

Facial Cosmetics Database and Impact Analysis on Automatic Face Recognition

Marie-Lena Eckert ^{#1}, Neslihan Kose ^{*2}, Jean-Luc Dugelay ^{*3}

[#] *Computer Science Department, TU Muenchen
Boltzmannstr. 3, 85748 Garching b. Muenchen, Germany*

¹marie-lena.eckert@mytum.de

^{*} *Multimedia Communications Department, EURECOM
450 Route des Chappes, 06410 Biot, France*

²kose@eurecom.fr

³jld@eurecom.fr

Abstract—Facial cosmetics, also called makeup, may change the appearance of a face that we could perceive. In order to contribute to studies in image processing related to facial cosmetics, a database is built which contains multiple images per person with and without applied cosmetics. Annotations provide detailed information about the amount and location of applied makeup for each picture. Furthermore, a classification approach is presented. The classification approach takes the altering effect of the applied cosmetics and the application area into account. Since automatic face recognition evolved to an important topic over the last decades and is affected by facial cosmetics, preliminary tests are done to evaluate their impact on automatic face recognition. The face as a whole as well as its most significant makeup application areas that are skin, eyes and mouth are investigated separately.

I. INTRODUCTION

Facial cosmetics are a common part of the daily life in many cultures since Ancient Egypt [1] and change the perception of a face. The impact of facial cosmetics on the perception of a face is exemplified in Figure 1.

Research for evaluating the impact of facial cosmetics is limited, mainly because of the lack of a public database that contains multiple images with and without applied facial cosmetics for a significant panel of women. In [2], the authors develop a system detecting the most suitable makeup for a given human face. Unfortunately, they cannot provide public access to the database that they used for their study. Some databases (e.g. [3], [4]) provide images with applied facial cosmetics but there are no corresponding images without which makes them inappropriate for face recognition.

In facial cosmetics studies, it is inevitable to know what the ordinarily used facial cosmetics are, where they are applied and which effect they cause. So far, no exhaustive specification exists.

Facial cosmetics can be applied in different ways. To evaluate the impact of facial cosmetics, an objective categorization approach beyond subjective opinions and preferences is needed. As facial cosmetics change the perception of the face, they play an important role in human face recognition [5].

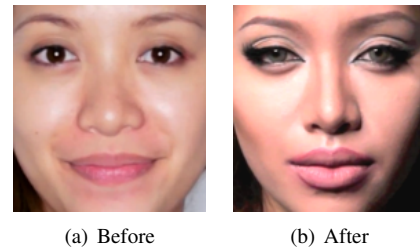


Fig. 1. Impact of Facial Cosmetics on the Perception of a Face. This figure is taken from [6].

Automatic face recognition evolved into an important topic over the last decades as its relevance in security applications, intelligent robotics and the entertainment industry continues to emerge. To improve the accuracy of face recognition systems, influencing factors such as facial pose and expression, illumination, aging, partial face occlusion and plastic surgery are investigated in several research studies (e.g. [7], [8]). Nevertheless, one rare investigation of the influencing factor facial cosmetics can be found in [9]. The authors state that this altering scheme comprises the accuracy of face recognition systems. They stress the need for a better understanding of the altering effects of facial cosmetics and the development of robust algorithms. Since the database used in their study is not publicly available, the obtained results are not reproducible. Neither the amount of applied facial cosmetics nor the cosmetics' application area are evaluated in detail.

The goal of this work is to contribute to studies related to facial cosmetics by providing a suitable database as well as an objective classification of facial cosmetics. The database contains multiple images with and without facial cosmetics for each person and provides detailed information about the application area and amount of applied cosmetics for each image. Especially studies related to automatic face recognition are affected by facial cosmetics. Therefore, in this study, preliminary tests concerning the impact of facial cosmetics on face recognition are conducted. In those tests, the popular and efficient face recognition technique Local Binary Patterns (LBP) [10] is used for evaluation to comply with the current state of face recognition. Since LBP is a texture-based tech-



Fig. 2. Reference Image and Makeup Series Images in their Original Form

nique and facial cosmetics have an influence on the texture, it is suitable to analyze the impact of facial cosmetics on face recognition with this technique.

