obtained by applying different values of the diffusion threshold (Δ), explained earlier in Section 2. The errors measured for each particular value of Δ are marked with asterisks on the graphs. As the DTCD delivers much better results than the DSA, it was adapted for the spatial analysis in the case of the texture-based methods. It can be observed that the proposed method (DSPF with DTCD) outperforms all other techniques, offering substantial error reduction for both investigated data sets. Overall, the minimal error (δ_{min}) is reduced by 8.95% for the ECU data set (from 24.91% to 15.96%) and by
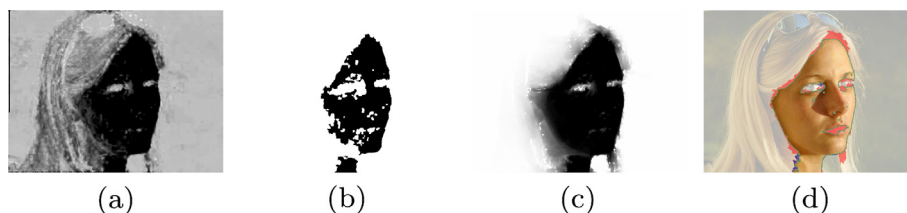

(a)   (b)   (c)   (d)

Fig. 5. An example of the distance transform using the DSPF skin map (a): the determined seeds (b), the skin map (c), and the obtained segmentation result (d).
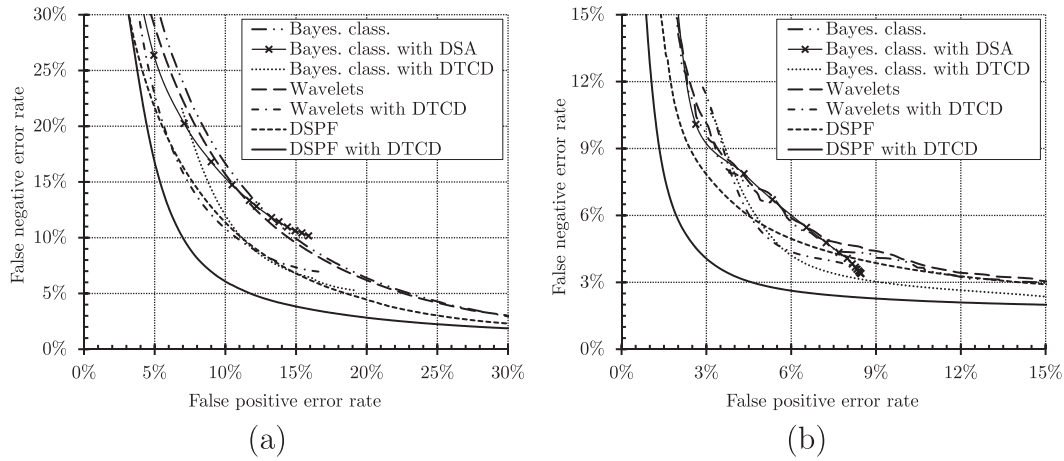
Table 1
Minimal error rate (δ_{min}) obtained for the ECU data set depending on the DSPF space dimensionality. Bold values indicate the best score.

	Number of DSPF dimensions (m)				
	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)
Fixed threshold	**20.95**	21.13	21.28	23.05	23.60
Spatial analysis	15.99	**15.96**	16.03	16.64	16.88

4.56% for the HGR data set (from 11.62% to 7.06%). Furthermore, the computation times are quoted for the ECU and HGR sets (an average image size is 516 × 542 pixels for ECU and 457 × 445 for HGR). Naturally, the Bayesian classifier offers a matchless processing speed, making it suitable for the most of real-time applications. Sequential implementation of the proposed method allows for

processing 2—3 frames per second. This is definitely faster than the wavelet-based method, which requires several seconds to process a single image. Also, it is worth noting that computing the BIFs involves many independent operations, hence this process can be easily parallelized and accelerated.

It can also be seen that in the case of the methods based on the discriminative features, the analysis of the skin probability maps delivers definitely better results compared with the analysis of the grayscale image. DSPF-L, which analyzes the texture of the luminance channel, delivers rather poor results. They are significantly better, when the BIFs are extracted both from the luminance and skin probability maps (DSPF-L+), and they are even better when the luminance is excluded at all (DSPF). Using the wavelet-based method, the analysis of the skin probability maps is slightly more effective in the case of the HGR set, and a little bit worse for

Fig. 6. ROC curves obtained for the ECU (a) and HGR (b) data sets.

