Face Finder: Filtering a Large Face Database using Scars, Marks and Tattoos

Unsang Park, Shengcai Liao, Brendan Klare, Jimmy Voss and Anil K. Jain
Dept. of Comp. Sci. & Eng. Michigan State Univ., E. Lansing, MI, USA 48824
{parkunsa, scliao, klarebre, vossjame, jain}@cse.msu.edu

Abstract

Facial marks (e.g. moles, freckles, and scars) are soft biometric traits that have limited utility in uniquely identifying individuals. However, facial marks can still play a crucial role both in filtering a large face gallery, as well as in distinguishing between ambiguous face comparisons such as identical twins. This research demonstrates that facial marks can indeed be used to speed up face matching and assist in individualizing identical twins. The facial mark detection process consists of two main stages: (i) Landmark detection using Active Shape Model (ASM) to represent face images in a common coordinate system, and (ii) a scale space extrema detection method to detect facial marks. All the detected marks are classified into six different types based on their morphology and color, and are encoded into a fixed length feature vector to facilitate efficient matching and retrieval. Experimental results on face retrieval using a mug shot data set consisting of 1,000 probe images (one per subject) and 100,000 gallery images (one per subject) demonstrate a significant improvement in retrieval time with only a slight loss of matching accuracy. We also demonstrate a symbolic query based retrieval capability using the facial mark based indexing scheme. The symbolic queries can be constructed based on the morphology, color, and location of a mark. This will facilitate the face retrieval process when only a verbal description of the face is available instead of a the face image as a query. Additional experiments on a twin data set of 404 images collected from 178 identical twins (89 twin pairs) further demonstrates improvements in differentiating identical twins using face marks.

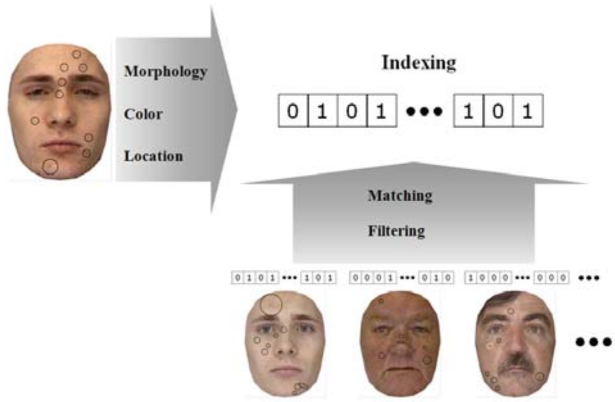
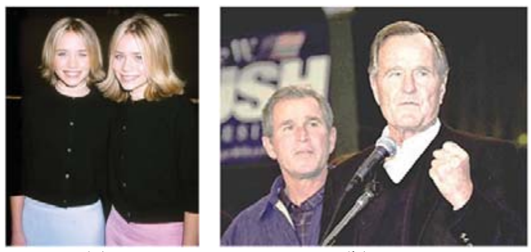


Figure 1. Schematic of the proposed mark based indexing system. The face index can be used to improve individuality or speed up the retrieval process.



(a) (b)

Figure 2. Face image pairs with very similar appearance that can confound automated face recognition systems (e.g., twins or families). Pictures are from (a) http://celebritygossips.ca/2011/03/ashley-mary-kate-olsen and (b) http://www.digalist.com/list/2382.

1. Introduction

In spite of the significant improvements in face recognition technology over the past two decades, unconstrained face recognition still remains a challenging problem due to large intra-subject variations and small inter-subject variations in the facial appearance. A number of studies are being conducted to improve the face recognition performance by developing robust feature representation schemes or advanced sensing technologies.