This paper is organized as follows. In Section II, the construction and structure of the database built for evaluating facial cosmetics are presented. Section III discusses a specification and classification approach of applied facial cosmetics. Preliminary tests concerning the impact of facial cosmetics on automatic face recognition are presented in Section IV. The conclusion is conducted in Section V including future works.

II. FACIAL COSMETICS DATABASE

This section deals with the construction of a database which can be used for diverse studies related to facial cosmetics. It contains multiple images with and without applied facial cosmetics for each person. The persons belong mainly to the Caucasian race and are between 20 and 30 years old. Hispanics, Asian and African Americans are also sparsely represented. In Section IV-B, the database is used to evaluate the impact of facial cosmetics on a basic face recognition system based on LBP.

A. Specification of Images

The database contains frontal facial images with little expression. It provides one reference image and one or more progressive makeup series per person. The latter is a series of images, where at each succeeding image more makeup is applied. The series always starts with a no makeup image as visualized in Figure 2. All kinds of makeup varying from slight to heavy are represented. The reference image does not belong to any makeup series and is taken in an other setting, resulting in variations of illumination, skin and hair conditions. As it can be observed from Figure 2, the reference image shows a different tone of skin, less dark circles under the eyes and a change in the eyebrows's shape.

Many applications require preprocessed images. Therefore, each image is available in its original and preprocessed version in the database. Preprocessing includes rotation and cropping as shown in Figure 3. Rotation and cropping combined ensure that the eyes in each image are exactly at the same position. The face area is extracted from the image in its original form through cropping, whereby background, hair and accessories are cut off.

The facial subareas, eyes and mouth, are altered the most by facial cosmetics. To evaluate them separately, eyes and

mouth are extracted and stored in separate images. The mouth images are rotated as well to ensure that the mouth corners are exactly at the same position in each image.

B. Acquisition of Images

It is important to collect images that are taken under similar conditions as images in real world scenarios (e.g. surveillance camera, snapshots). To achieve this, makeup tutorial videos are used as source. Makeup tutorials are well suitable for obtaining progressive makeup series. Most used videos are found on YouTube, some on websites of facial cosmetic companies. More information is available at <http://fcd.eurecom.fr/>. YouTube is a convenient source, since many persons inherit an own channel and therefore multiple videos of the same person can be found quickly. The number of online videos increases daily.

To ensure different conditions for reference and test images as described in the previous Section II-A, at least two videos are required for each person.

C. Structure of the Database

Each image in the database contains detailed information in its filename of form XXXX-N-ABCDEFGH. XXXX specifies the identification number of the person and N the identification number of the video related to the person. ABCDEFGH refers to an 8-digit makeup code which is described in Table I. The makeup code specifies the quantity of applied makeup at eight different locations (see Figure 4).

The database currently contains 389 images, whereas 50 are reference images and 339 belong to makeup series. 109 of the makeup serie images are classified into the makeup category 'no', 147 into 'slight', 54 into 'intermediate' and 79 into 'heavy' which are described later in Section III-B2. Each image is available in its original (Figure 2) and cropped form (Figure 3). The sizes of the face, eyes and mouth images are 150×140 pixel, 60×140 pixel and 40×70 pixel. For each original form of an image, the coordinates of the eye centers and mouth corners are manually annotated and provided in a textfile. If necessary, users can rotate and crop the images such as it suits their individual application best.

III. FACIAL COSMETICS

In this section, commonly used facial cosmetics, also called makeup, are specified. The specification results from the observation of numerous makeup tutorials. To evaluate facial cosmetics in detail, they must be classified into categories. In



Fig. 3. Reference Image and Makeup Series Images in their Rotated and Cropped Form

this paper, a first classification approach based on the effect and exact location of applied cosmetics is proposed.

