Adaptive Digital Makeup

Abhinav Dhall, Gaurav Sharma, Rajen Bhatt,
and Ghulam Mohiuddin Khan

Samsung Delhi R&D, D5 Sec. 59, Noida
{abhinav.d,rajen.bhatt,mohish.khan}@samsung.com, grvsharma@gmail.com

Abstract. A gender and skin color ethnicity based automatic digital makeup system is presented. An automatic face makeup system which applies example based digital makeup based on skin ethnicity color and gender type. One major advantage of the system is that the makeup is based on the skin color and gender type, which is very necessary for an effective makeup. Another strong advantage is that it applies automatic makeup without requiring any user input.

1   Introduction

Good looks can be deceiving. Improving looks has always been of special interest to humans. Imagine a business video conferencing scenario; shabby looks can leave a bad impression on the viewer. The user is well served if the system can automatically apply digital makeup and enhance the caller's face. Different makeup suites different people as all faces are different and have varied makeup demands. For example a lip-stick shade which may suite a lady from some place may not exactly look good on lady from some other place having a different skin tone. Therefore the type of makeup must be adapted to the type of face. Not just female but male faces also require some touching up to remove small spots, pimples etc. and also perhaps improving the appearance of smokers' lips.

Digital makeup or artificial makeup has strong applications in various scenarios such as video conference, digital cameras, digital albums etc. In this work we pre-sent a new framework for ethnicity based skin tone and gender adaptive digital makeup for human faces. Before applying makeup straight forward as described in some earlier works our system access the type of makeup required based on gender and skin tone type based on skin ethnicity. The system first performs skin segmentation using fuzzy rules and locates the face using Active shape models and facial feature detector on probable face candidate areas. Before applying ASM scalable boosting based gender classification is performed on the face. Random sampling is performed on the detected face skin pixels and SVM based skin tone ethnicity classification is performed. Now with the available gender and skin tone type information and the facial feature the system applies case specific digital makeup. A database of k images in each category was taken and makeup-statistics in terms of HSV and alpha values were mined from them. This data was taken as reference and operations were performed for applying the makeup.

The paper is organized as follows. First we review the related work in Sec. 2. Then we describe our method in Sec. 3. Finally we give experimental results in Sec. 4.

2   Related Work

Robust automatic digital makeup systems are still in nascent stages. There is no such single system which can be termed as complete facial makeup system. In 2006 Microsoft was granted a patent [8] which outlines a general digital makeup system for video conferencing solutions. The system locates the face and the facial features and then applies various enhancement filters on the face such as contrast balancing, histogram equalization, eye dark circle removing. It also proposes blurring the background such that the caller/person looks clearer with respect to the background. In another approach blood haemoglobin and melanin values of a face were changed on the bases of a physical model generated from haemoglobin and melanin values [14]. Then in an-other work by Ojima et al. [12] "before" and "after" makeup examples are used. And foundation makeup procedure is used. Face retouching softwares have also been developed such as MyPerfectPicture [11] which manually ask the user to define the key facial points of the face and then select various parameters for touching. In their work Tommer et al [9]. used facial a facial attractiveness engine has been suggested which is trained with the help ph human raters. It extracts the distance between various facial feature points and maps that to a so called face space. This space is then searched for a larger predictive attractiveness rating and then points are edited using 2D warp. This basically alters the shape of the face and maps it to a better looking face as rated by the users earlier.

3   Proposed Framework

We present a system which first categorizes the face and then applies the digital makeup. Various steps are performed in order to extract and categorize the face. In order to identify the facial features, the first step is to perform fuzzy skin color segmentation, then haar feature based face, eyes and mouth detectors are used to extract rough areas. Now the skin pixels are classified into its ethnicity class using support vector machines. Next scalable boosting approach is used to categorize the face into male female. Active shape models are then used to extract lip and eyes from the earlier haar feature detected regions. Then the system matches the color moment of the face with the subset images from the database. The system starts with applying pre processing Gaussian smoothing filter. The whole process is depicted in Figure 1.

