

On a Taxonomy of Facial Features

Brendan Klare and Anil K. Jain

Abstract—After nearly a decade of intensive research in face recognition, no standard organization exists for grouping the salient information available in 2D face images into feature categories. At the same time, human verification of a subject’s identity based on facial images lacks a consistent methodology. In this paper we propose a taxonomy of available facial features that: (i) serves as a precursor to studies on the individuality of facial features, (ii) follows a similar well established and accepted organization for fingerprint features, and (iii) contains features computable by both machines and humans as well as by machines alone. This manuscript is intended as a strawman of an organization of facial features, that would hopefully lead to a standardization of such features. Such a facial feature organization will (i) enable studies on the individuality of facial features, which has important ramifications for the acceptance of expert testimony in legal proceedings for determining the identity of an individual from a facial photograph, and (ii) help standardize the framework of commercial face recognition systems.

I. INTRODUCTION

We propose to organize the vast gamut of facial features leveraged in automated and manual face recognition into three levels: Level 1, Level 2, and Level 3. Level 1 features consist of gross facial characteristics that are easily observable in a face, such as skin color, gender, and the general appearance of the face. Level 2 features consist of localized face information that requires specialized cortex processing, such as the structure of the face, the relationship among facial components, and the precise shape of the face. Level 3 features consist of certain irregularities in the facial skin, which includes micro features such as facial marks, skin discoloration, and moles. An example of this proposed feature grouping can be found in Figure 1.

Our categorization of facial features intends to provide a better understanding and standardization of both manual and automated face recognition processes. The benefit of this categorization is two fold: (i) facilitating an individuality measure for face images that can be used in legal testimony, and (ii) improving the accuracy of commercial matchers through a more careful selection of facial features. The current fingerprint feature categorization [1], accepted by both forensic scientists as well as fingerprint vendors, served as a guiding principle for our categorization of facial features. Compared to face recognition, fingerprint matching has over

100 years of history and success. Furthermore, features used in automatic fingerprint matchers (AFIS) are compact and have a physical interpretation in terms of the ridge flow patterns in the fingerprint. Indeed human fingerprint examiners essentially use the same features that are utilized in state-of-the-art AFIS. This is not necessarily true for face recognition; features extracted by humans have not yet been precisely described, and thus cannot be compared to features extracted by machines. Salient features in fingerprints are categorized into three levels: Level 1 features encompass the global structure or ridge pattern (e.g. arch, loop, whorl). Level 2 features consist of minutiae location and orientation, and are primarily used for matching. Level 3 features consist of information available at higher spatial resolutions, such as dots, incipients and ridge width. An example of these fingerprint features can be found in Figure 2. The analogy between these widely accepted fingerprint feature levels and the proposed face feature levels will be established in this manuscript.

A major benefit of the proposed facial feature taxonomy is that the same feature levels can be defined for both face recognition engines as well as human face examiners. The lack of a well defined and accepted method used in human face identification is being noticed as automated face recognition systems continue to mature [5]. The rapid growth in the use of face images captured from surveillance cameras in legal proceedings in courts has also drawn into question the methods by which human face examiners determine a person’s identity using typically low quality video frames [6]. The absence of a defined set of face features prevents: (i) a generally well accepted method of human face examination, and (ii) an understanding of the statistical uniqueness of face features derivable by humans [5], and (iii) a likelihood of a false association occurring in automated face recognition systems. Ongoing studies on the individuality of fingerprints [7] are also motivated by challenges to fingerprint evidence in court cases. A report from the National Academy of Sciences on forensics [8] highlights the need for such individuality studies not only for fingerprints but for other biometric traits as well. A recent volume on forensic facial comparison [9] also mentions this report among and other motivating factors for developing face individuality models. Our proposed organization of facial features will assist in conducting a study on the individuality of facial features.

II. FACE FEATURE LEVELS

This section will describe in detail each of the three face feature levels. A summary of these features is provided in Table I.