A. Specification of Facial Cosmetics

Facial cosmetics are mainly applied on female faces to increase their attractiveness. Attractiveness is affected by youthfulness, symmetry [11] and averageness [12]. Youthfulness is composed of healthy, lively and also childlike characteristics [13] (e.g. soft skin, large eyes and small nose). All three attributes can be influenced by applying facial cosmetics. Their positive effects on attractiveness is shown in [14].

The application of facial cosmetics can be split up into eight main steps. Those steps are defined based on the observation of numerous makeup tutorials used for constructing the database in Section II. The tutorials are provided by ordinary women as well as professional makeup artists. They present different ways to apply makeup ranging from every-day to more exclusive makeup. Table I summarizes the eight main makeup steps including associated used products and their impacts.

Those makeup steps form an 8-digit makeup code. Each makeup step can be applied with a slight or heavy amount, which is indicated in the makeup code by the digit 1 or 2. No application of makeup is indicated by the digit 0.

B. Classification of Applied Facial Cosmetics

The application of facial cosmetics is classified in two groups which are;

1) *Classification by facial subareas*: The face is divided into its most significant subareas where makeup can be applied. Those are skin, eyes and mouth. The makeup steps foundation and cheek belong to skin, whereas lips belong to mouth. Most steps though belong to the eye area, namely brow, shadow, upper line, mascara and lower line.

Each subarea is considered separately. This gives insight concerning its sensitiveness to facial cosmetics and its contribution to the overall results. For each subarea, the applied cosmetics are classified by their altering effect as described in the following.

TABLE I
EIGHT MAKEUP STEPS

Step	Products	Effect
Foundation (skin)	Foundation, concealer, powder.	Coverage of skin flaws and natural shadows; change of skin color, face and nose shape; reduction of shininess.
Cheek	Bronzer, blush.	Emphasis and change of perceived position of cheekbones; lively and youthful appearance.
Brow (eyebrow)	Eyebrow pencil.	Accentuation, shaping and change of perceived position of eyebrows; increase of contrast.
Shadow (upper eyelid, eyebrow bone)	Eyeshadow.	Emphasis of lid crease; increase of depth; change of perceived size and color of eyes.
Upper line (upper eyelid)	Eyeliner, kajal.	Enhancement of contours; change of perceived size of eyes; thickening of lashline.
Mascara (lashes)	Mascara, false lashes.	Increase of density, swing and length of lashes; illusion of bigger eyes.
Lower line (lower eyelid)	Eyeliner, kajal.	Enhancement of contours; change of perceived size of eyes.
Lips	Lip liner, lipstick, lip gloss.	Enhancement of contours; change of perceived size, color and contrast; shininess.

2) *Classification by impact on facial appearance*: The effect of facial cosmetics depends on how much and in which way the perceived size and color of facial features are changed. Three main categories (see table II) were established after examining numerous makeup tutorials used in section II. Each single makeup step with its applied amount of makeup leads to one of the makeup categories. The rules in figure 4 describe

TABLE II
CATEGORIES OF THE IMPACT OF FACIAL COSMETICS

Category	Description
No	No makeup applied, no effect
Slight	Slight changes in color and perceived shape
Intermediate	Strong enhancement of contours and characteristic features
Heavy	Estrangement from personal characteristics, mask-like and artificial