Conventional face recognition systems typically encode the face images by utilizing either local or global texture features. Local techniques first detect the individual components of the human face (i.e, eyes, nose, mouth, chin, and ears), prior to encoding the textural content of each of these components (e.g., EBGM and LFA) [15] [12]. Global (or holistic) techniques, on the other hand, consider the entire face as a single entity during the encoding process (e.g., PCA, LDA, etc.) [4]. In these approaches, small distinctive marks on the face are often considered as noise and not explicitly used in the matching process. However, as shown in the Face Recognition Vendor Test (FRVT 2006) [13], the skin details such as facial marks are important to achieve a high recognition accuracy. At the same time, facial marks can be used to filter the database to speed up the retrieval process or differentiate identical twins. Facial marks belong to a more general term, Scars, Marks, and Tattoo (SMT), which is gaining increasing attention for its utility of subject identification under non-ideal conditions for forensic investigation [1].

There have been only a few studies utilizing facial marks for recognition purposes [10][6][14]. Previous studies on facial marks have primarily focused on improving the face recognition performance using standard face image data sets. Park and Jain [10] showed that facial marks can be used in differentiating identical twins using semi-automatic method an a small twin data set.

We propose an automatic facial mark detection method and utilize the marks to filter a large face database for efficient image retrieval and to differentiate identical twins. We show the utility of facial marks in three different ways.

- Indexing face images to filter or retrieve candidate face images from a large database.
- Symbolic queries constructed based on the morphology, color, and location of a mark to retrieve candidate face images.
- A combination of face marks with a face matcher to better individualize subjects, especially identical twins.

Fig. 1 shows the schematic of the proposed facial mark based indexing system. The index is constructed based on the morphology, color, and location of marks observed in each face image. Fig. 2 shows example images where the inter-subject variation is very small because of the familial relationship. We will show that the proposed indexing scheme helps to individualize identical twins.

The rest of this paper is organized as follows: sections 2 and 3 describe our mark detection and classification process, respectively; section 4 presents the mark based indexing scheme, and section 5 provides experimental results and discussions. Section 6 summarizes our contributions and lists some directions for future work.

2. Facial Mark Detection

The proposed mark detection method is based on the scale space analysis [7] that detects local extrema in scale space representation of an input image. To avoid detecting local extrema around primary facial features (e.g., eye brows, eyes, nose and mouth), facial landmark detection and masking process is applied as in [10]. The complete facial mark detection procedure is illustrated in Fig. 3. The overall mark detection process consists of (i) primary facial landmark detection (ASM), (ii) mapping to the mean shape, (iii) mask construction, (iv) scale space extrema detection on non-masked region, and (v) post processing. We will discuss each step of the mark detection process briefly in the following sections. More detailed explanations of the mark detection process, except the scale space extrema detection, can be found in [10].

2.1. Primary Facial Feature Detection

We use Active Shape Model (ASM) [9] to automatically detect 68 landmarks that delineate the primary facial features: eye brows, eyes, nose, mouth, and face boundary (Fig. 3). For more reliable detection of the 68 landmark points, we first detect three landmark points (two eyes and nose tip) using a commercial SDK [2] and use them for the initial alignment of the ASM model. The 68 ASM landmarks do not include the forehead region because of the poor localization of the forehead region mostly due to the hair. Therefore, we interpolate the forehead landmarks from the 68 points using a reduced morphable model approach as in [11]. The final number of landmarks are 82, where 13 additional points are for delineating the forehead.

2.2. Mapping to the Mean Shape

Given the landmarks in our gallery database, we select 1,000 face shapes and calculate the mean face shape. Each face image is mapped to the mean shape using the Barycentric coordinate system [5] as in [10].

2.3. Mask Construction

We construct a generic mask and derive a user specific mask [10] to suppress the local extrema detection around the primary facial features. The user specific mask covers small misfit of the landmarks around the primary facial features and suppress false positives around small wrinkles or beards that are connected to the primary facial features.

