Investigating the Periocular-Based Face Recognition Across Gender Transformation

Gayathri Mahalingam, *Member, IEEE*, Karl Ricanek, Jr., *Senior Member, IEEE*, and A. Midori Albert

***Abstract*— This paper introduces a novel face recognition problem domain: the medically altered face for gender transformation. A data set of >1.2 million face images was constructed from wild videos obtained from You Tube of 38 subjects undergoing hormone replacement therapy (HRT) for gender transformation over a period of several months to three years. The HRT achieves gender transformation by severely altering the balance of sex hormones, which causes changes in the physical appearance of the face and body. This paper explores that the impact of face changes due to hormone manipulation and its ability to disguise the face and hence, its ability to effect match rates. Face disguise is achieved organically as hormone manipulation causes pathological changes to the body resulting in a modification of face appearance. This paper analyzes and evaluates face components versus full face algorithms in an attempt to identify regions of the face that are resilient to the HRT process. The experiments reveal that periocular face components using simple texture-based face matchers, local binary patterns, histogram of gradients, and patch-based local binary patterns out performs matching against the full face. Furthermore, the experiments reveal that a fusion of the periocular using one of the simple texture-based approaches (patched-based local binary patterns) out performs two Commercial Off The Shelf Systems full face systems: 1) PittPatt SDK and 2) Cognetic FaceVACs v8.5. The evaluated periocular-fused patch-based face matcher outperforms PittPatt SDK v5.2.2 by 76.83% and Cognetic FaceVACS v8.5 by 56.23% for rank-1 accuracy.**

***Index Terms*— Periocular recognition, face recognition, medical alteration, plastic surgery, disguise, gender transformation, hormone replacement therapy, transgender.**

I. Introduction

THE face is the most expressive part of the human body. From the face one can determine a number of attributes or characteristics of a person. For example the face conveys identity, lineage, sex, race, ethnicity, mood, feelings, etc. The face is powerfully expressive, and hence, it can be challenging to extract information in a robust and efficient way. Over the years the face has provided unique challenges to the biometrics community. These challenges are often described as A-PIE which stands for aging, pose, illumination, and expression. There has been a wealth of research conducted over the last 30 years to resolve these challenges. In this work, we introduce a very unique challenge, which is recognition under gender transformation due to hormone replacement therapy.

Manuscript received March 15, 2014; revised July 14, 2014, September 7, 2014, and September 26, 2014; accepted September 26, 2014. Date of publication October 2, 2014; date of current version November 12, 2014. This work was supported in part by the FBI Biometric Center of Excellence and in part by the Army Research Laboratory in conjunction with the Center for Advanced Studies in Identity Science. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Aly A. Farag.

The authors are with the Face Aging Group Research Laboratory, Department of Computer Science, Institute for Interdisciplinary Studies in Identity Sciences, University of North Carolina at Wilmington, Wilmington, NC 28403 USA (e-mail: mahalingamg@uncw.edu; ricanekkg@uncw.edu; albertm@uncw.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2014.2361479

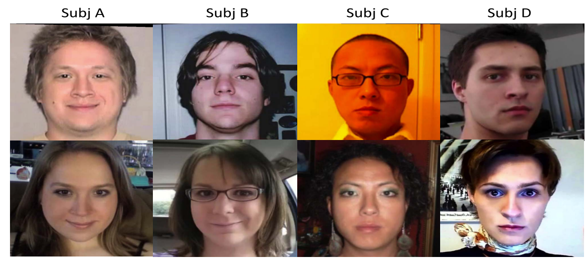
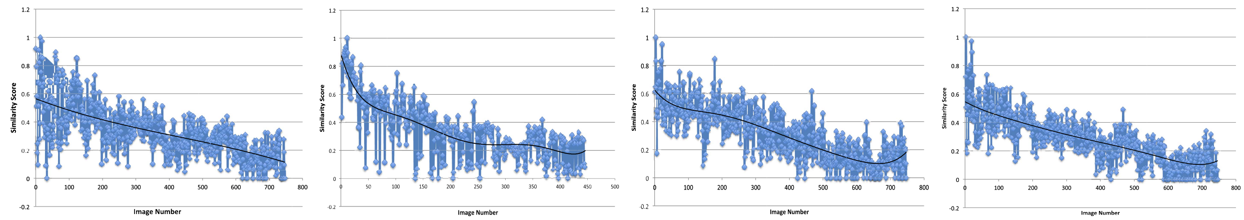


Fig. 1. Sample images of subjects showing their facial appearance before and after gender transformation (HRT). Row 1 contains pre-HRT images while row 2 contains post-HRT examples. Subjects have undergone at least 1-year of hormone replacement therapy (HRT).

Face recognition performance from the Good, the Bad, and the Ugly problem [9] indicates that more work is needed for face recognition to address the non-ideal, or PIE, scenarios. Face verification research in the literature [10]–[12] has shown that the verification performance decreases with an increase in the age span between the match pairs, and that the main–anecdotally derived–causes are due to large facial shape and texture changes. Coincidentally, changes in facial shape and texture is also observed for subjects who undergo gender transformation through hormone replacement therapy (HRT).

HRT based gender transformation affects face fat distribution thus resulting in changes to face shape and texture. Reduced fat distribution can allow for fine wrinkles and lines to become prominent whereas an increase in fat distribution stretches the dermis thereby reducing the prominence of wrinkles and lines. For example, female to male gender transformation causes the face to become more angular (masculine) by reducing the fat distribution in the face. (The reduction in fat cells is caused by a shrinkage of the cells not an eradication of the fat cells.) In addition, the skin is either thinned (male to female) or thickened (female to male), thus introducing texture variations to the face region. It has also been shown that the factors influencing skin aging process are significantly improved as a result of HRT [13]. The result of HRT gender transformation is an increase in the *within-class* variation and reduction in *between-class* variation.

1556-6013 © 2014 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.



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

Fig. 2. Similarity scores (from PittPatt v5.2.2 SDK) indicating performance degradation as the subject goes through hormone alteration for gender transformation. X-axis shows the n^{th} image matched to the original (pre-HRT) and Y-axis is the normalized similarity score. (A) Subject A. (b) Subject B. (c) Subject C. (d) Subject D.

It can be argued that gender transformation can be considered a variant of face disguise, however, disguise falls under the broader category of biometric obfuscation [17], which refers to the deliberate alteration of the face for the purpose of masking one's identity. Transgender persons do not undergo HRT for the purpose of biometric obfuscation, however, the question that still remains is, "*will someone use HRT for the purpose of masking or creating a new identity?*". The juror is out, but researchers will now have data with which they can use to develop face matching systems for this disguise variant. One has to only review Figure 2 to ascertain the problem presented by gender transformation. The figure clearly illustrates reduction in similarity score over the course of time (HRT). This graphic was obtained by comparing the first image in the sequence to the remaining images of the person as they under went HRT for a period of several months. The similarity scores are obtained using the Pittsburgh-Pattern Face Recognition SDK v5.2.2, PittPatt SDK, [34]. It is clearly evident that the similarity between the images of the same subject decrease indicating an increase in the *within-class* variation. Hence, a recognition/verification system should take into account the variations caused by gender transformations or gender invariant features of the face in order to provide better recognition performance.

The problem of face recognition across gender transformation is unique and different from other medical alterations such as plastic surgery. Although, it seems that both HRT and plastic surgery introduces texture and shape variations to the face, the facial changes that are introduced by these two are in contrast with each other owing to their significant difference.

In plastic surgery, the skin textural changes are attributed and can include either, or in some cases both, surgical removal of facial scars, acnes, or skin resurfacing. The visible signs of aging such as sagging in the middle of the face, deep creases below the lower eyelids, deep creases along the nose extending to the corner of the mouth, etc. are removed surgically through face lift and brow lift procedures and more recently with a combination of fat injections to the fat padding that exists below the muscle. The recent trend in enhancing (enlarging) fat padding produces is more natural face alteration. In HRT, the skin textural changes are attributed to either loss or gain of fat distribution in the face. The reduction of fat distribution results in a more angular, masculine, face near the cheek and jaw line and also allowing for fine wrinkles and lines to become more prominent, thus changing the texture patterns of the face. Also, the thinning/thickening of the skin and changes in the level of hormones in the aging process influence the skin textural changes.