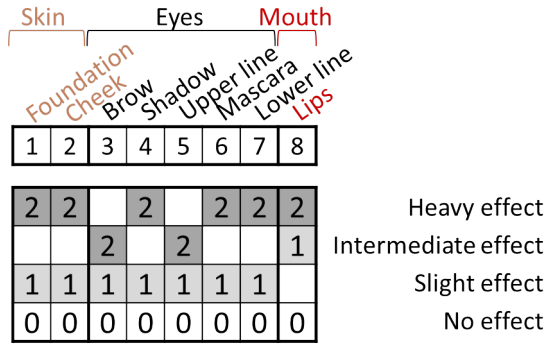


Fig. 4. Makeup Classification Rules Depending on Steps

the classification of the applied cosmetics per facial subarea. This allows the automated deduction of the categories based on the makeup code. The highest makeup category of all steps belonging to a certain subarea defines the makeup category of this given subarea. The classification of the eye area is performed as follows: If shadow, mascara or lower line are applied heavily, it leads to a classification of the eye area into the category 'heavy', because those steps have a mask-alike or artificial effect. Heavy application of brow or upper line indicate the category 'intermediate', since the shape of eyes is enhanced. If any makeup step in the eye area is applied slightly, it leads to a classification into the category 'slight'. If the foundation or cheek makeup step is applied heavily, the skin area is classified into the category 'heavy'. Foundation uniforms, evens and flats the facial shape. Natural shadows and marks are covered. Properly placed highlights change the apparent shape and size of facial features, e.g. nose. The cheek makeup step changes the perceived position, color and size of cheekbones. This leads to a distinction from the personal characteristics and therefore has an artificial or mask-alike effect. There is no skin makeup category 'intermediate'. The slight application of lipstick, lip liner or lip gloss enhances the contours and contrast of the mouth area and therefore leads to a classification into the category 'intermediate'. Heavy application of lip makeup however causes changes in size and shape and therefore leads to the category 'heavy'. There is no mouth makeup category 'slight'.

The makeup category of the face as a whole is equal to the highest category of the subareas. However, there is one special case, because the subareas' categories contribute unequally to the face makeup category. Mouth makeup has little impact on the overall perceived amount of makeup. Intermediate mouth makeup leads only to a makeup category 'slight' for the face as a whole.

IV. EVALUATING THE IMPACT OF FACIAL COSMETICS ON AUTOMATIC FACE RECOGNITION

In this section, the impact of facial cosmetics on automatic face recognition is evaluated. First, the test setup is described while the results are presented in the next part.

A. Test Setup

The makeup serie images are presented to the face recognition system in order to be classified with the identity of the person shown in the image. Each identity is defined by a reference image. The reference images in security or entertainment applications are usually taken in a different setting than the test images which is also ensured in the database. No makeup series is then recognized better than others due to same illumination, skin and hair conditions as in the reference image.

The accuracy of the face recognition system is here measured by the identification rate (IDR). The IDR indicates the percentage of correctly identified persons. The face recognition is conducted by applying Local Binary Patterns (LBP) with a block size of 10×10 pixels and a neighborhood with $(P, R) = (8, 2)$. P denotes the number of considered sampling points within radius R . As discussed in Section III-B, the application of facial cosmetics can be categorized by application area and impact of makeup. The face as a whole and skin, eye and mouth subareas are examined separately. For each of them, makeup images from the category 'no', 'slight', 'intermediate' and 'heavy' are evaluated. To evaluate skin makeup, only images where solely skin makeup has been applied are considered. Skin makeup is usually applied before other steps. The database currently contains 55 images with solely skin makeup. Since 53 of them are classified into the makeup category 'slight', the number of heavy skin makeup images is not large enough in current database to represent the category 'heavy' correctly. To evaluate it as well, the size of the database would have been more important. Hence, only slight skin makeup is evaluated here. For evaluating eye and mouth makeup, the cropped images of the eye and mouth area are tested.

B. Results

In this section, the identification rates resulting from the tests are presented. First, reference images with no makeup are used. In the second part, the reference images are replaced by a subset of intermediate makeup images.

1) *Test with No Makeup Images as References:* The database currently contains 50 reference images with no makeup. Each of them are taken in an other setting than the makeup serie images.