2.4. Scale Space Extrema Detection

Lindeberg proposed that local maxima over multiple image scales of normalized Gaussian derivatives (i.e., $\sigma^2\nabla^2 G$) reflects a characteristic size of a local structure [7]. This allows for a blob detection with automatic scale selection, which is invariant with image scale. Motivated by this, we detect facial marks via scale space analysis. The scale space extrema detection starts by constructing a normalized multiscale representation of the face image by convolving the input images, $I(x, y)$, with a Laplacian of Gaussian (LoG)

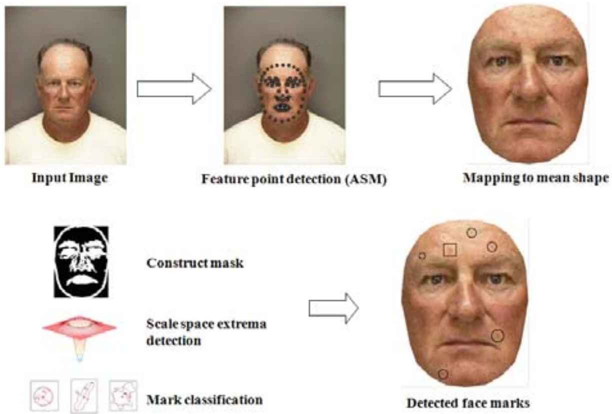


Figure 3. Schematic of the mark detection process.

filter with a sequence of σ_k as

$$D(x, y, \sigma_k) = \sigma_k^2 \nabla^2 G(x, y, \sigma_k) * I(x, y), \;\; k = 1, 2, \cdots, n, \tag{1}$$

where $\sigma_k^2 \nabla^2 G$ is the scale-normalized Laplacian of Gaussian operator, and $\sigma_k = k\sigma_0, k = 1, 2, \cdots, n$, with σ_0 being a constant value (=$\sqrt{2}$) for the initial scale.

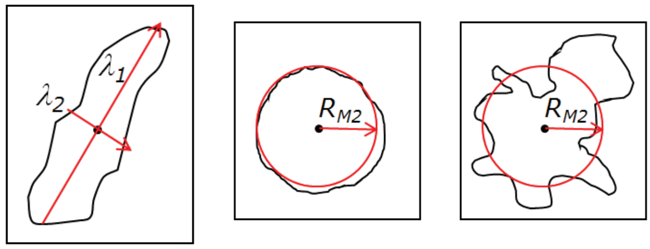
Next, we detect local extrema over both spatial and scale direction in every $3 \times 3 \times 3$ image block[1] [8]. The detected candidate mark locations have the following properties.

- A detected location contains candidate facial marks.
- The detected scale (σ_k) indicates the size of the corresponding facial mark.
- The absolute value of $D(x, y, \sigma)$ reflects the strength of the response. We use this strength as the confidence value to select stable marks.
- The sign of $D(x, y, \sigma)$ provides an evidence for the type of detected facial mark. Positive (negative) sign represents a dark (bright) facial mark with brighter (darker) surrounding skin.

2.5. post processing

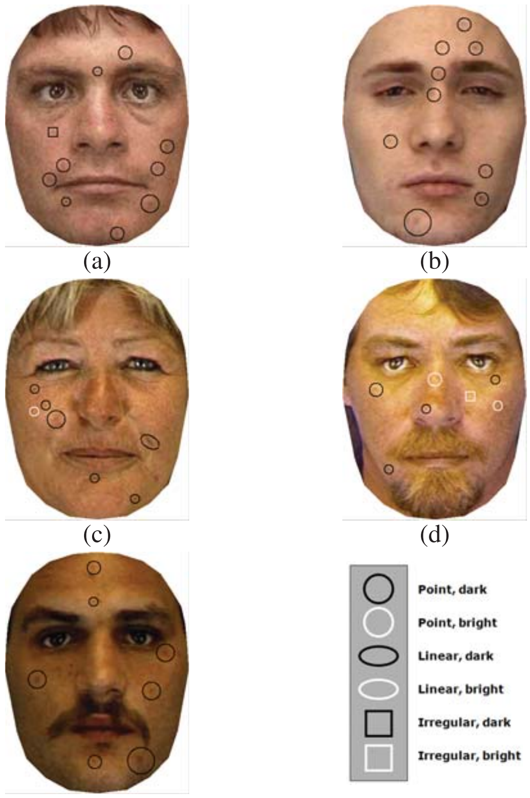
Finally, we apply a post-processing step to refine the mark detection results: (i) Delete candidates with response strength below a threshold T, (ii) delete candidates that overlap with the masked area which contains primary facial features, and (iii) sort the remaining candidates in a descending order of the corresponding response strength. Up to top K (e.g. 10) candidates are retained as the final facial marks.