3.1   Fuzzy Learning Based Skin Segmentation

Skin segmentation is based on low complexity fuzzy decision tree constructed over B, G and R color space. Skin and non-skin data was generated from various
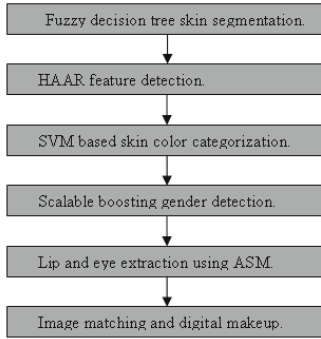
Fig. 1. Block diagram of the system

facial images constituting various skin tones, ages and gender. Sparse number of rules is generated by our skin segmentation system which is very fast. Fuzzy decision trees are powerful, top-down, hierarchical search methodology to extract easily interpretable classification rules [2] [3]. Fuzzy decision trees are composed of a set of internal nodes representing variables used in the solution of a classification problem, a set of branches representing fuzzy sets of corresponding node variables, and a set of leaf nodes representing the degree of certainty with which each class has been approximated. We have used our own implementation of fuzzy ID3 algorithm [2] [3] for learn-ing a fuzzy classifier on the training data. Fuzzy ID3 utilizes fuzzy classification entropy of a possibility distribution for decision tree generation. The overall skin non-skin detection rate comes out to be 94%. Fig. 2 shows fuzzy decision tree using fuzzy ID3 algorithm for the skin-non skin classification problem. Skin Binary Map Image (SBI) is generated which contains skin and non-skin information.
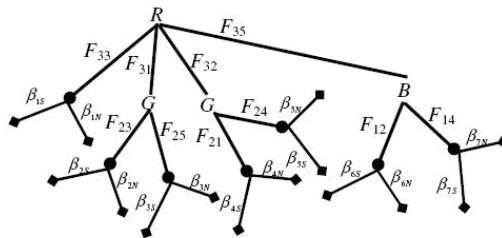

Fig. 2. Fuzzy decision tree

Next a connected component analysis is performed on the binary map ISBI and then face detection is applied on the skin segmented blobs.

Fig. 3. Output of the fuzzy based skin segmentation

3.2 Haar Based Facial Feature Extraction

Once face probable area is deduced using the skin detection, HAAR feature based Viola Jones face detector [15] is used to detect face on the reduce space. Then in the lower half of the face, similar detector for mouth is applied. This gives the approximate mouth area. Similarly, in the upper half of the face haar based eye detector is applied. The detected eye rectangles are referred to as EyeRectL and EyeRectR maps and the mouth area is referred to as LipRect map. This was done using the Intel OpenCV library [13]. Figure 4 depicts all the steps involved in facial feature extraction.
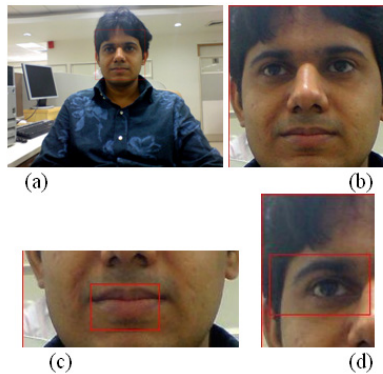

Fig. 4. Depicts the output sequence for face and facial feature detection steps. (a) Original image. (b) Face extracted using haar feature face detector. (c) Mouth region detected using haar feature based mouth detector. (d) The detected eye by the face detector.

3.3 Skin Ethnicity Classification

Next step in the method is classification into different ethnicity. Random skin pixels are picked from the face area and radial basis function kernel based SVM classification is performed. We perform a one versus one classification. This gives the skin tone class which provides the ethnicity information. Introduced by Vapnik [4] in 1995, Support vector machine is a set of related supervised learning

methods used for classification and regression. SVM constructs a hyper plane and maximizes the mar-gin between two set of vectors in n dimensions. Two parallel hyper planes are constructed and margin is maximized by pushing them towards the data vectors, the one which achieves maximum distance from data points of both the classes gets maximum margin and generally larger the margin then lower is the generalization error of the classifier.

Three classes based on ethnicity, European, Asian and African were defined. Training data was constructed for these individual classes and a radial bases function kernel was trained. Training and testing were used using the libsvm library [5].