Table 2
Skin detection errors obtained using different methods. The gain (G_{sa}) presents the error reduction score attributed to the spatial analysis (DSA or DTCD) within each group. Bold values indicate the best score in each column. Italics indicate the published baseline methods.

Method ↓	ECU data set						HGR data set					
	t (ms)	δ_{fn} (%)	δ_{fp} (%)	$\delta_{fp}^{(SE)}$ (%)	δ_{min} (%)	G_{sa} (%)	t (ms)	δ_{fn} (%)	δ_{fp} (%)	$\delta_{fp}^{(SE)}$ (%)	δ_{min} (%)	G_{sa} (%)
Bayes. class.	**6**	9.97	14.93	14.93	24.91	–	**4**	6.72	4.90	4.90	11.62	–
with DSA	176	12.81	12.20	10.15	25.01	−0.10	84	**3.39**	8.46	6.78	11.85	−0.23
with DTCD	167	8.72	12.22	11.66	20.94	3.97	89	4.35	5.84	4.62	10.19	1.43
Wavelets	5510	10.45	14.00	14.00	24.45	–	4010	5.90	6.00	6.00	11.89	–
with DTCD	5664	10.25	10.53	8.39	20.78	3.67	4105	4.92	5.24	4.41	10.16	1.73
Wavelets (s.p.m.)	6119	11.44	13.84	13.84	25.28	–	4034	6.04	5.70	5.70	11.74	–
with DTCD	6256	12.04	10.53	12.44	22.57	2.71	4125	5.97	3.74	3.03	9.71	2.03
DTF	425	8.44	11.94	11.00	20.39	–	295	3.98	3.89	2.15	7.87	–
with DTCD*	582	**6.28**	9.97	7.75	16.25	4.14	389	4.34	3.39	1.65	7.72	0.15
DSPF-L	273	19.96	15.78	15.41	35.75	–	181	24.50	7.21	2.45	31.71	–
with DTCD*	425	13.49	13.40	11.33	26.89	**8.86**	313	17.67	5.37	1.77	23.03	**8.68**
DSPF-L+	486	10.10	11.25	10.56	21.35	–	342	8.20	3.74	2.08	11.94	–
with DTCD*	662	7.47	**9.08**	7.19	16.56	4.79	467	5.29	**2.39**	1.61	7.68	4.25
DSPF	257	9.65	11.48	10.49	21.14	–	176	6.17	4.30	2.15	10.47	–
with DTCD*	432	6.75	9.21	**7.01**	**15.96**	5.18	316	3.95	3.11	**1.60**	**7.06**	3.41

*local costs (ρ_P) obtained using Eq. (11).
Bayes. class. – Bayesian classifier (Jones and Rehg, 2002).
DSA – diffusion-based spatial analysis (del Solar and Verschae, 2004).
DTCD – distance transform in a combined domain (Kawulok, 2013).
Wavelets – wavelet-based method (Jiang et al., 2007).
Wavelets (s.p.m.) – wavelet-based method applied to the skin probability maps.
DTF – discriminative textural features (Kawulok, 2012).
DSPF-L – the BIFs extracted from the luminance.
DSPF-L+ – the BIFs extracted both from the luminance and skin probability maps.
DSPF – the BIFs extracted from the skin probability maps.

the ECU images. The table also includes the comparison with the DTF method (Kawulok, 2012). Although this earlier approach performs better if the distance transform is not applied ($\delta_{\min} = 20.39\%$ compared with 21.14% for DSPF), the improved BIF extraction scheme offers lower error than DTF with DTCD. Moreover, the computation time is also reduced here due to a smaller number of the basic image features.

Furthermore, in Table 2 we quote the *gain* (G_{sa}) attributed to the spatial analysis. It can be seen that the gain is larger when the DTCD local costs are computed in the discriminative domain as given in Eq. (11). G_{sa} is the largest in the case of DSPF-L, but here the minimal errors are the highest anyway. In the case of DSPF, the gain equals 5.18% and 3.41% compared with 3.97% and 1.43% for the Bayesian classifier for the ECU and HGR sets, respectively. This shows that the introduced discriminative feature space constitutes a better domain for the spatial analysis than the raw skin probability map.