[1]The LoG response of each pixel is compared with its neighboring values which forms a $3 \times 3 \times 3$ block.



(a) Linear (b) Circular (c) Irregular

Figure 4. Schematic of the morphology based mark classification.



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

Figure 5. Example mark detection and classification results. (a), (b), (c), (d), (e) are example images with marks. (f) Symbols used to denote six different mark classes.

3. Facial Mark Classification

For each detected local extrema, we find a bounding box of size proportional to the associated scale. Pixels in the bounding box are binarized with a threshold value selected as the mean of the surrounding pixels. A local extrema is darker or brighter than its surrounding region, so the average value of the surrounding area can serve to effectively segment the blob in the center area of the bounding box. We then classify the blobs in a hierarchical fashion: linear vs.

all, followed by circular (point) vs. irregular. For the linearity classification of a blob, two eigen values λ_1 and λ_2 are obtained from the eigen decomposition of the spatial coordinates of blob pixels. When λ_1 is significantly larger than λ_2 the mark is decided as a linear blob. For the circularity detection, we calculate the second moment of the blob pixels, M_2. A circle, R_{M_2}, with radius M_2 will enclose most of the blob pixels if they are circularly distributed. Therefore, a decision can be made based on the ratio of the number of pixels within and outside of R_{M_2}. The color of the blob is decided according to the sign of $D(x, y, \sigma)$ as explained in section 2.4.

The schematic of blob classification is shown in Fig. 4. Fig. 5 illustrates five examples of the facial detection results using the proposed mark detection and classification method. We observe that the proposed method is robust with noise and estimates the size and class of the mark reasonably well.

4. Facial Mark Based Indexing

The detected facial marks are encoded into a 48-bin histogram representing the morphology, color, and location of facial marks. To encode the location information of facial marks, the face image in the mean shape space is subdivided into eight different regions as shown in Fig. 6. Each mark is encoded by a six digit binary number representing its morphology and color. When there is more than one mark in the same region, a bit by bit summation is performed. The six bin values are concatenated for the eight different regions in the order as shown in Fig. 6 to generate the 48-bin histogram. If a mark is observed on the borderline of the face segments, it is included into both regions considering the variations of the segments across multiple face images of the same subject. Given indices obtained from face images, histogram intersection method is used to calculate the matching scores or filter candidate face images[2]. Let $H^1(i)$ and $H^2(j)$ be the two histogram representing the mark indices, then the histogram intersection is calculated as $\sum_{k=1}^{48} \left(H^1(k) \& H^2(k)\right)$, where $\&$ represents the logical *and* operation. The score range of the mark index based matching is 0∼48.

5. Experimental Results

5.1. Database

We used two different data sets, named, DB1 and DB2, to evaluate the effectiveness of the proposed mark detection method. DB1 is a subset of the Pinellas County Sheriff Office (PCSO)[3] face mug shot database, consisting of

[2]When both entries are non-zero, the intersection becomes one, and zero otherwise.

[3]This database can be obtained by contacting Pinellas County Sheriff Office, but it is not generally considered as a public domain database.

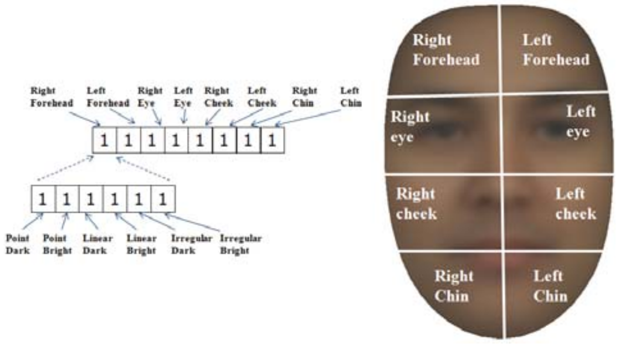


Figure 6. Schematic of the mark based indexing scheme.