Facial shape changes due to plastic surgery can be introduced to the face components such as forehead, eyelid, nose, chin, lip and ears. These face components are either reshaped or restructured through surgical procedures causing an overall shape change to the face and local texture changes around the face component.

It is to be noted that HRT introduces wrinkles and lines in the skin in contrast to the plastic surgery, which tries to, removes them. Also, the shape changes due to HRT are mainly due to changes in the fat distribution, while in plastic surgery it is due to surgical procedures. Individuals undergoing HRT may also undergo surgical procedures for changes in their facial appearance, thus rendering the problem of face recognition across HRT to be unique and much difficult than individuals undergoing surgical procedures only. Aggarwal *et al.* [14], from their experimental evaluation have shown that the recognition accuracies from each individual face components were much lower when compared with the full face. This clearly indicates the significant shape changes of the face components before and after the surgery. Also, Liu *et al.* [38] have shown that the best performance is achieved in identifying an eyelid surgery indicating the fact that the eye region undergoes significant changes in plastic surgery. This is in contrast with the primary assumption in our work that the periocular region is invariant to changes due to HRT.

Despite the high adoption rate of face recognition for a number of security and consumer applications, face recognition under gender transformation has not been researched by the community. This paper introduces the problem of face recognition under the presence of this new covariate, *gender transformation*. Gender transformation occurs by down selecting the natural sex hormone of a person in replacement for its opposite. This is known medically as hormone replacement therapy; however, more broadly this can be described as hormone alteration or medical alteration. A face component-based recognition framework is proposed in a hope to improve the recognition performance. The face component framework has been selected to explore regions of the face that may be resilient to the face alterations caused by this medical procedure. The authors hypothesize that components will out perform full face as these components may undergo significantly less changes than the sum of their parts, full face.

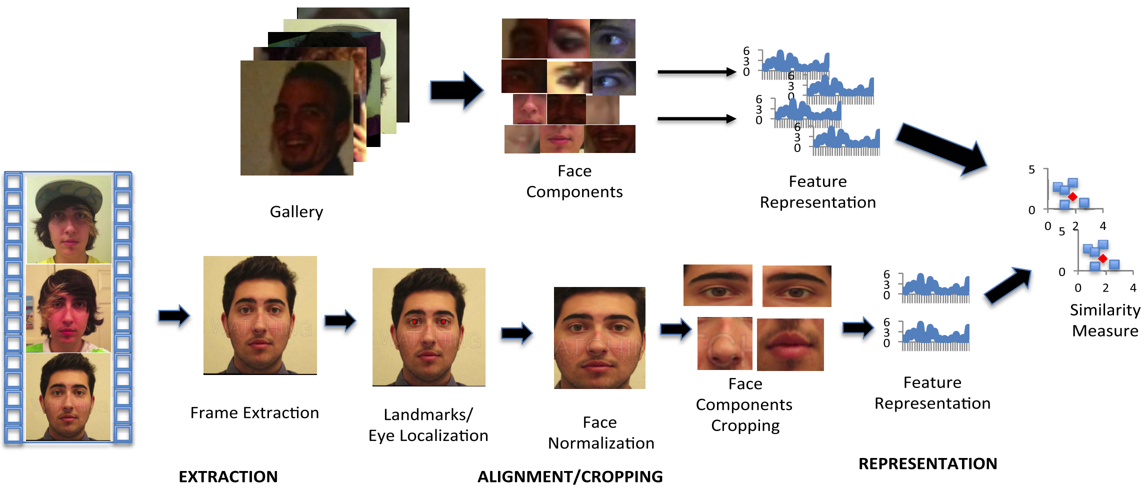


Fig. 3. Illustration of the component-based recognition framework used for this work.

Further, the authors believe that the periocular region, which is anchored by the eye orbit and brow ridge, absent of fat paddings (exception, fat pads around the eyeball), and where the epidermis, skin, is sufficiently thin for both male and female, will be sufficiently stable as compared to the nose, mouth, or full face. Figure 3 shows the framework proposed in this work.

It is to be noted that the periocular region as defined in this paper refers to the region that includes the eyes, the eyebrows, and the periorbital region (soft tissue region contained around the eye-orbit). The framework is evaluated with a transgender dataset (organized by the authors) consisting of images of 38 subjects taken under unconstrained environment across time and gender transformation. This work extends the conference version [35] by dramatically increasing the number of subjects and investigating the full spectrum of the major face components. The source of the images comes from YouTube©videos. The images are extracted from frames of video sequences and only those frames that present a frontal face, where the face is the dominant object in the frame. Figure 1 show sample images from the dataset. It is unknown whether these subjects have undergone facial surgery in addition to HRT. In addition to these variations, the images include other covariates such as pose, expressions, illumination, aging, makeup, etc. that are not expressly handled in this research work. Unlike other works [15], [16], [37], this work aims at investigating the authors hypothesis' that face components will outperform full face, and more particular, that the periocular region will outperform all face components, and finally, that the periocular using simple texture features will dramatically outperform the selected Commercial Off The Shelf Systems (COTS).

A. Contribution

The key contributions of this paper are as follows.

- To extend the work performed in the conference paper [35] and hence, the hypothesis. In [35] the authors theorized that eyes of an HRT transgender would demonstrate less changes than the full face when measured through the lens of texture-based face matchers. This work examined the major face components against each as well as against the full face to determine areas of resilience. Further, the number of subjects were more than doubled.
- A challenging dataset that includes more than 1.2 million face images of 38 subjects (an extended version of the dataset used in [35] that spanned several months to three years. The dataset by itself includes complexities from the aspects of A-PIE, occlusions, weight loss/gain, and more important the gender variations.
- The detailed analysis is performed from the perspective of both medical (anatomy) studies and through extensive experiments. The medical study analyzes the facial differences between male and female, the facial changes due to hormonal changes, and the feasibility of the periocular region as an HRT invariant feature.
- An extensive set of experiments under different scenarios using three feature descriptors on the various face components and the full face are performed. Additional experiments were performed to understand the effect of pose on the performance of the matchers. The statistical significance of the dataset and the recognition results is provided through appropriate tests. In this work, we also provide comparison on the recognition performance of the periocular region with the commercial face recognition systems PittPatt v5.2.2 [34] and Cognetic FaceVACS v8.5.

The remainder of this paper is organized as follows. In Section II, we provide a detailed anatomical analysis on the feasibility of the periocular region as a useful biometric trait for recognizing individuals undergoing facial changes due to HRT. In Section III we detail our protocol in the extraction of frames and our approach for aligning and extracting features from the periocular region and other face parts. In Section IV discusses in detail the face-based recognition experiments (both closed set and open set recognition) using the periocular region representations. Results from experiments on

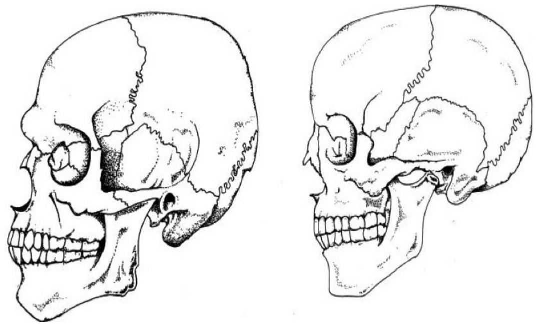


Fig. 4. Male (left) and female (right) skulls showing the difference in the glabellar region.

face images with orientation changes and occlusions are also detailed and discussed. Statistical analysis of the experimental results are also discussed in this section. Finally, in Section V, the results are summarized and the potential use of the periocular region in practical applications is discussed.

II. Anthropometry Analysis on Feasibility of Periocular Region

The difference between females and males in terms of body size and morphology is traditionally known as sexual dimorphism, and is pertinent to both soft tissues (skin and muscles) as well as hard tissues (skeletal elements). Here, the aim is to provide a description of sexual dimorphism in the adult facial skeleton, followed by a consideration of the question of the effects of hormonal changes on these bony sex differences. Specifically addressed are the changes, if any that occur during the transsexual change process when a female shifts to male and when a male shifts to female through the use of hormone supplementation. Finally discussed is the finding that most facial changes evident in transsexual individuals are manifest largely through artificial augmentation of the face (i.e., plastic surgery).