Several examples of the skin maps and corresponding segmentation results obtained using different methods are demonstrated in Fig. 7. Images (III.) and (IV.) are from the HGR data set, and the rest come from the ECU set. Correctly detected skin regions (i.e. true positives) are annotated with green boundaries in Fig. 7(e)–(h), a red tone indicates false positives, and a blue tone – false negatives. True negatives are shown with faded colors. It can be seen from the skin maps (Fig. 7(b)–(d)) that the proposed

technique increases the separability between the skin and non-skin regions (c), which is even intensified in (d). As a result, the proposed method reduces both false positives and false negatives significantly. Although the final outcomes still contain some errors, they are definitely smaller compared with the results obtained using the alternative techniques.

The false positive rates obtained for both data sets after excluding the skin regions from processing (δ_{fp}^{SE}) are presented in Table 2. For the pixel-wise methods, the presence of skin pixels does not influence the results obtained for non-skin pixels, hence $\delta_{fp}^{SE} = \delta_{fp}$. However, the approaches that involve spatial analysis or detection in the discriminative feature space are sensitive to the presence of skin pixels, so the obtained false positive rates are different. This is caused by two general reasons. Firstly, in many images the thresholds P_α or P_{ref} are not exceeded by the non-skin pixels, resulting in $\delta_{fp}^{SE} = 0$, and secondly, when these thresholds are exceeded by some non-skin pixels, the number of false positives increase, because the distance map is scaled and more non-skin pixels exceed the acceptance threshold P_{acc}. The overall scores indicate that in most cases, including DSPF with DTCD, $\delta_{fp}^{SE} < \delta_{fp}$, which means that if no skin regions are present in an image, the false positive rate is likely to be lower than if the image contains some skin pixels. This observation is also supported with some results obtained for non-skin images presented in Fig. 8. It can be seen that for the majority of the images, the false positives were eliminated using the proposed
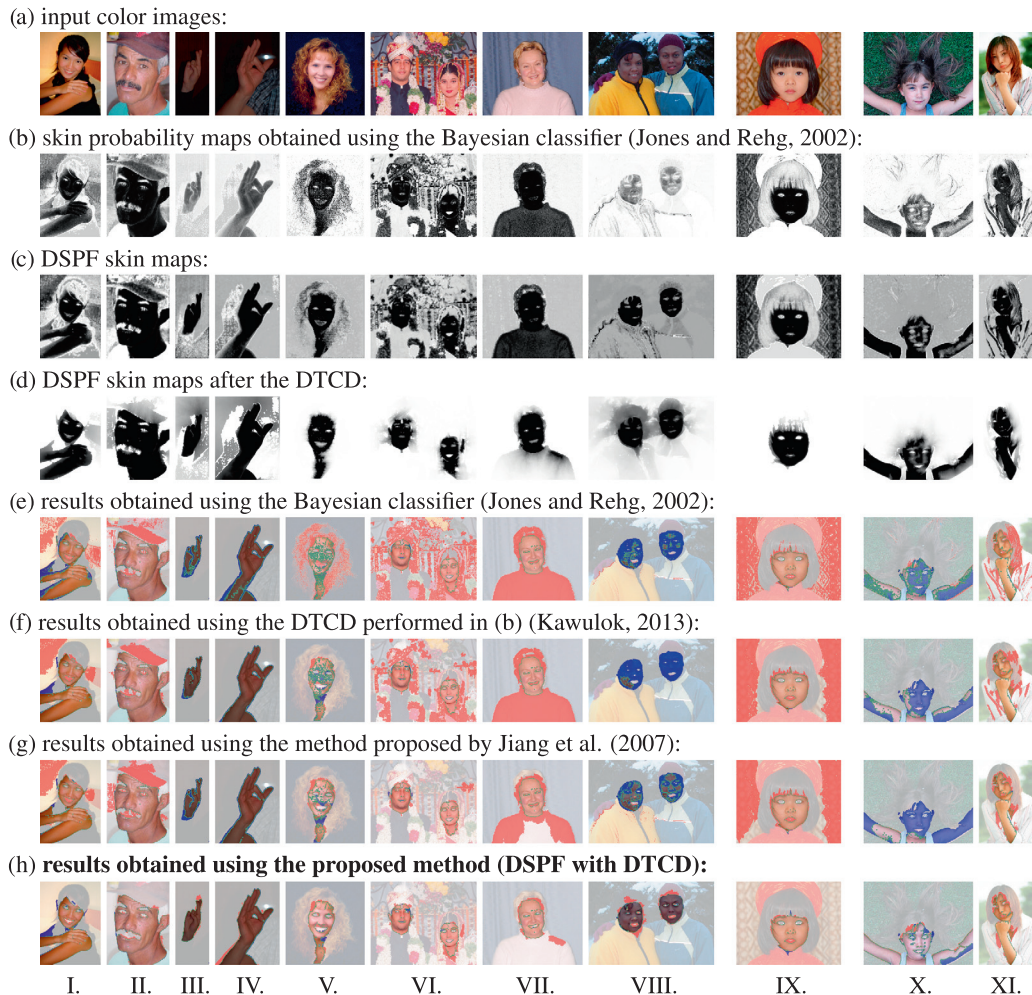
Fig. 7. Examples of skin probability maps (b–d) and the results (e–h) obtained using different methods (red tone: false positives; blue tone: false negatives; faded color: true negatives; boundary of the skin region: green). Images (III.) and (IV.) are from the HGR set and the remaining examples are from the ECU image data set. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
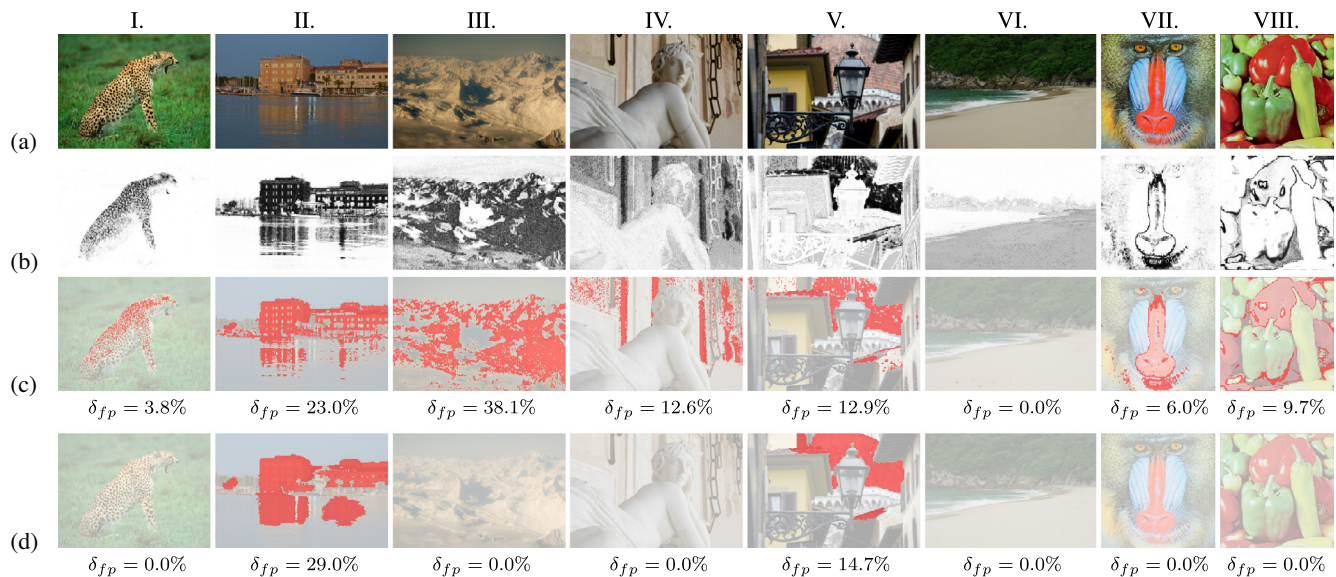
Fig. 8. Examples of skin detection outcome for images without any skin presence: original images (a), skin probability maps obtained using the Bayesian classifier (b), fixed-threshold outcome (c) and results obtained using DSPF with DTCD (d).