1,000 probe images (one per subject) and 100,000 different gallery images (one per subject). DB2 is a twin data set consisting of 404 images of 178 identical twins (89 twin pairs) collected by University of Notre Dame in 2009. Database DB1 is used for the filtering and symbolic query based retrieval experiments using the proposed mark based indexing scheme. Both DB1 and DB2 are used to evaluate the effectiveness of the indexing scheme to improve facial individuality.

5.2. Large Scale Face Image Retrieval

Some of the facial marks are not stable due to the instability of marks (e.g., pimples or zits) or the poor image quality (e.g., low resolution or image blur). Therefore, the mark based index cannot be used to filter the database with the same high confidence as the demographic information such as gender or ethnicity. We propose a conditional filtering process using a leading COTS face recognition engine (FaceVACS [3]) as a measure of the quality of the retrieved results. The conditional filtering process includes the following steps.

i Mark index based filtering is first applied to retrieve a short list of gallery images.

ii FaceVACS is applied to this list to sort gallery images according to their similarity to the probe image in a descending order.

iii The maximum match score from FaceVACS in the filtered list is subjected to the following test; if it is lower then a threshold value, t_{FV}, then FaceVACS is used to match all the images in the gallery (the correct mate(s) is accidentally filtered out); otherwise, the filtered list serves to generate the retrieval result.

In the worst case scenario, the matching time of this scheme will be slightly increased because of the additional mark index based matching in addition to the FaceVACS.

We have measured the average matching time using a small database with 1,000 probe and 1,000 gallery both for FaceVACS and mark index based matcher (template generating time is not included). The entire matching time on DB1 is extrapolated as 6 hours and 12.5 minutes with FaceVACS and mark index, respectively. Therefore, it is expected that applying mark index based filtering first followed by FaceVACS on the small set of candidate images will lead to a more efficient search process. In the filtering experiment with DB1, the overall matching time is decreased by 1.1 hour (18.4% speed up) with 7.1% loss of matching accuracy (t_{FV}=0.6), or 24.8 min (7% speed up) with 0.5% loss of rank-1 matching accuracy (t_{FV}=0.9)[4]. Fig.7 shows the accuracy-penetration curve of the mark index scheme. The accuracy represents the probability of a correct answer is included in the short retrieval result and the penetration represents the average number of images in the filtered list. Ideally, a higher accuracy with minimal penetration rate is desirable. In general, the instability of some of the marks requires a large penetration rate to achieve a high retrieval accuracy.

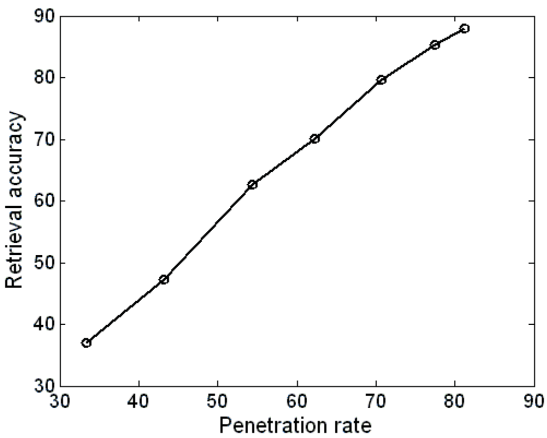


Figure 7. Accuracy vs. penetration curve.

5.3. Symbolic Query Based Retrieval

There are situations when the gallery database needs to be searched in the absence of a query image. These situations occur when there is no surveillance or any other type of camera available to capture a suspect's facial image during a criminal activity. However, a witness to the crime is able to provide a verbal description in the form a statement like "white Caucasian male, about 30 years old, tall with a dark mole on the right cheek." Our goal is to build a face retrieval system that can utilize the demographic information as well as the facial mark information to drastically filter the gallery. It is hoped that the witness will now be ale to examine this short list of gallery face images and positively identify the suspect. Our mark based index scheme can be used for this non-image (or symbolic query) based retrieval. The mark index is constructed by any marks observed in the suspect's face in terms of the morphology, color, and location. This mark index is then compared with the indices of the gallery images to retrieve the candidate images.