A. Facial Skeleton Sex Differences

In general, females are typically described as gracile relative to males, who are usually more robust. The terms gracile and robust in biological anthropology and human osteology refer to muscle mass linked to the bony features to which the muscles attach. That natural muscle mass tends to differ between females and males, simply due to biology/genetics, means that the female skull, for example, is characterized by more gracile features, while the male skull is more robust.

The primary differences between the male and female skulls appear at the forehead region, the superior rim in the eye orbital region, the cheek bones, and the chin. Females are characterized as having a more vertical forehead than males, whereas males' frontal bone tends to slop slightly posteriorly. Further, the supraorbital brow ridge (the "bump" of bone just superior to the eye orbits) is robust in males, and either gracile or virtually absent in females [22]. This sex difference is best seen in from a lateral view (Figure 4).

The superior rim in the eye orbital region, known as the superior orbital margin, is regarded as thin or sharp for females and thick, blunt, or round for males [22], [23]. Moreover Dempf and Eckert [24] noted that in females the orbits appear larger and are higher up on the face. In the malar or cheek region, the male malar (zygoma) or cheekbones are more robust. But, they have a flatter appearance than in the female. Thus, when viewing the face, females seem to display more prominent cheekbones [29]. Thus, the prominent facial differences between male and female is characterized at the forehead and the jaw region, when viewed from the frontal view.

B. Hormonal Influence on the Female and Male Facial Skeleton

A comprehensive review of the literature yielded a paucity of studies indicating the effects of hormone therapy on adult facial skeletal remodeling. Although the circulating levels of testosterone and estrogen are often assumed to be associated with the masculinity and feminity, there is not much empirical evidence to verify this notion. A study [25] on judging the level of masculinity using the digital composites of men known to have high and low testosterone levels showed that the participants identified high testosterone faces as more masculine 53% of the time.

Estrogen is known to play a role in skeletal maturation and resultant adult bony proportions [26]. However, other hormones, such as Human Growth Hormone (HGH) and insulin-like growth factor, contribute to normal craniofacial and dental growth and development [27]. Becking *et al.* [28] explained that the role of hormone therapy in initiating the transsexual change process in adults, from male to female, results in changes to body hair, breast size, the appearance and texture of the skin, distribution of body fat, and the size and function of the reproductive organs. And, while androgens are largely known to have a greater impact on overall bone mass than estrogen [29], and are tied to improvements in visual memory [32], while vascular function improves in male-to-female transsexuals supplemented with estrogen [33], no further information could be found on the facial skeletal effects of androgen or estrogen increases or decreases in the male-to-female or female-to-male transsexual process.

Other studies ([3], [4]) addressing the role of hormones on adult facial appearance due to hormonal changes, discuss the soft tissue changes, with no mention of body remodeling. Much of the literature on changes in facial appearance in transsexual individuals reports the changes as surgical changes. These surgeries involve changes such as forehead reduction and remodeling, adjusting the hairline via scalp advance, eyebrow lift, rhinoplasty (shaping the nose), bimaxillary osteotomies (removing bone from the cheek area), cheek implants (fat pads), genioplasty (chin reduction), mandibular angle reduction, lip enlargement, and thyroid shave to reduce the *Adam's apple* ([28], [29]).

The above findings suggest that the sex differences in facial appearance are mainly in the forehead, jaw, and the cheek bone region of the face, and that these changes are largely derived from surgeries. It is to be noted that the *periocular region remains robust to hormonal changes over*

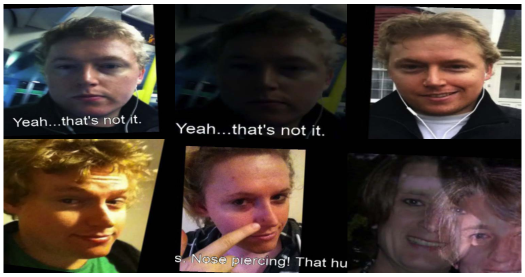


Fig. 5. Sample frames of a subject extracted by the ffmpeg decoder.

time and hence, can be considered as a reliable biometric trait in identifying individuals undergoing medical alterations. In addition, the structural complexity of the face changes over time in terms of color, texture, and shape, while the periocular region undergo minor variations over time in terms of shape and the spatial location of the eyes [21]. Also, the periocular region of the human face includes the most dense, complex, and discriminative features of the face such as contour, eyelids, eyebrow, etc. making it suitable for it to be considered as a robust biometric trait.

III. Face Component Extraction and Representation

The extraction of the eyes, nose, and mouth regions from the video sequences and its representation involves the following steps:

- Extraction of frames with valid face images from the video sequences.
- Alignment and cropping of the face components by registering the face image using the eye center coordinates.
- Representation of the face components using TPLBP, LBP, and HOG feature descriptors.

This section provides details of each of these steps in the extraction and representation of the eyes, nose, and mouth regions from a face image. Figure 3 shows the framework for face component extraction and representation of a face image.

A. Face Image Extraction

Face image extraction from real world videos has been a challenge to the face recognition community as the videos are taken under uncontrolled conditions such as variations in pose, expressions, illumination, and occlusions. These variations need to be considered in order to evaluate the robustness of any face recognition framework. It is also necessary to include a standard protocol in the frame extraction and processing steps, so that a fair comparison of the performance of various face recognition algorithms on the same dataset can be evaluated. In this section, we present a protocol that utilizes standard techniques to extract face images from the video frames.

The frames from the video sequences are extracted using the ffmpeg decoder at the rate of 29.3fps, which is the standard for video sequences. Figure 5 shows sample frames extracted by the ffmpeg decoder. The frames illustrate the challenges such as pose variations, expressions, frame transitions, illumination variations, and occlusions. Face detection is then performed using the Viola-Jones face detector [20] as implemented in OpenCV v2.4.6. The frames that were not detected by the face detector are discarded. It is to be noted that the face detector generates false positive-detects a face when there is not one-as well as false negatives-misses a face when it is present in the frame. In order to eliminate the false positives the frames undergo a second detector; this detector is an eye detector [1]. The frames in which no eyes are detected are discarded during the second pass as not containing a face. This process is efficient as it offers the detection of eye-coordinates from the face images as well as detect the false positives.

B. Alignment and Cropping

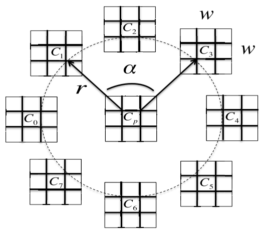
The alignment of the face region is performed using the eye-center coordinates extracted by the eye detector [1]. The alignment and cropping of the face components are achieved by means of geometric normalization. The coordinates of the eye centers are used to scale, rotate, and crop the face to a set size. These geometric transformation are performed such that the centers of the eyes are horizontally aligned and placed on standard pixel locations thereby maintaining a pseudo fix interocular distance.

Once the face image is aligned the components of the face can be obtained by defining a cropping boundary using the standard eye center locations. Cropping of the left and right periocular region is accomplished by defining a cropping boundary around each eye center. Consider the face on an xy-plane. The bounding box is created by fixing a ratio of distance for the height and width of the eyes from the eye center coordinates. The nose and mouth region are defined by a cropping boundary whose width is the distance between the two eye centers and the height being 60 pixels. The height of the nose is measured from the horizontal line that connects the two eye centers and the height of the mouth is measured from the height of the nose. The extracted face component images are then resized to 64×64 pixels.

C. Representation

The aligned and cropped periocular images are represented individually using the Three-Patch Local Binary Patterns (TPLBP) [7], Local Binary Patterns (LBP) [6], and Histogram of Oriented Gradients (HOG) [5]. While the LBP, and HOG are pixel based feature descriptors, the TPLBP is a patch-based feature descriptor that extracts features from local patches around a central patch. The descriptor from each patch is concatenated to form one global descriptor for the entire image.