approach, however otherwise the false positive error was larger compared with the Bayesian classifier (images II. and V.).

6. Conclusions and future work

In this paper a new method that substantially improves the skin detection outcome has been proposed. This was achieved by exploiting the discriminative features extracted from the skin probability maps, and using them in our spatial analysis scheme. The presented experimental results proved that our method achieves much better results comparing with current state-of-the-art approaches.

A potential drawback of the presented method lies in a validation-based approach towards selecting the basic features prior to learning the DSPF space. Our ongoing research is focused on adapting the feature selection schemes to optimize the type of the features used, as well as the scales at which they should be computed. Finally, our plan is to propose a parallel implementation of the proposed method, which would make it applicable in real-time systems.

References

Abin, A.A., Fotouhi, M., Kasaei, S., 2009. A new dynamic cellular learning automata-based skin detector. Multimed. Syst. 15 (5), 309–323.

Argyros, A.A., Lourakis, M.I.A., 2004. Real-time tracking of multiple skin-colored objects with a possibly moving camera. In: Proceedings of ECCV. LNCS, vol. 3023. Springer, pp. 368–379.

Baltzakis, H., Pateraki, M., Trahanias, P., 2012. Visual tracking of hands, faces and facial features of multiple persons. Mach. Vis. Appl. 23, 1141–1157.

Bilal, S., Akmeliawati, R., Salami, M.J.E., Shafie, A.A., 2012. Dynamic approach for real-time skin detection, J. Real Time Image Proc. Springer-verlag.

Cheddad, A., Condell, J., Curran, K., Mc Kevitt, P., 2009. A skin tone detection algorithm for an adaptive approach to steganography. Signal Process., 0165-1684 89 (12), 2465–2478.

Chenaoua, K., Bouridane, A. 2006. Skin detection using a Markov random field and a new color space. In: Proceedings of IEEE ICIP, pp. 2673–2676.

Chen, M.-J., Chi, M.-C., Hsu, C.-T., Chen, J.-W. 2003. ROI video coding based on H.263+ with robust skin-color detection technique. In: IEEE International Conference on Consumer Electronics, pp. 44–45.

Clausi, D., Jernigan, M., 1998. A fast method to determine co-occurrence texture features. IEEE Trans. Geosci. Remote Sens. 36 (1), 298–300.

Conci, A., Nunes, E., Pantrigo, J.J., Sánchez, Á. 2008. Comparing color and texture-based algorithms for human skin detection. In: Proceedings of ICEIS, pp. 166–173.

Dadgostar, F., Sarrafzadeh, A., 2006. An adaptive real-time skin detector based on Hue thresholding: a comparison on two motion tracking methods. Pattern Recognit. Lett. 27 (12), 1342–1352.

del Solar, J.R., Verschae, R. 2004. Skin detection using neighborhood information. In: Proceedings of IEEE FG, pp. 463–468.

Forsyth, D.A., Fleck, M.M., 1999. Automatic detection of human nudes. Int. J. Comput. Vis. 32, 63–77.

Fotouhi, M., Rohban, M., Kasaei, S. 2009. Skin detection using contourlet-based texture analysis. In: Proceedings of IEEE ICDT, pp. 59–64.

Fritsch, J., Lang,S., Kleinehagenbrock, M., Fink, G., Sagerer, G. 2002. Improving adaptive skin color segmentation by incorporating results from face detection. In: Proceedings of IEEE ROMAN, pp. 337–343.

Greenspan, H., Goldberger, J., Eshet, I., 2001. Mixture model for face-color modeling and segmentation. Pattern Recognit. Lett., 0167-8655 22, 1525–1536.

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

Ikonen, L., Toivanen, P., 2007. Distance and nearest neighbor transforms on gray-level surfaces. Pattern Recognit. Lett. 28 (5), 604–612.

Jiang, Z., Yao, M., Jiang, W., 2007. Skin detection using color, texture and space information. Proc. IEEE FSKD 3, 366–370.

Jones, M., Rehg, J., 2002. Statistical color models with application to skin detection. Int. J. Comput. Vis. 46, 81–96.

Kakumanu, P., Makrogiannis, S., Bourbakis, N.G., 2007. A survey of skin-color modeling and detection methods. Pattern Recognit. 40 (3), 1106–1122.

Kawulok, M., 2008. Dynamic skin detection in color images for sign language recognition. In: Proceedings of ICISP. LNCS, vol. 5099. Springer, pp. 112–119.

Kawulok, M., 2010. Energy-based blob analysis for improving precision of skin segmentation. Multimed. Tools Appl. 49 (3), 463–481.

Kawulok, M., 2012. Texture analysis for skin probability maps refinement. In: Proceedings of MCPR. LNCS, vol. 7329. Springer, pp. 75–84.

Kawulok, M. 2013. Fast propagation-based skin regions segmentation in color images. In: Proceedings of IEEE FG, 1–7.

Kawulok, M., Kawulok, J., Smolka, B., 2012. Discriminative textural features for image and video colorization. IEICE Trans. Inf. Syst. 95-D (7), 1722–1730.

Kawulok, M., Kawulok, J., Nalepa, J., Papiez, M. 2013. Skin detection using spatial analysis with adaptive seed. In: Proceedings of IEEE International Conference on Image Processing (ICIP 2013), 3720–3724.