We used a 100 face images from the probe images in DB1 to evaluate the symbolic query based retrieval method. We manually marked distinctive marks on the 100 face images, and derived symbolic queries based on the labeled marks. For example, a point dark mark can be interpreted as a mole, and a white spot can be interpreted as a blemish. Some of the example queries are listed below.

- A mole on the right side of forehead.
- A mole on the left side of the chin.
- a scar on the right cheek

...

Since we know which symbolic query is generated from which subject in the probe data, and we also have the mate images in the gallery database, we can evaluate these symbolic queries in terms of the retrieval accuracy. About 48% of the symbolic queries successfully retrieved correct mates in the filtered results. This is a still good performance considering that a conventional image based face matcher cannot operate with these symbolic queries. Fig. 8 shows example retrieval results from symbolic queries. The retrieved face images of the correct subjects from whom the symbolic queries are constructed are enclosed with red boxes.

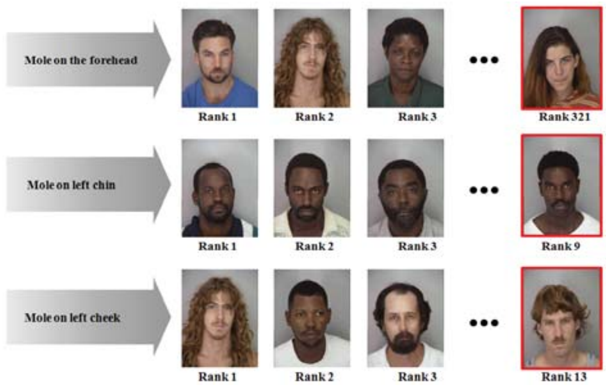


Figure 8. Example retrieval results using symbolic queries.

5.4. Individualizing Identical Twins

In DB2, there are about 2.27 images per subject, collected from 178 identical twins. Given N_{DB2} images in

[4]The match scores of FaceVACS are in the range [0,1].

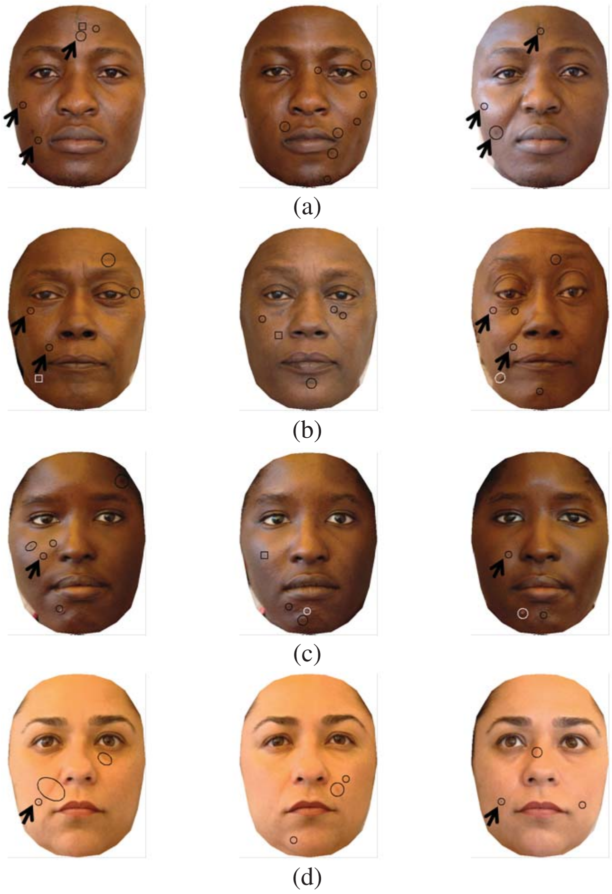


Figure 9. Example matching results with four different identical twins. Images in the first and third columns belong to the same subject and images in the second column are the twins. The first, second, and third columns correspond to the probe, incorrect match using FaceVACS only, and correctly matched true mates using FaceVACS and mark index, respectively. Black arrows in each row show the facial marks correctly detected and classified to contribute to individualize the identical twins.