The TPLBP is a variant of LBP, where a central patch encompassing a pixel location at its center is compared with its neighboring patches to generate the feature descriptor. The TPLBP descriptor is produced by comparing the values of three neighboring patches to produce a single bit value in the descriptor code. For each pixel in the image, a patch of size $w \times w$ centered on the pixel, and S neighboring patches uniformly distributed in a circle of radius r around it considered. Two neighboring patches that are α patches apart



(a)

$$
\begin{aligned}
&3P-LBP_{r,S,3,2}(p)= \\
&f(d(C_0,C_p)-d(C_2,C_p))2^0+ \\
&f(d(C_1,C_p)-d(C_3,C_p))2^1+ \\
&f(d(C_2,C_p)-d(C_4,C_p))2^2+ \\
&f(d(C_3,C_p)-d(C_5,C_p))2^3+ \\
&f(d(C_4,C_p)-d(C_6,C_p))2^4+ \\
&f(d(C_5,C_p)-d(C_7,C_p))2^5+ \\
&f(d(C_6,C_p)-d(C_0,C_p))2^6+ \\
&f(d(C_7,C_p)-d(C_1,C_p))2^7
\end{aligned}
$$

(b)

Fig. 6. Figure shows the computation of the three-patch LBP. (a) Three patch LBP. (b) Three patch LBP code computation.

are compared with the center patch and the descriptor code bit is set based on the neighboring patch that is more similar to the center patch. The TPLBP code is given by the following formula:

$$TPLBP_{r,m,w,\alpha}(p) = \sum_{i}^{S} (f(d(C_i, C_p) - d(C_{i+\alpha \mod m}, C_p))2^i \tag{1}$$

where C_i and $C_{i+\alpha}$ are two patches along the ring and C_p is the central patch. The function $d(.,.)$ is any distance function between two patches (e.g., L_2 norm of their gray level differences) and f is defined as:

$$f(x) = \begin{cases} 1 & \text{if} \quad x \geq \tau \\ 0 & \text{if} \quad x < \tau \end{cases} \tag{2}$$

where τ is set to a value slightly larger than zero in order to provide stability in uniform regions [18]. Figure 6 shows the computation of TPLBP code for a pixel.

Patch-based approaches have been shown to provide state-of-the-art capabilities in similarity learning of faces and of general images [19]. While LBP and TPLBP are, by design, more robust to photometric changes than HOG, being a gradient based representation, is robust to additive photometric changes, whereas both LBP and TPLBP are robust to any monotonic photometric transformation. All three are similarly robust to pose changes as all three pool values in spatial histograms. The advantages of these descriptors is that they are more discriminative since they operate on patches of images and hence can capture subtle texture changes well.

IV. Experiments

The following experiments were designed in order to investigate the effectiveness of the component-based representation, and ultimately periocular-based, in face recognition of subjects undergoing HRT. The primary baseline for these experiments is a holistic face representation. In this paper, "holistic face representation" refers to the features extracted from globally aligned full face image, which is the most common approach used by face recognition systems. Recognition experiments on the extended transgender dataset individually and combined with the LFW dataset [36] are also performed.

The cropped images undergoes a noise removal stage, where the Wiener filter [8] is applied. No other preprocessing is included, that is we don't attempt to correct for pose, or perform any advanced illumination mitigation, etc. The TPLBP, LBP, and HOG face representations are then computed using the aligned, cropped face image. A radius of 1 and a neighborhood size of 8 pixels is used for both TPLBP and LBP.

A. Dataset

To the best of our knowledge, there exists no dataset in the literature that includes images of a subject taken during the period of gender transformation. Hence, we collected these images from YouTube©videos that are compilations of the images taken during various time periods of the gender transformation as reported by the content uploader. It is unknown whether the subjects also went through a surgery in order to alter the face further. The images of the subjects span across years with a minimum of one year. For example, in one case it was a picture a day for three years and in others it was a picture a week or a random sampling over a year or more. The *transgender dataset*, as we call it includes more than a million extracted face images of 38 subjects . The complexity of this dataset can be clearly seen from the recognition performance of the COTS system (see figure 2). Figure 1 shows sample images of subjects before and after the transition.

B. Experimental Setup

This section details the experiments conducted to investigate the advantages of periocular-based representations, as opposed to the holistic representations. Both identification (1:N matching) and verification (1:1 matching) experiments are performed for the component-based and holistic representations.

For the component-based verification experiments, feature vectors are extracted using TPLBP, LBP, and HOG feature descriptors and individually applied to the aligned and cropped left, right periocular, nose, and mouth regions. The similarity between two feature vectors is measured as the Euclidean distance between two feature vectors. For the task of verification, two images are considered to be from the same subject if the Euclidean distance between their feature vectors is below a threshold a set value. The normalized Euclidean distance is converted to a similarity score by simply subtracting the distance from one. The periocular region is measured from the viewers perspective and not the subjects, i.e. the right periocular is the subject's left eye-region and the left periocular is the subject's right eye region. The fusion of the left and right periocular region is performed at the score level, which is obtained by a weighted combination of the similarity scores from the left and right periocular region. The following weight factors were used for this work; left 0.7 and right 0.3. The optimal value for the weight is determined from the scores obtained from our experiments for the left and right periocular regions. The authors would like to note that this body of work is not focused on optimizing the performance of the approaches discussed, and hence, the authors did not investigate the best distance metric to use or the best fusion method. The Euclidean distance and the binary-weighted score-level fusion was chosen for their simplicity.

TABLE I

COMPONENT-BASED FACE RECOGNITION. TABLE SHOWS RANK-1 RECOGNITION ACCURACY OF VARIOUS DESCRIPTORS ON THE TRANSGENDER DATASET

Approach	Left Eye	Right Eye	Fusion	Nose	Mouth	Full Face
TPLBP	51.05%	50.17%	**52.44%**	44.57%	37.75%	46.69%
HOG	42.81%	40.97%	**44.52%**	37.84%	29.64%	38.00%
LBP	30.81%	31.12%	**35.68%**	20.7%	21.94%	28.14%

The verification performance is measured in terms of the Receiver Operating Characteristic (ROC) curve, False Acceptance Rate (FAR) as the x-axis and the True Acceptance Rate (TAR) as the y-axis. The verification rate is also measured in terms of Equal Error Rate (EER), which is defined as the error rate when FAR and TAR is equal.

In addition to verification experiments, identification experiments were performed for both the closed and open set scenarios. The closed set identification was setup as follows: the gallery is constructed from the first (youngest temporally) one-third of images of each subject. The remaining two-thirds of the subject is used as probes.

C. Identification Results

The component-based recognition is performed using the extracted gallery and probe sets from the verification experiment. Euclidean distance measure is used to compute the match score between a pair of images and the pair with the least distance score is determined as the match. A similar matching procedure is followed for full-face and score level fusion approaches.

The rank-1 recognition accuracy for all the approaches is shown in Table I. The results are consistent with the assumption that the periocular region is the most discriminative feature. However, it is to be noted that the performance is significantly affected by the presence of other factors such as illumination and occlusions. The poor performance of the nose and mouth region may indicate that these face areas are subject to more changes due gender transformation. Of course, PIE does play a role for these regions, but, PIE is also associated with the periocular region. It is important to note that the very simple fusion approach used works well–results in a higher performance score than the individual periocular components—in some scenarios and not so in others. *Again this work is not focused on optimizing the simple texture approaches for face matching but, rather to demonstrate that the selection of the appropriate face regions and a good feature representation can vastly out perform commercial face recognition engines.*

The superior performance of the fused approach over the baseline (full face) and other face component representations is likely due to several factors. One of these factors is the local alignment of the periocular region using the eye centers, which results in a better localization of the region of interest as the Viola-Jones detector used has significant variability between face images of congruent quality. It has been demonstrated that eyebrows are very discriminate. Hence, focusing on the eyes, eyebrows, and the region around it with the highest degree of inter-personal variations improves the recognition accuracy. Prior work in cognitive and face recognition has indicated that the nose and mouth are also facial features with high discrimination, however, the mouth is highly deformable, which negatively affects discrimination [30], [31], [35]. The other discriminable features of the face (nose, mouth, etc.) shows significant variations with medical alterations, gender variations, etc.; hence in this scenario, these features of the face contain considerably less discriminative information. This is also evident from the recognition accuracies from the nose and mouth region. In addition, the parameters of the feature descriptors is an additional factor to be considered that can be tuned for better description of the periocular region.