B. Klare and A.K. Jain are with the Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, U.S.A {klarebre, jain}@cse.msu.edu

Anil Jain’s research was partially supported by WCU (World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology(R31-2008-000-10008-0).

The authors would like to thank Unsang Park for his valuable comments.

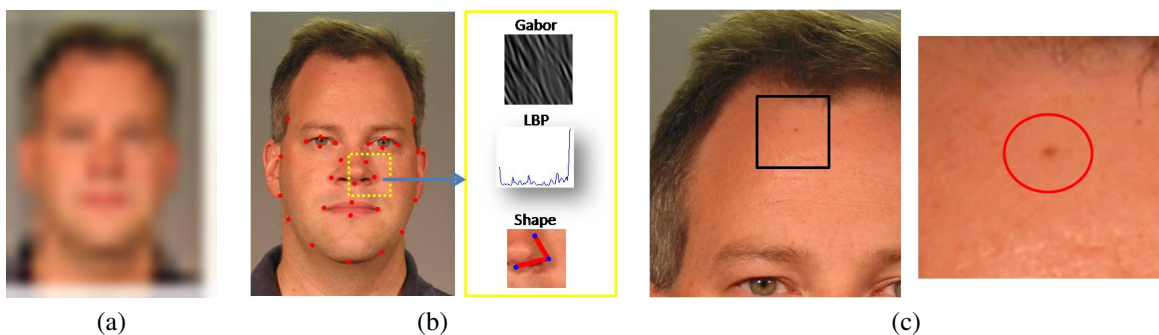


Fig. 1. Examples of the proposed three levels of facial features. (a) Level 1 features contain low dimensional appearance information that is useful for determining ethnicity, gender, and the general shape of a face. Such information is useful for rapidly eliminating match candidates that do not satisfy these attributes. (b) Level 2 features require detailed processing for face recognition. Information regarding the structure and the specific shape and texture of the face is used to make an accurate determination of the subject’s identity. This information extraction requires image processing algorithms for automatic methods, and processing in the fusiform face area of the cerebral cortex (which is believed to be dedicated to face recognition [2], [3]) for human recognition. (c) Level 3 features include marks, moles, scars, and other irregular micro features of the face. This information is useful in forensic identification and to resolve ambiguities among identical twins. Images used with permission from FRGC 2.0 [4]

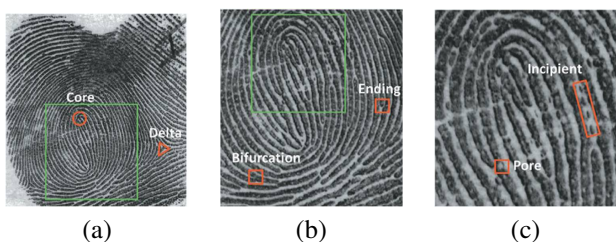


Fig. 2. A fingerprint image and its (a) Level 1, (b) Level 2, and (c) Level 3 features. Similar feature levels have been successful in organizing the information useful for matching fingerprints, facilitating studies on the individuality of fingerprint features.

A. Level 1

Level 1 facial features encompass the global nature of the face, and can be extracted from low resolution face images (< 30 interpupillary pixel distance (IPD)). In automated face recognition, Level 1 features include appearance-based methods such as PCA (Eigenfaces [10]) and LDA (Fisherfaces [11]). For example, these features can generally discriminate between: (i) a short round face and an elongated thin face; (ii) faces possessing predominantly male and female characteristics; or (iii) faces from members of different ethnicities. Level 1 features cannot, however, accurately identify an individual over a large population of candidates. This is illustrated in Figure 3, where a query image can easily be differentiated from a subject that has a very different appearance, but cannot be distinguished from a more similar looking subject.

Level 1 facial features derivable by humans and machines are the gender, race, and general age. The postulated feed-forward nature of human face recognition also uses Level 1 features, where the initial layers can quickly discard a match candidate if they have a largely different facial appearance [12].

Level 1 face features are quite analogous to Level 1 fingerprint features. In each of these two traits, Level 1 features are simple to compute even with low resolution images (see Section V). However, Level 1 features alone are

generally only useful for indexing or reducing the search space. Level 1 features should be explicitly leveraged to improve the matching speed by using them in early stages of a cascaded face recognition system.