As shown in Figure 5(a) the ability to identify a person's face decreases with increasing amount of makeup. The IDRs for the categories 'no', 'slight' and 'heavy' decrease almost linearly, while the one for the makeup category 'intermediate' shows a small increase (i.e. does not decrease linearly) but is still lower than the one for the 'slight' category in the graph. The statement from [9], that facial cosmetics degrade the accuracy of face recognition systems, is hereby confirmed. The results are in line with the expectations. Slight makeup test images achieve a little lower IDR than no makeup images. Slight makeup has little effect on human perception and maintains a natural look as defined in table II. Intermediate makeup in contrast is clearly visible and therefore provides a lower IDR.

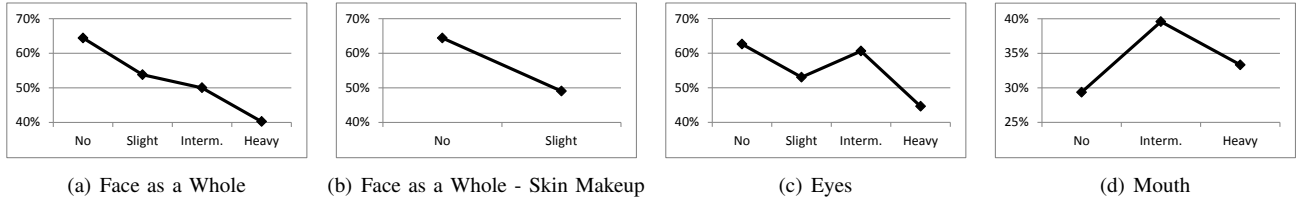


Fig. 5. IDR for Reference Images with No Makeup

Heavy makeup estranges persons from their characteristic appearance and is perceived as mask-alike. Hence, the IDR for faces with heavy makeup reach only two third of the IDR for faces with no makeup. To gain detailed insights into the results, the face is divided into its three most significant makeup application areas, namely skin, eyes and mouth. As shown in Figure 5(b), skin makeup decreases the face recognition ability. As explained in Section III-B2, skin makeup uniforms the face and changes the perceived position, color, shape and size of facial features. This leads to a distinction from the personal characteristics and therefore prevents face recognition.

Figure 5(c) shows the IDR for eye recognition. It follows the same scheme as Figure 5(a) in terms that the IDRs for the categories 'no', 'slight' and 'heavy' decrease almost linearly, while the one for the makeup category 'intermediate' results in a local peak. The peak is higher than in Figure 5(a). Compared to the categories 'slight' and 'heavy', intermediate makeup images provide the highest IDR, which is almost as high as the IDR for images where no makeup is applied. Apparently, intermediate makeup in the eye area has no negative impact on face recognition. It enhances characteristic features and contours, which leads to better distinguishable eye shapes. Heavy eye makeup however influences eye recognition adversely. The eye area is dominated by large applications of unnaturally colored makeup which have a estranging and unifying effect. Shape, size and contours of the eyes get less distinctive. Slight eye makeup has only a slight negative impact on eye recognition. Some makeup steps as brow and upper line enhance the contours of facial characteristics, while others as eye shadow, mascara and lower line do not.

The IDRs for the mouth area are shown in Figure 5(d). The overall IDR curve is lower than for all other facial parts. This leads to the conclusion, that the mouth area contributes little to face recognition. The low IDRs are caused by facial expression which affect predominantly the mouth area. Nevertheless, it is shown that application of facial cosmetics in the mouth area leads to better identification results. This is explained by its contour enhancing and contrast increasing effect. Again, intermediate makeup leads to the best results. Heavy makeup also includes changes of size which leads to a lower IDR compared to intermediate makeup. There is no makeup category 'slight' heavy for the mouth area as explained in Section III-B2.