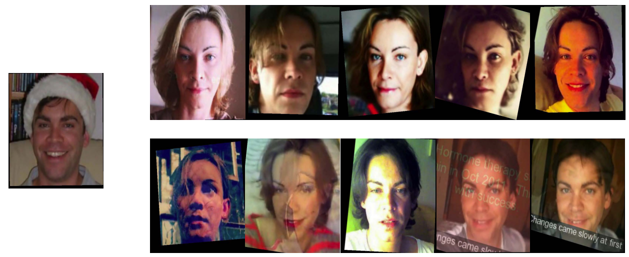


Fig. 7. Examples of face images from the Transgender database in which poor performance was obtained by the periocular-based approach proposed in this paper and the holistic approach. Image on the left shows the face before HRT transformation. Top row shows examples of poor performance by the holistic approach and bottom row shows examples of poor performance by the periocular-based approach.

TABLE II

NON-PARAMETRIC VALUES FROM THE KRUSKAL WALLIS H TEST

Descriptor	Periocular Vs Nose	Periocular Vs Mouth	Periocular Vs Full Face
TPLBP	$H = 6.437$ $p = 0.011$	$H = 4.760$ $p = 0.029$	$H = 17.29$ $p = 0.004$
HOG	$H = 4.661$ $p = 0.031$	$H = 5.129$ $p = 0.024$	$H = 3.621$ $p = 0.016$
LBP	$H = 4.246$ $p = 0.039$	$H = 4.628$ $p = 0.031$	$H = 1.621$ $p = 0.025$

Figure 7 displays face images in which both the periocular-based approach and the holistic approach performed poorly. As the figure illustrates these images of failures (from the periocular-based approach) have high degrees of PIE (either pose, illumination, and/or expression). This highlights an inherent problem with the periocular approach used in this work, which has to deal with out-of-plane rotation. Further, this figure shows that the holistic failure modes tend to occur when the face becomes more transgendered.

TABLE III
OPEN SET RECOGNITION. TABLE SHOWS RANK-1 RECOGNITION ACCURACY OF DESCRIPTORS ON THE TRANSGENDER+LFW DATASET

Approach	Left Eye	Right Eye	Fusion	Nose	Mouth	Full Face
TPLBP	46.48%	45.47%	**50.19%**	39.06%	38.77%	45.34%
HOG	37.31%	35.40%	**38.28%**	32.66%	29.21%	34.16%
LBP	27.00%	27.58%	**31.31%**	20.03%	20.84%	27.28%

D. Statistical Analysis

In order to illustrate the statistical significance of the recognition results, Kruskal Wallis H-test was conducted on the rank-1 recognition accuracies of the periocular region individually compared with other face components and the full face region for all the three descriptors. The rank-1 recognition accuracies for each video sequence (the rank-1 accuracy of each frame of the video sequence and for a total of 38 videos) were computed for each descriptor and used for the test. Table II shows the values of H, p-value, and the degrees of freedom for each of the descriptors. The p-values from all the comparisons (Periocular Vs Nose, Periocular Vs. Mouth and Periocular Vs Full Face) shows the significance of the periocular region as a useful biometric trait when compared with other face regions. The strongest presumption against the null hypothesis occurs for the comparison of the periocular region vs. full face region. This clearly shows that the periocular region can provide superior performance than the full face with the presence of gender variations.

E. Open Set Recognition

It is important for any face recognition system to reduce the risk of false positives during the recognition process. Most of the recognition systems focus on closed set recognition in which every class that is being tested is known to have a representation in the gallery. However, real world scenarios involve open set recognition that is most prominent with unknown classes. In order to show the robustness of the proposed approach in reducing the number of false positives, an open set recognition experiment was performed. It is to be noted that a general open set recognition experiment involve unknown classes in the probe set and known classes in the gallery set. The terminology *open set recognition* used in this work represent the experiment that included a gallery set with both known and unknown classes and the system was tested using the known classes (images from the transgender dataset). The prime motivation behind such a setup is to determine whether the known classes from the probe set are matched accurately to the known classes in the gallery.

The gallery included 1068 images of 400 subjects from the LFW dataset [36] in addition to the set of images from the transgender dataset (gallery set from the close set recognition). The probe set from the closed set experiment was used as the probe set for the open set experiment. The images from the LFW dataset were preprocessed with the same technique (1. eye center detection using PittPatt SDK, 2. face image alignment using geometric normalization) used to preprocess the images from the transgender dataset. This ensures that the face components are fairly extracted from both the datasets. Table III shows the recognition accuracies from all the face components and the holistic approach for all the descriptors in this experiment.

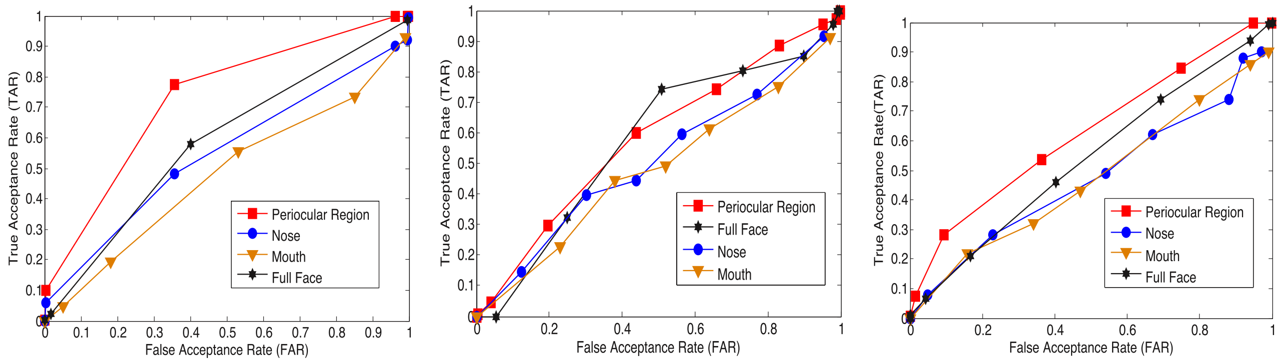
It is to be noted, that there is a degradation in the performance of the open set recognition when compared with the closed set performance, which is to be expected. However, this performance variation is insignificant in the case of periocular fusion and full face approaches for the TPLBP descriptor. This indicates that complex feature representations can effectively capture the discriminative features of the face and thus improve the recognition performance. One interesting observation is that there is insignificant variation in the performance for both the open set and closed set recognition for the mouth region from all the descriptors. This could be due to the fact that the mouth region provides less discriminative features (thus giving low recognition rates) and hence resulting in a simple feature vector for most of the images.

F. Verification Results

The motivation behind the verification experiments is to study the effectiveness of the component approach to that of the full face in verifying the identity of subjects across their gender transformation period. Operationally one could consider a scenario in which a transgender person could cause a false non-match with an official identification document, e.g. drivers license, passport, or national identification card. *Should a transgender person be required to update their official identification documents? If so, when? Who is responsible for enforcement of the update?*

The gallery and probe images from the identification experiment are utilized for the verification experiment. The task of verification is performed by categorizing a pair of images as intra-personal (belonging to the same subject) or extra-personal (belonging to different subjects). Each probe image is matched with the gallery images and a similarity score is obtained. Each image pair is determined to be an intra-personal pair, if the similarity score is higher than a pre-determined threshold. Otherwise, the image pair is determined to be extra-personal.

Figure 8 shows the verification performance of the periocular based fusion approach and the holistic approach. Table IV shows the EER (equal error rates) for the left, right, periocular fusion, and holistic approach. From the results it is evident, that the performance of the proposed approach (periocular-based fusion) is better than the holistic approach. This indicates the potential use of the periocular region as a biometric trait in scenarios where the full face is not available.



(a) (b) (c)

Fig. 8. ROC curves, performance of the components and the holistic approach against the simple descriptors. (a) TPLBP. (b) LBP. (c) HOG.