DB2, we used each image as a probe in turn and the rest of the images ($N_{DB2} - 1$) as gallery to evaluate the identification accuracy (leave-one-out evaluation). The rank-1 identification accuracy improves from 90.8% (FaceVACS) to 92.1% by the weighted score-sum fusion of FaceVACS and mark index. A simple weighted score-sum method is used for the fusion. Similar experiment with DB1 reveals the rank-1 matching accuracy improves from 56.3% to 57.2% after fusion. This demonstrates that the facial marks are more effective in individualizing subjects with similar appearance (i.e., identical twins) than regular subjects. Fig. 9 shows example matching results with DB2, where FaceVACS failed, but the fusion of FaceVACS and mark index succeeded.

6. Conclusions and Future Work

We have proposed a new facial mark detection, classification, and indexing scheme using the facial marks. We have shown that the proposed facial based indexing scheme (i) helps filter the face image database to expedite the retrieval process with minimal loss of face recognition accuracy by employing a conditional filtering method, (ii) enables symbolic query based face image retrieval, and (iii) improves the individuality both in normal and identical twin face image database. The mark index scheme can also be coupled with demographic information as shown in [10].

Contrary to the previous studies that used the facial marks implicitly or with poor accuracy, we presented the most effective method so far to explicitly utilize individual facial mark. We are currently working to further improve the robustness of the facial mark detection and classification process and to develop more effective facial mark based index scheme.

References

[1] FBI Next Generation Identification, http://www.fbi.gov/about-us/cjis/fingerprints_biometrics/ngi/ngi2. 2

[2] Face Recognition Software Development Kit (PittPatt SDK), http://www.pittpatt.com. 2

[3] FaceVACS Software Developer Kit, Cognitec Systems GmbH, http://www.cognitec-systems.de. 4

[4] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Trans. PAMI*, 19(7):711–720, 1997. 2

[5] C. J. Bradley. *The Algebra of Geometry: Cartesian, Areal and Projective Co-ordinates*. Bath: Highperception, 2007. 2

[6] D. Lin and X. Tang. From macrocosm to microcosm. In *Proc. CVPR*, pages 1355–1362, 2006. 2

[7] T. Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):79–116, 1998. 2

[8] D. G. Lowe. Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004. 3

[9] S. Milborrow and F. Nicolls. Locating facial features with an extended active shape model. *ECCV*, 2008. http://www.milbo.users.sonic.net/stasm. 2

[10] U. Park and A. K. Jain. Face matching and retrieval using soft biometrics. *IEEE Transactions on Information Forensics and Security*, 5(3):406–415, 2010. 2, 6

[11] U. Park, Y. Tong, and A. K. Jain. Age invariant face recognition. *IEEE Transactions on Pattern Anal. Mach. Intell.*, 32(5):947–954, 2010. 2

[12] P. S. Penev and J. J. Atick. Local feature analysis: a general statistical theory for object representation. *Network: Computation in Neural Systems*, 7:477–500, 1996. 1

[13] P. J. Phillips, W. T. Scruggs, A. J. O'Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott, and M. Sharpe. Face Recognition Ven-

dor Test 2006: FRVT 2006 and ICE 2006 Large-Scale Results, Tech. Report NISTIR 7408, NIST, 2007. 2

[14] J. S. Pierrard and T. Vetter. Skin detail analysis for face recognition. In *Proc. CVPR*, pages 1–8, 2007. 2

[15] L. Wiskott, J.-M. Fellous, N. Kruger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Trans. PAMI*, 19(7):775–779, 1997. 1