The IDR curve of the face as a whole is most similar to the one corresponding to the eye area, although the IDR curve for the eye area is overall little higher. That implies that the

eye area is the most important facial subarea when it comes to face recognition, which is also stated in [7]. The eyes are the most distinctive facial part. The reduced peak at the face IDR curve is explained as follows. Intermediate makeup enhances characteristic features and contours, which leads to a better recognition of eyes and mouth. Skin makeup though decreases the recognition ability because of its unifying effect and changes in perceived size and shape of facial features. Slight and heavy makeup have a negative impact on face, skin and eye recognition. Images of the makeup category 'no' result generally in the best recognition rates. Mouth makeup steps are out of line because they have a positive effect on recognition. However, the mouth area has little impact on the face recognition.

The aim of evaluating the impact of facial cosmetics is to improve face recognition in reality. The previous results state that intermediate makeup plays an extraordinary role in face recognition. The question arises if intermediate makeup in reference images could have a positive effect on face recognition.

2) Test with Intermediate Makeup Images as References:

In this section, images with intermediate makeup are used as reference images. The database provides makeup series for 23 persons containing intermediate makeup images. The remaining makeup series provide only images belonging to the makeup categories 'no', 'slight' or 'heavy'. Therefore, the evaluation is conducted with 23 reference images originating from makeup series. They consequently do not underlie other conditions than the test images but rather possess the same camera settings, skin and illumination conditions. Since the reference images are from the same session, it is less realistic compared to the test with no makeup images as references case and this is why we obtain almost perfect result (100%) in Figure 6 for intermediate images. 175 test images are available for the specified 23 persons. 51 of them are classified into the makeup category 'no', 54 into 'slight', 30 into 'intermediate' and 40 into 'heavy'.

The IDR curve for face recognition with intermediate makeup reference images is higher compared to the one resulting from no makeup in the reference images. This is caused by the feature enhancing effect in the reference images due to intermediate makeup and the similarity of the reference images to the test images based on the same makeup series as origin. Here, increasing makeup does not lead to decreasing IDRs. The IDRs for the categories 'no', 'slight' and 'heavy' are almost linearly increasing, while the one for the makeup

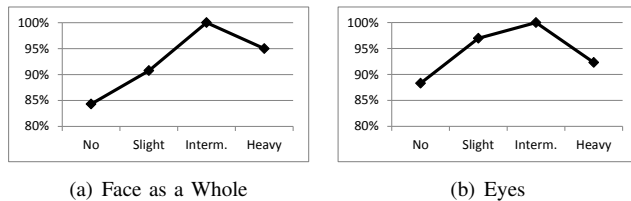


Fig. 6. IDR for Reference Images with Intermediate Makeup

category 'intermediate' results again in a high peak as shown in Figure 6(a). Facial images with intermediate makeup have the highest IDR, namely 100%. Images where no makeup is applied result in the lowest IDR. The IDRs for the eye area show a similar structure, which reinforces the statement, that the eye area dominates the face recognition. Eyes with intermediate makeup are recognized the best while eyes with no makeup provide the lowest IDR.

V. CONCLUSION

Facial cosmetics are commonly and daily used in many cultures and may alter the perception of facial features.

To propose a study on the impact of facial cosmetics on face recognition, a database is built and a categorization of the most common facial cosmetics and an objective classification approach are developed.

This database is used for preliminary evaluation of the impact of facial cosmetics on face recognition. If the reference images do not include any make-up, facial cosmetics do have a negative impact on automatic face recognition. Surprisingly, the application of facial cosmetics in the reference images can help to obtain better identification rates. Intermediate make-up increases both interclass and intraclass variation but the increase is higher for interclass variation therefore the impact is positive. Facial cosmetics create intraclass variations. If they are only applied at one of the image sets, they obviously lead to a decrease of the accuracy of face recognition systems. In [15], it is underlined that intraclass variations between reference and test images greatly affect the accuracy of face recognition systems. In both cases, the eye subarea plays a major role in face recognition while the mouth area is not pivotal.