TABLE IV

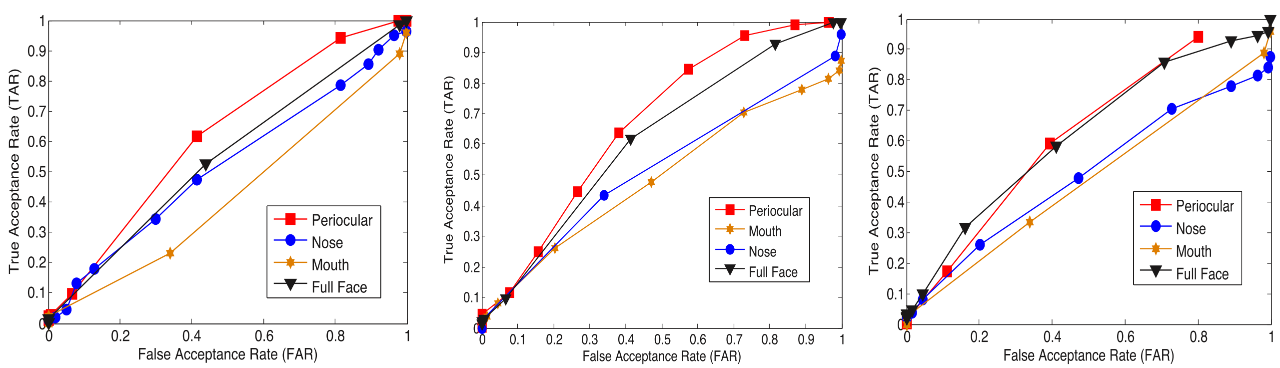
EER FOR COMPONENTS AND FULL FACE. TABLE SHOWS THE MEAN EER ACROSS ALL SUBJECTS

Approach	Left Eye	Right Eye	Fusion	Nose	Mouth	Full Face
TPLBP	0.3571	0.3748	**0.3530**	0.4247	0.4583	0.4059
HOG	0.4083	0.4098	**0.3985**	0.4404	0.4827	0.4106
LBP	**0.3741**	0.4166	0.3823	0.4359	0.4614	0.3860

TABLE V

COMPONENT-BASED FACE RECOGNITION. TABLE SHOWS RANK-1 RECOGNITION ACCURACY ON TRANSGENDER DATASET UNDER POSE VARIATIONS

Approach	Left Eye	Right Eye	Fusion	Nose	Mouth	Full Face
TPLBP	52.23%	49.95%	**57.79%**	43.28%	39.24%	46.39%
HOG	30.90%	31.48%	**36.03%**	24.31%	25.82%	27.99%
LBP	42.92%	40.42%	**46.72%**	34.79%	33.47%	37.86%



(a) (b) (c)

Fig. 9. ROC curves showing the verification performance of the component-based and the holistic approach under pose variations using three feature descriptors. (a) TPLBP. (b) LBP. (c) HOG.

An interesting aspect with the results is that the left periocular region (from the user's perspective) performs better than the right and the score based fusion for all the descriptors. This indicates that these descriptors captures side based information from the periocular region. This could also potentially indicate the discriminative power of the left periocular region than the right one. However, there exists no evidence in cognitive science and anthropology that indicates this discrimination between the eye regions. In addition, this also likely indicate the rotation invariance property of the LBP that compensates for variations in pose between the eye regions.

G. Robustness to Changes in Facial Pose

Current face recognition systems face challenges from images with significant variations in pose, illumination, and expressions. The state of the art systems perform well under controlled conditions where the above mentioned variations are eliminated from the images, either during acquisition or in preprocessing with illumination mitigation and expression and pose re-projections. This section aims at demonstrating the capability of the periocular-based region in performing well under the presence of facial pose variations as compared to the other components and full face.

TABLE VI

COMPONENT-BASED VERIFICATION AND FULL FACE VERIFICATION RESULTS UNDER POSE VARIATIONS. TABLE SHOWS THE AVERAGE EER

Approach	Left Eye	Right Eye	Fusion	Nose	Mouth	Full Face
TPLBP	0.3585	0.3745	**0.3521**	0.4139	0.4325	0.4089
HOG	0.4063	0.4096	**0.3956**	0.4296	0.4483	0.4123
LBP	**0.3721**	0.4187	0.3818	0.4206	0.4519	0.3867

TABLE VII

COMPONENT-BASED FACE RECOGNITION. TABLE SHOWS RANK-1 RECOGNITION ACCURACY OF VARIOUS DESCRIPTORS ON THE TRANSGENDER DATASET UNDER NO POSE VARIATIONS

Approach	Left	Right	Fusion	Nose	Mouth	Full Face
TPLBP	47.56%	50.97%	**55.51%**	43.21%	45.76%	45.70%
HOG	42.43%	43.05%	**45.09%**	40.28%	39.27%	37.96%
LBP	30.35%	29.64%	**34.26%**	26.92%	26.34%	28.22%

TABLE VIII

COMPONENT-BASED VERIFICATION AND FULL FACE VERIFICATION RESULTS UNDER NO POSE VARIATIONS. TABLE SHOWS THE AVERAGE EER

Approach	Left	Right	Fusion	Nose	Mouth	Full Face
TPLBP	**0.3691**	0.3869	0.3770	0.4216	0.4476	0.4137
HOG	**0.3899**	0.4166	0.3961	0.4102	0.4253	0.3997
LBP	0.4219	**0.4134**	0.4267	0.4304	0.4431	0.4136

The same alignment procedures were used as above. The gallery includes images with both frontal and off-pose images and the probe is divided into two sets, 1) images with pose variations and 2) images with no pose variations. The pose of the face image, and hence, the face components is determined using the PittPatt SDK v5.2.2 [34]. Both face verification and identification experiments are performed on the two sets.

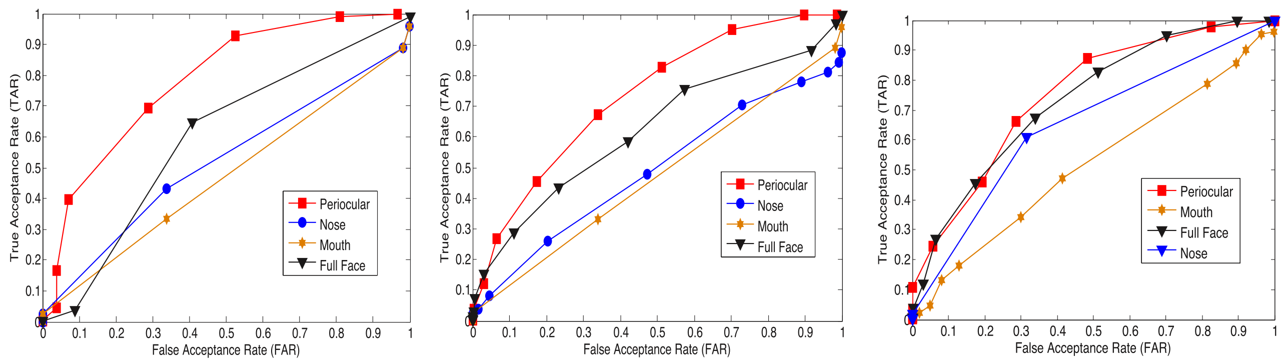
The rank-1 recognition rates for the probe with pose variations are shown in the Table V and Table VII shows the rank-1 recognition rates for the probe with no pose variations. The EER rates from the verification experiments with and without pose variations are shown in Tables VI and VIII, respectively. Figures 9 and 10 show the ROC curves obtained for the periocular based approach and the holistic approach for all the descriptors with and without pose variations, respectively. From the recognition rates it is evident that the pose variations cause insignificant differences with performance of the TBLBP periocular-based approach. This insignificant variations in performance is likely due to several factors. First, the gallery includes both frontal and non-frontal pose images that can match with the probe. However, it is to be noted that not all pose variations are included in the gallery and hence the insignificant variations in performance is attributed to the robustness of the periocular region. The local alignment of the eye region also contributes toward an accurate alignment and extraction of the periocular region improving the recognition performance.

It can be seen that there is some asymmetrical performance difference between the left and the right periocular regions for all the descriptors. This phenomenon is seen in many, if not all, periocular work. An exact understanding of this phenomenon is not in the scope of this paper, however, the authors suspect that the difference is a result of the algorithms and the starting position of analysis. Also, the impact of pose variations can be seen in the performance of the holistic approach when compared with the periocular-fusion based approach. The effects of pose variations on other face components can be clearly seen from their respective performances.

It is to be noted that the performance of the nose and mouth regions of the face under pose and no pose variations is lower when compared with the periocular region as well as the holistic approach. This indicates an significant influence of the pose variations on these face components in the name of occlusions and shape variations. In addition, it is to be noted that the performance of these face components under no pose variations is also lower when compared with the periocular region. One possible reasoning is due to the dearth of discriminative features available from the nose region unlike the periocular region and the presence of expressions for the mouth region that changes their shape significantly.