3.4   Gender Classification

Since for applying efficient automatic digital makeup gender classification is necessary for male and female have different makeup requirements. Gender classification is done on the faces detected in the skin bound regions bearing ethnicity information by the haar feature detector. Each face is classified into male or female using Scalable boosting learned classifier models during training as defined by S. Baluja et Al. [1]. The author in this paper used the approach for image categorization; we modified it for gender detection. Classifier model chosen for gender classification of a face de-pends on its predicted ethnicity. Scalable boosting uses simple pixel comparison features for gender classification. Five types of pixel comparison operators are pi ¿ pj, pi within 5 units (out of 255) of pj, pi within 10 units (out of 255) of pj, pi within 25 units (out of 255) of pj, pi within 50 units (out of 255) of pj. There exist weak classifiers for each pair of different pixels in the normalized face image for each comparison operator. For each i and j pixel of the face image one of the above features is chosen which gives the best gender classification results for the dataset. The feature chosen acts as a weak classifier yielding binary results 1or 0 de-pending on whether the comparison is true or false respectively. The output corresponds to male if it is true and to female if false. For a 4848 pixel images, this means there are 2523042303 distinct weak classifiers. This yields extremely large number of classifiers to consider. Thus we need to minimize the number of features that need to be compared when given a face image in run time, while still achieving high identification rates. This is done using Adaboost algorithm [7]. The accuracy for European, Asian and African class were 93.3 %, 91.7 % and 90.2 % for 500 classifiers.

3.5   Facial Feature Extraction

Then the Active Shape Models by Cootes et al. [6] are used on these three maps generated from individually. This is done because ASM may not fit properly on the face and we want the exact eyes and lip. Active shape models (ASMs) are shape based statistical models of objects which iteratively fit to the object in a new scenario. In ASM the shapes are constrained by the point distribution statistical shape model to vary only in ways seen in a training set of labeled examples. The shape of an object is rep-resented by a set of points (controlled by

Fig. 5. The subset images of category specific faces that were used for training in gender detection taken from the face database. Row one and two are from European skin category, third is African and fourth and fifth are Asian.

the shape model). The ASM algorithm aims to match the model to a new image. The ASM library by Stephen Milborrow [10] was used for experimentation. Two ASM models were trained for eye and lip each. Region filling is done on the control points obtained from ASM.
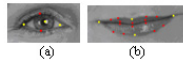

Fig. 6. (a) The output using ASM eye model and (b) The output using ASM mouth model

3.6   Reference Database

A database of example sample images was created. For each skin tone ethnicity type, Hue and Saturation color information of k example images each of male and female was stored. For example in the European skin tone category there is k male and k fe-male images which have been chosen as representing different type of European skin tones. Information on the type of makeup that should be applied on these images is stored along with the image skin color Hue Saturation color moment. The information contains after makeup skin tone Hue-Saturation values, lip stick color and alpha values in case of women, eye makeup color and its alpha values.

3.7   Digital Makeup

The first step in digital makeup is applying Gaussian smoothening and morphological dilation to the input image in order to remove small marks, pigmentations and moles. This RGB skin color image is then converted in HSV color space. The color moment CM (Skin Color) is computed over the new color space skin color image IHSV. This CM (SC) is then compared to the pre-stored color moment of skin color of the sub set images in the database. This sub set is derived on the basis of skin ethnicity type and gender type calculated earlier on the input face. For example: Asian male or African female etc. The image with minimum difference is chosen as the reference image and referenced to as IREF. The after

makeup Hue and Saturation values stored with IREF are used as the target values for IHSV. Hue and Saturation values are balanced on the basis of these pre-stored database values. Figure 7 shows these outputs.

Fig. 7. (a) Input skin sample (b) Output skin sample after Hue and Saturation balancing

Next step is lip shading. In case of female faces lipstick is applied to the lip area. Lipstick is applied as a rasterization operation (ROP). The target lip color LIPRGB and alpha values ALPHA are taken from the data present with the closest image matched earlier. The RGB MouthMap image is extracted using the MouthMap and control points computed earlier. For each pixel on lip a new color value is calculated as NewRGB = OLDRGB * (MAX-ALPHA) + LIPRGB * ALPHA. This preserves the texture of the lip. Figure 8, demonstrates the lip stick applying operation. In case of men separate color values are used which don't give the lipstick touch to the lips but makes their texture smooth which is in tandem with their skin color. This is especially useful in improving the looks of a smoker's lips as depicted in figure 8(b).