Kawulok, M., Nalepa, J., Kawulok, J. 2014. Skin region detection and segmentation in color images. In: Celebi, M.E., Smolka, B. (Eds.), Advances in Low-Level Color Image Processing. Lecture Notes in Computational Vision and Biomechanics, vol. 11. Springer, in press.

Kovac, J., Peer, P., Solina, F. 2003. Human skin color clustering for face detection. In: Computer as a Tool, vol. 2, pp. 144–148.

Kruppa, H., Bauer, M.A., Schiele, B., 2002. Skin patch detection in real-world images. In: Proceedings of DAGM Symposium on Pattern Recognition. LNCS, vol. 2449. Springer-Verlag, pp. 109–117.

Kukharev, G., Nowosielski, A., 2004. Fast and efficient algorithm for face detection in colour images. Mach. Graphics Vis., 1230-0535 13, 377–399.

Lee, J.-S., Kuo, Y.-M., Chung, P.-C., Chen, E.-L., 2007. Naked image detection based on adaptive and extensible skin color model. Pattern Recognit., 0031-3203 40, 2261–2270.

Nalepa, J., Grzejszczak, T., Kawulok, M. 2014. Wrist localization in color images for hand gesture recognition. In: Man-Machine Interactions 3. Advances in

Intelligent Systems and Computing, vol. 242, Springer International Publishing, pp. 79–86.

Ng, P., Pun, C.-M. 2011. Skin color segmentation by texture feature extraction and K-mean clustering. In: Proceedings of IEEE CICSyN, pp. 213–218.

Ojala, T., Pietikinen, M., Menp, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. 24 (7), 971–987.

Phung, S.L., Chai, D., Bouzerdoum, A. 2003. Adaptive skin segmentation in color images. In: Proceedings of IEEE ICASSP, pp. 353–356.

Phung, S., Bouzerdoum, A., Chai, D., 2005. Skin segmentation using color pixel classification: analysis and comparison. IEEE Trans. Pattern Anal. Mach. Intell. 27 (1), 148–154. http://dx.doi.org/10.1109/TPAMI.2005.17.

Radlak, K., Smolka, B. 2012. A novel approach to the eye movement analysis using a high speed camera. In: Second International Conference on Advances in Computational Tools for Engineering Applications (ACTEA'12), Zouk-Mosbeh, Lebanon, pp. 139–144.

Seber, G., 1984. Multivariate Observations. Wiley, New York.

Soriano, M., Martinkauppi, B., Huovinen, S., Laaksonen, M., 2000. Skin detection in video under changing illumination conditions. In: Proceedings of IEEE ICPR 1, 839–842.

Stern, H., Efros, B. 2002. Adaptive color space switching for face tracking in multi-colored lighting environments. In: Proceedings of IEEE FG, ISBN 0-7695-1602-5, pp. 249–254.

Taqa, A., Jalab, H., 2010. Increasing the reliability of skin detectors. Sci. Res. Essays 5 (17), 2480–2490.

Tsekeridou, S., Pitas, I. 1998. Facial feature extraction in frontal views using biometric analogies. In: Proceedings of EUSIPCO, pp. 315–318.

Wang, X., Zhang, X., Yao, J. 2011. Skin color detection under complex background. In: Proceedings of IEEE MEC, pp. 1985–1988.

Yang, G., Li, H., Zhang, L., Cao, Y. 2010. Research on a skin color detection algorithm based on self-adaptive skin color model. In: Proceedings of IEEE ICCIIS, pp. 266–270.

Yogarajah, P., Condell, J., Curran, K., McKevitt, P., Cheddad, A., 2012. A dynamic threshold approach for skin segmentation in color images. Int. J. Biometrics 4 (1), 38–55.

Zafarifar, B., Martiniere, A., de With, P. 2010. Improved skin segmentation for TV image enhancement, using color and texture features. In: Proceedings of IEEE ICCE, pp. 373–374.

Zhang, M.-J., Wang, W.-Q., Zheng, Q.-F., Gao, W. 2004. Skin-color detection based on adaptive thresholds. In: Proceedings of IEEE ICIG, ISBN 0-7695-2244-0, pp. 250–253.

Zhu, Q., Cheng, K.-T., Wu, C.-T., Wu, Y.-L. 2004. Adaptive learning of an accurate skin-color model. In: Proceedings of IEEE FG, ISBN 0-7695-2122-3, pp. 37–42.