B. Level 2

Level 2 features are features that are explicit to face recognition, and require more detailed face observations. These features are locally derived and describe structures in the face that are only relevant in face recognition (as opposed to general object recognition) due to their spatial uniqueness. Examples of such face features in automated face recognition include the use of Gabor wavelets in elastic bunch graph matching (EBGM) [13], local binary patterns (LBP) [14], SIFT feature descriptors [15], [16], point distribution models [17], texture appearance models [17], biologically inspired R&P features [18], [16], and explicit face geometry [19] (which includes the Bertillon system [20]).

Level 2 features are essential for face recognition. There is strong evidence that suggests face recognition activity in humans takes place in the fusiform face area [2], [3], which is a cortical region that appears to be dedicated to face recognition. In an attempt to replicate human visual processing for face recognition, the use of Level 2 biologically inspired features in the form of Gabor wavelets have been successfully utilized in machine face recognition [16]. Along with other features such as the local binary patterns and gradient-based methods, these features are face specific provided they are defined with respect to their spatial coordinates on the face. For example, EBGM extracts Gabor descriptors at specified locations of the face [13], and LBP and SIFT-descriptor methods extract these descriptors at uniformly distributed locations on a face that has been normalized using the eye coordinates [14], [15].

It is important to note that the proposed feature taxonomy does not consider how face recognition is performed using these features. While nearest neighbor classification is the most common method, novel methods to classify face images are often proposed. For example, Kumar et al. proposed

TABLE I
EXAMPLE FEATURES FROM EACH OF THE THREE DIFFERENT LEVELS OF FACIAL FEATURES

		<i>Source</i>
		Machine Only
Level 1	Humans and machine gender, race, age	appearance-based methods (PCA, LDA, etc.)
Level 2	anthropometric features	distribution-based feature descriptors (LBP, SIFT, etc.), shape distribution models, texture descriptors
Level 3	moles, scars, freckles, birth marks	high spatial frequency

using “simile” and “attribute” classifiers for face recognition [21]. While this method for face recognition ultimately makes a decision on a subject’s identity using a series of non-thresholded SVM outputs, the SVM operates on image gradients and normalized pixel features (Level 2 features).

Level 2 face features are analogous to minutiae location and orientation in fingerprint recognition. In both face and fingerprint, the Level 2 features are defined with respect to a particular spatial coordinate reference, and in each case the local features can generally be computed independently of one another.

While Level 2 features are the most discriminative face features, and are predominantly used for face recognition, certain matching scenarios exist in which they alone are not sufficient. One example is face recognition in monozygotic twins [22] (i.e. identical twins). Because the facial appearance of monozygotic twins is nearly identical at medium resolutions (roughly 20 to 100 IPD), Level 2 features alone are generally not sufficient for such a task. Another example where Level 2 features alone may be insufficient is age-invariant face recognition [23]. As humans age, the bone structure (in early aging) and cartilage (in late aging) of the face expands and the skin wrinkles, causing both the facial shape and texture to change.

While humans extract “biological features” to recognize faces, we are limited in our knowledge of how to precisely describe these features. As a result, expert testimony for face recognition in the legal system is generally restricted to the geometric Level 2 features, such as face measurements and ratios (e.g. the ratio of the distance between the eyes and the nose width). These anthropomorphic methods were applied to systematic face recognition, prior to the advent of fingerprint identification, in the Bertillon system [25], [20]. While the uniqueness of such anthropometric features is not currently leveraged in face examination, anthropometric features: (i) have gained informal acceptance in the legal system, and (ii) are computable by both humans and machines [17]. Thus, despite the fact that anthropometric-based face recognition (i) is not a typical approach for automated face recognition, and (ii) is currently used without a consistent and proven methodology in court cases [6], *a thorough examination of their uniqueness must be undertaken*. Such a study would be guided by similar statistical studies on the uniqueness of fingerprints [7] that are critical for the acceptance of fingerprint evidence in the legal system. The