Fig. 8. (a) The input and output of a female Asian lip and (b) The input and output of an Asian male having smokers lip

4 Experimental Results

Figure 9 show the outputs for four different cases. Kindly note in Figure 9 (a) the lips color has changed and the texture has been maintained considerably. This is as Asian female, the skin tone now looks more red. The dark regions below the eyes have been considerably suppressed due to smoothening and skin color balancing. In Figure 9(b) the skin color has been improved and the lips have been applied with similar color so as to give a smoothened lip and remove the smoker's lips effect. In 9(c) the output has been taken in an office environment which shows the effectiveness of performing skin color segmentation and haar features as the initial steps. In Figure 9 (d) the subject is an African female, the skin color is now lighter, improved and smoothened and a light
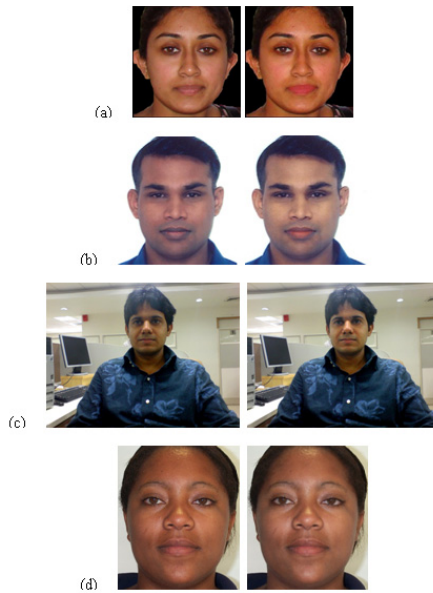
Fig. 9. (a) Input and output image of an Asian female after applying digital makeup (b) Input and Output image after applying digital makeup of a dark Asian man. (c) Input and output image after applying digital makeup. (d) Input and output image after applying digital makeup on an African female.

colored lip stick suiting the skin color tone has been applied by the system. The oily skin effect has also been removed.

5 Conclusion and Future Work

We presented a system for digital makeup. The system is based on two important factors (a) dependency of skin tone on ethnicity and (b) gender of the subject. The system customizes the makeup based on these two factors/parameters and retrieves makeup values via color matching from a pre defined data set. The reference image is choosen on the bases of the two parameters and color information. We employed ro-bust machine learning technique (a) fuzzy decision tree based skin color skin segmentation, (b) HAAR feature based face, eye and lip detection, (c) SVM with skin pixel color ethnicity categorization. (d) ASM based lip and eye extraction. Then we used fundamental image processing techniques for improving the appearance of the skin region and enhancing lips. The system is fast and we are currently exploring optimizations in order to implement it on embedded platform with real time performance. Eyes are relatively more complicated to be manipulated with fast and simple operations; we are currently pursuing this as a future work.

References

1. Baluja, S.: Automated image-orientation detection: a scalable boosting approach. Pattern Anal. Appl. 10(3), 247–263 (2007)
2. Bhatt, R.B., Gopal, M.: Erratum: "neuro-fuzzy decision trees". Int. J. Neural Syst. 16(4), 319 (2006)
3. Bhatt, R.B., Gopal, M.: Frct: fuzzy-rough classification trees. Pattern Anal. Appl. 11(1), 73–88 (2008)
4. Boser, B.E., Guyon, I., Vapnik, V.: A training algorithm for optimal margin classifiers. In: COLT, pp. 144–152 (1992)
5. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines (2001), `http://www.csie.ntu.edu.tw/~cjlin/libsvm`
6. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models-their training and application. Computer Vision and Image Understanding 61(1), 38–59 (1995)
7. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55(1), 119–139 (1997)
8. Microsoft Corporation. System and method for applying digital make-up in video conferencing us20060268101 (2006)
9. Leyvand, T., Cohen-Or, D., Dror, G., Lischinski, D.: Data-driven enhancement of facial attractiveness. ACM Trans. Graph. 27(3) (2008)
10. Milborrow, S., Nicolls, F.: Locating facial features with an extended active shape model. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 504–513. Springer, Heidelberg (2008)
11. MyPerfectPicture. Myperfectpicture (2009), `http://www.myperfectpicture.com/`
12. Ojima, N., Yoshida, K., Osanai, O., Akasaki, S.: Image synthesis of cosmetic applied skin based on optical properties of foundation layers. In: International Congress of Imaging Science, pp. 467–468 (1999)
13. Pisarevskyand, V., et al.: Opencv, the open computer vision library (2008), `http://mloss.org/software/view/68/`
14. Tsumura, N., Ojima, N., Sato, K., Shiraishi, M., Shimizu, H., Nabeshima, H., Akazaki, S., Hori, K., Miyake, Y.: Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. ACM Trans. Graph. 22(3), 770–779 (2003)
15. Viola, P.A., Jones, M.J.: Rapid object detection using a boosted cascade of simple features. In: CVPR (1), pp. 511–518. IEEE Computer Society, Los Alamitos (2001)