Both the makeup categorization approach and the established database are suitable for face recognition and related application areas, e.g. beauty evaluation, computer-suggested makeup or automated makeup detection. Based on the provided images and detailed information about the application area and amount of applied makeup, algorithms can be trained to recognize different kinds of makeup. If makeup is detected, facial subareas can be weighted differently to improve face recognition results as similarly performed in [7] for facial occlusions. To obtain more stable and detailed results, the size of the database can be increased. The definition of the makeup category 'slight' requires further refinement.

REFERENCES

[1] M. Angeloglou, *A History of Make-up*. Macmillan, 1970. [Online]. Available: <http://books.google.fr/books?id=4tzfAAAAAAJ>

[2] K. Scherbaum, T. Ritschel, M. Hullin, T. Thormählen, V. Blanz, and H.-P. Seidel, "Computer-suggested facial makeup," *Comp. Graph. Forum (Proc. Eurographics 2011)*, vol. 30, no. 2, 2011.

[3] "Face recognition homepage." [Online]. Available: <http://www.face-rec.org/databases>

[4] "Face-place." [Online]. Available: <http://tarlab.cnbc.cmu.edu/face-place>

[5] S. Ueda and T. Koyama, "Influence of make-up on facial recognition," *Perception*, vol. 39, no. 2, pp. 260–4, 2010. [Online]. Available: <http://www.biomedsearch.com/nih/Influence-make-up-facial-recognition/20402247.html>

[6] [Online]. Available: <http://www.youtube.com/watch?v=rNQ2oDEyVNw>

[7] R. Min, A. Hadid, and J.-L. Dugelay, "Improving the recognition of faces occluded by facial accessories," in *FG 2011, 9th IEEE Conference on Automatic Face and Gesture Recognition, March 21-25, 2011, Santa Barbara, CA, USA*, Santa Barbara, UNITED STATES, 03 2011. [Online]. Available: <http://www.eurecom.fr/publication/3314>

[8] N. Erdogmus, N. Kose, and J.-L. Dugelay, "Impact analysis of nose alterations on 2D and 3D face recognition," in *MMSP 2012, IEEE International Workshop on Multimedia Signal Processing, 17-19 September, 2012, Banff, Canada*, Banff, CANADA, 09 2012. [Online]. Available: <http://www.eurecom.fr/publication/3760>

[9] A. Dantcheva, C. Chen, and A. Rosant, "Can Facial Cosmetics Affect the Matching Accuracy of Face Recognition Systems," in *BTAS12, IEEE Fifth International Conference on Biometrics: Theory, Applications and System*, Washington, USA, 2012.

[10] T. Ahonen, A. Hadid, and M. Pietikinen, "Face description with local binary patterns: Application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[11] D. I. Perrett, D. M. Burt, I. S. P. Voak, K. J. Lee, D. A. Rowland, and R. Edwards, "Symmetry and human facial attractiveness," *Evolution and Human Behavior*, vol. 20, pp. 295–307, 1999.

[12] G. Rhodes, A. Sumich, and G. Byatt, "Are Average Facial Configurations Attractive Only Because of Their Symmetry?" *Psychological Science*, vol. 10, pp. 52–58, 1999.

[13] L. Z. McArthur and K. Apatow, "Impressions of baby-faced adults," *Social Cognition*, vol. 2, pp. 315–342, 1984.

[14] A. Dantcheva and J.-L. Dugelay, "Female facial aesthetics based on soft biometrics and photo-quality," in *ICME 2011, IEEE International Conference for Multimedia and Expo, July 11-15, 2011, Barcelona, Spain*, Barcelona, SPAIN, 07 2011. [Online]. Available: <http://www.eurecom.fr/publication/3368>

[15] H. Sellahewa and S. Jassim, "Image-quality-based adaptive face recognition," *Instrumentation and Measurement, IEEE Transactions on*, vol. 59, no. 4, pp. 805–813, 2010.