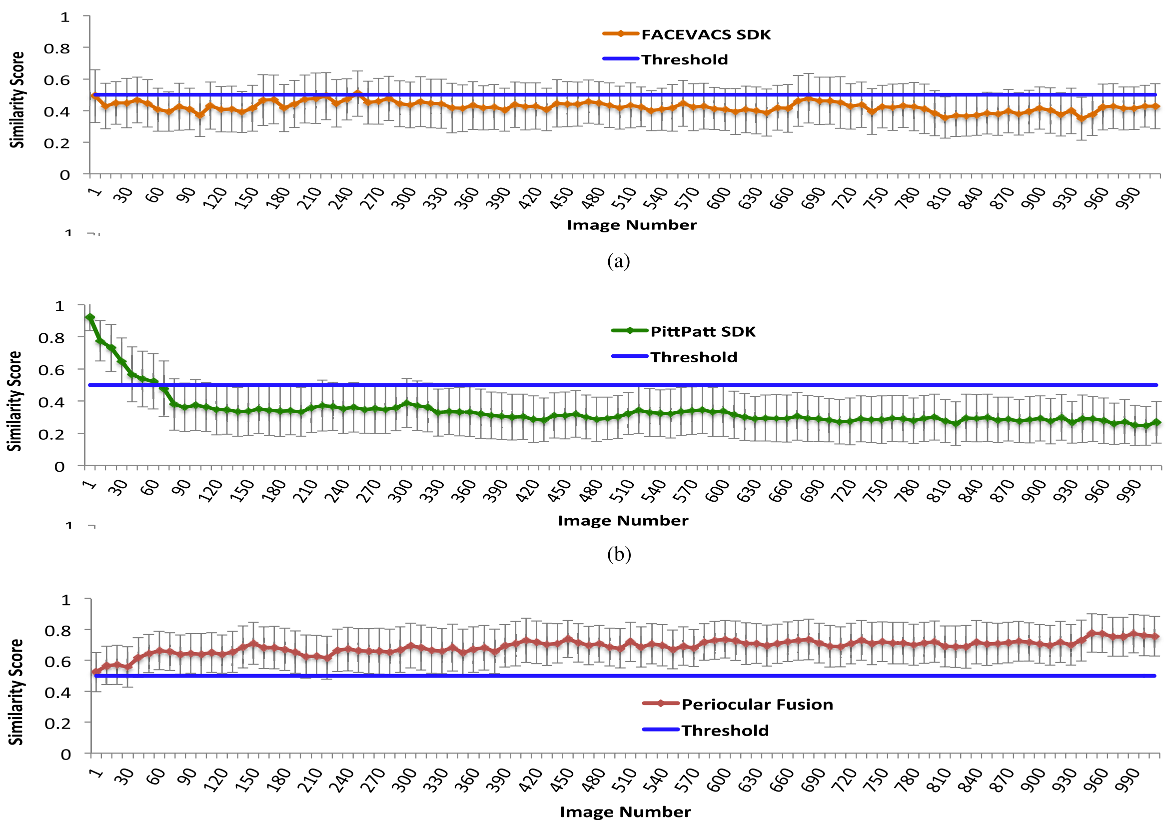
H. Comparison With COTS Systems

In order to illustrate the performance of COTS systems on the transgender dataset, recognition experiments were performed on the extended transgender dataset using PittPatt SDK v.5.2.2 and Cognetic FaceVACS v.8.5. The gallery and probe set from the closed set experiment explained in section IV-C are retained for this experiment. However, the full face images were used for training and testing the COTS systems, while the aligned periocular images represented using the TPLBP descriptor are used to test the proposed system. The aim of this experiment is to determine how well the COTS systems can identify an intra-personal (images from the same subject) pair of images. Hence, the gallery and probe of the same subject were matched to determine



(a) (b) (c)

Fig. 10. ROC curves showing the verification performance of the component-based and the holistic approach under no pose variations using three feature descriptors. (a) TPLBP. (b) LBP. (c) HOG.



(a) (b) (c)

Fig. 11. Comparison of similarity scores from the COTS systems (PittPatt SDK v5.2.2 and Cognetic FaceVACS v8.5) and TPLBP periocular fusion. Figure shows the average similarity scores for (a) FaceVACS, (b) PittPatt, and (c) TPLBP respectively. X-axis shows the n^{th} image matched to the original (pre-HRT) and Y-axis is similarity score. The std. deviation of the scores from all the videos are shown as the errors bars.

the similarity scores from the COTS systems. These similarity scores can represent the robustness of the COTS systems in matching the images of the same subject across age and gender variations.

Table IX shows the recognition accuracies from the COTS systems and the proposed periocular-based fusion approach. It is to be noted that the performance of the proposed approach is superior than both the COTS systems. This clearly shows that the existing face recognition systems may fail or perform poorly on challenging datasets that include medical alterations. On the other hand, it is also evident that the periocular region can provide robust description with much complex descriptors and hence improved performance.

TABLE IX

RANK-1 RECOGNITION ACCURACIES FROM THE COTS SYSTEMS AND THE PROPOSED APPROACH ON THE EXTENDED TRANSGENDER DATASET

FaceVACS	PittPatt [34]	Periocular Fusion
36.99%	29.37%	57.79%

Figure 11 shows the similarity scores obtained for the COTS systems and the proposed approach. The similarity scores from the COTS systems indicate the significant facial variations across HRT resulting in a poor performance by these systems. Also, it is to be noted that these scores are much lower than the threshold level (50% chance), unlike the periocular-based fusion approach. The scores from the proposed approach tend to increase besides facial variations during the course of HRT, substantiating the invariant nature of the periocular region to these changes.

V. Conclusion

The motivation behind this work is to demonstrate the reliability and robustness of the periocular region based representation in automated face recognition on subjects undergoing gender transformation using hormone replacement therapy. The potential of the periocular region is studied through various recognition and verification experiments on a challenging dataset and performance comparisons with COTS systems. Indeed, face recognition accuracies of the proposed periocular region based representation greatly exceed the accuracies obtained using the other face components and holistic representations. In addition, more precise approach for the periocular region alignment extends the possibility of improving the recognition accuracy of a recognition system. Further, this work demonstrates that the periocular is a much better region for face matching than the other major components as well as the full face, at least for this problem domain.

This work has demonstrated that simple texture features combined with a naive fusion for the periocular can be a formidable face matcher. So, formidable that it outperforms the COTS systems used as a baseline. Performance improvements were seen of *76.83%* and *56.23%* for rank-1 accuracy over PittPatt SDK v5.2.2 and Cognetic FaceVACS v8.5 respectively.

Further this work crystalizes the problem of HRT based medial treatment on the face, illustrating that the impacts of this treatment fundamentally disrupts face recognition engines. This disruption occurs quit earlier in its use, i.e. the COTS face matchers began to fail after only a few months. It is further believe by the authors that the changes, at least, for the short term use are reversibile. That means that some enterprising spoofer could self medicate with hormone drugs to fool a face recognition systems whether for access control or to gain entry to a foreign country.

In conclusion, we have demonstrated the effectiveness of the periocular region as a useful biometric trait for the unique scenario of recognizing individuals across gender transformations, and in future possibly across other medical alterations. The periocular may be a viable replacement for full face matchers under the right circumstances.

References

[1] D. S. Bolme, B. A. Draper, and J. R. Beveridge, "Average of synthetic exact filters," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 2105–2112.

[2] H. Kopka and P. W. Daly, *A Guide to LaTeX*. 3rd ed. Harlow, U.K.: Addison-Wesley, 1999.

[3] G. Hall and T. J. Philips, "Estrogen and skin: The effects of estrogen, menopause, and hormone replacement therapy on the skin," *J. Amer. Acad. Dermatol.*, vol. 54, no. 4, pp. 555–568, 2005.

[4] M. T. Patriarca *et al.*, "Effects of topical estradiol on the facial skin collagen of postmenopausal women under oral hormone therapy: A pilot study," *Eur. J. Obstetrics Gynecol. Reproductive Biol.*, vol. 130, no. 2, pp. 202–205, 2007.

[5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 886–893.

[6] T. Ojala, M. Pietikäinen, and D. Harwood, "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions," in *Proc. 12th IAPR Int. Conf. Pattern Recognit.*, vol. 1. 1994, pp. 582–585.

[7] L. Wolf, T. Hassner, and Y. Taigman, "Descriptor based methods in the wild," in *Proc. Real-Life Images Workshop Eur. Conf. Comput. Vis. (ECCV)*, Oct. 2008.

[8] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Hoboken, NJ. USA: Wiley, 1949.

[9] P. J. Phillips *et al.*, "An introduction to the good, the bad, and the ugly face recognition challenge problem," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recognit.*, Mar. 2011, pp. 346–353.

[10] N. Ramanathan and R. Chellappa, "Face verification across age progression," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2006, pp. 387–394.

[11] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs, "Face verification across age progression using discriminative methods," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 1, pp. 82–91, Mar. 2010.

[12] G. Mahalingam and C. Kambhamettu, "Face verification of age separated images under the influence of internal and external factors," *Image Vis. Comput.*, vol. 30, no. 12, pp. 1052–1061, 2012.

[13] P. G. Sator, J. B. Schmidt, M. O. Sator, J. C. Huber, and H. Honigsmann, "The influence of hormone replacement therapy on skin ageing: A pilot study," *Maturitas*, vol. 39, no. 1, pp. 43–55, 2001.

[14] G. Aggarwal, S. Biswas, P. J. Flynn, and K. W. Bowyer, "A sparse representation approach to face matching across plastic surgery," in *Proc. IEEE WACV*, Jan. 2012, pp. 113–119.

[15] F. Juefei-Xu and M. Savvides, "Unconstrained periocular biometric acquisition and recognition using COTS PTZ camera for uncooperative and non-cooperative subjects," in *Proc. IEEE Workshop Appl. Comput. Vis. (WACV)*, Jan. 2012, pp. 201–208.

[16] U. Park, R. R. Jillela, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 96–106, Mar. 2011.

[17] S. Yoon, J. Feng, and A. K. Jain, "Altered fingerprints: Analysis and detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 3, pp. 451–464, Mar. 2012.

[18] M. Heikkiia, M. Pietikainen, and C. Schmid, "Description of interest regions with center-symmetric local binary patterns," in *Proc. 5th Indian Conf., Comput. Vis., Graph. Image Process.*, 2006, pp. 58–69.

[19] E. Nowak and F. Jurie, "Learning visual similarity measures for comparing never seen objects," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.

[20] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2001, pp. I-511–I-518.

[21] E. Masoro and S. Austad, *Handbook of Biology of Aging*, 7th ed. New York, NY, USA: Academic, 2010.

[22] T. D. White, M. T. Black, and P. A. Folkens, *Human Osteology*, 3rd ed. New York, NY, USA: Academic, 2011.

[23] K. R. Burns, *Forensic Anthropology Training Manual*, 3rd ed. London, U.K.: Pearson, 2012.

[24] R. Dempf and A. W. Eckert, "Contouring the forehead and rhinoplasty in the feminization of the face in male-to-female transsexuals," *J. Craniomaxillofac Surgery*, vol. 38, no. 6, pp. 416–422, 2010.

[25] I. S. Penton-Voak and J. Y. Chen, "High salivary testosterone is linked to masculine male facial appearance in humans," *Evol. Human Behavior*, vol. 25, no. 4, pp. 229–241, 2004.

[26] V. Rochira, A. Balestrieri, M. Faustini-Fustini, and C. Carani, "Role of estrogen on bone in the human male: Insights from the natural models of congenital estrogen deficiency," *Molecular Cellular Endocrinol.*, vol. 178, nos. 1–2, pp. 215–220, 2001.

[27] K. E. Simmons, "Growth hormone and craniofacial changes: Preliminary data from studies in Turner's syndrome," *Pediatrics*, vol. 104, pp. 1021–1024, Oct. 2001.

[28] A. G. Becking, D. B. Tuinzing, J. J. Hage, and L. J. Gooren, "Transgender feminization of the facial skeleton," *Clin. Plastic Surgery*, vol. 34, no. 3, pp. 557–564, 2007.
[29] E. Van Caenegem *et al.*, "Low bone mass is prevalent in male-to-female transsexual persons before the start of cross-sex hormonal therapy and gonadectomy," *Bone*, vol. 54, no. 1, pp. 92–97, 2013.
[30] J. Shepherd, G. Davies, and H. Ellis, "Studies of cue saliency," in *Perceiving and Remembering Faces*. New York, NY, USA: Academic, 1981.
[31] J. Sadr, I. Jarudi, and P. Sinha, "The role of eyebrows in face recognition," *Perception*, vol. 32, no. 3, pp. 285–293, 2006.
[32] E. Gómez-Gil, S. Cañizares, A. Torres, F. de la Torre, I. Halperin, and M. Salamero, "Androgen treatment effects on memory in female-to-male transsexuals," *Psychoneuroendocrinology*, vol. 34, no. 1, pp. 110–117, 2009.
[33] G. New *et al.*, "Long-term estrogen therapy improves vascular function in male to female transsexuals," *J. Amer. College Cardiol.*, vol. 29, no. 7, pp. 1437–1444, 1997.
[34] *PittPatt Face Recognition SDK, Pittsburgh Pattern Recognition*. [Online]. Available: http://www.pittpatt.com
[35] G. Mahalingam and K. Ricanek, Jr., "Is the eye region more reliable than the face? A preliminary study of face-based recognition on a transgender dataset," in *Proc. IEEE 6th Int. Conf. Biometrics, Theory, Appl. Syst. (BTAS)*, Sep./Oct. 2013, pp. 1–7.
[36] L. Wolf, T. Hassner, and Y. Taigman, "Effective unconstrained face recognition by combining multiple descriptors and learned background statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 10, pp. 1978–1990, Oct. 2011.
[37] J. Zou, Q. Ji, and G. Nagy, "A comparative study of local matching approach for face recognition," *IEEE Trans. Image Process.*, vol. 16, no. 10, pp. 2617–2628, Oct. 2007.
[38] X. Liu, S. Shan, and X. Chen, "Face recognition after plastic surgery: A Comprehensive study," in *Lecture Notes in Computer Science*, vol. 7725. Berlin, Germany: Springer-Verlag, 2013, pp. 565–576.

Gayathri Mahalingam is currently a Post-Doctoral Researcher with the Institute for Interdisciplinary Studies in Identity Sciences, University of North Carolina at Wilmington, Wilmington, NC, USA. She received her Ph.D. degree in computer science from the University of Delaware, Newark, DE, USA, in 2012, and the master's degrees in mathematics and computer science from the University of South Florida, Tampa, FL, USA, and the University of Delaware, in 2005 and 2007, respectively. Her research interests include biometrics, pattern recognition, and computer vision. She has authored several publications in highly peer-reviewed biometrics conferences and journals.

Karl Ricanek, Jr. is currently a Professor of Computer Science with the University of North Carolina at Wilmington, Wilmington, NC, USA, where he joined the faculty in 2003. He is the Director of the Institute for Interdisciplinary Studies in Identity Sciences and Face Aging Group Research Laboratory, and the Codirector of the ODNI affiliated Center for Academic Excellence CASIS. He received the Ph.D. degree in electrical engineering, the master's degree in electrical engineering, and the bachelor's degree from North Carolina Agriculture and Technical State University, in 1999, 1993, and 1992, respectively. He has authored over 70 articles in biometrics and related fields. He is an Editor of the IEEE Computer Identity Sciences Column, which highlights current trends in identity sciences and biometrics.

A. Midori Albert is currently a Professor and Forensic Anthropologist with the Department of Anthropology, University of North Carolina at Wilmington, Wilmington, NC, USA, where she joined the faculty in 1995. She received the Ph.D. degree in anthropology from the University of Colorado at Boulder, Boulder, CO, USA, in 1995, and the master's degree in anthropology and the bachelor's degree in psychology from the University of Florida, Gainesville, FL, USA, in 1993 and 1990, respectively. Her educational background involves human skeletal identification methods and techniques and an understanding and exploration of normal human skeletal variability. These specializations led her to link with colleagues in computer science, statistics, and mathematics, where her primary research focus has been on adult age-related craniofacial morphological changes and their effects on computer automated face recognition technologies since 2003. Together with her colleagues, she collaborates on research with the Institute for Interdisciplinary Studies in Identity Sciences, University of North Carolina at Wilmington. Students often actively engage in research learning opportunities with Dr. Albert as interests in human biometrics and identity recognition continue to evolve. She also consults regional law enforcement agencies on skeletal human identification cases.