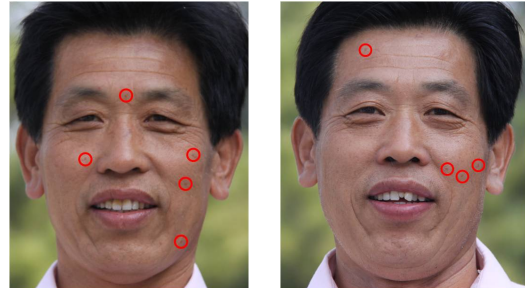


Fig. 4. Face images of two identical twins. While the Level 1 and Level 2 features are the same, it is seen here the facial mark information contained in the Level 3 features offer discriminating information.

following quotation from Spaun [5] speaks specifically to this need:

A major challenge in identifying people by their facial morphology and geometry is the lack of statistical analysis available to document the percentage of uniqueness. The lack of specific statistics restricts the facial image examiner within the justice system to opinion based testimony.

C. Level 3

Level 3 features contain unstructured, micro level features on the face, which includes scars and facial marks. Only recently has this identifiable information been explicitly considered for face recognition [26]. One challenging face recognition problem where Level 3 features are critical is discrimination of monozygotic twins. Because identical twins are extremely difficult for even humans to distinguish, the presence of any small identifying mark could be the difference between successful and mistaken identification. Research in the medical community has shown that while the number of moles (or nevus) in monozygotic twins is correlated, the locations of these moles are not [27] (see Figure 4). Level 3 features have been shown to also improve the matching accuracy in standard face recognition scenarios [26].

Level 3 features in the form of marks should be easy to extract by both humans and computers. Given a good quality face image, the presence of freckles, moles, marks, and scars can be manually marked. An automated approach to mark extraction is also viable [26], though more attention is needed to develop robust solutions. For high resolution images (> 100 IPD) machines are also able to extract

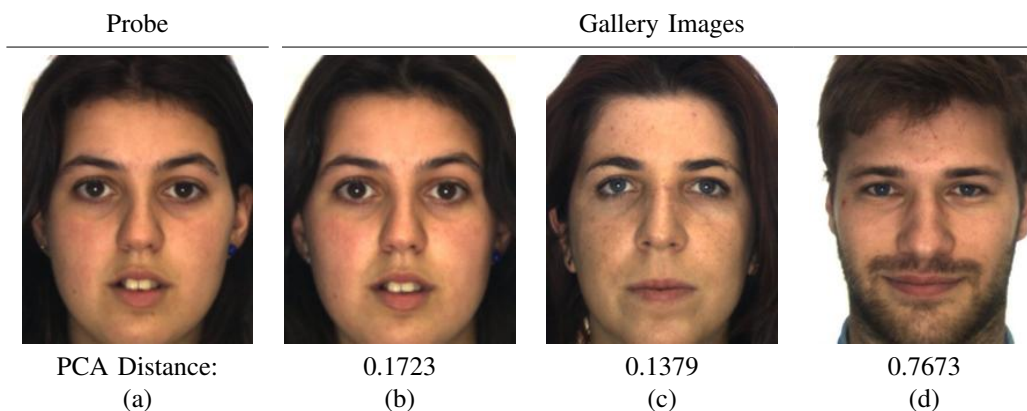


Fig. 3. An example of how Level 1 features can easily filter out faces that exhibit large differences, but cannot distinguish faces that possess many similarities. Above, the probe image (a) was matched using a 500 dimensional PCA representation computed using over 14,000 images, with the (normalized) PCA L_2 distances listed below each gallery image. Using this Level 1 representation, the face in (a) matched well to an image of a similar looking subject (c) than its true mate (b), but was easily differentiated from other subjects that looked largely different (d). The information in Level 1 features is sufficient for quickly discarding some subjects (d), but more advanced Level 2 features are needed to discriminate between similar looking subjects (c). These images are from the AR face database [24].

micro texture information, though very few studies have been conducted to explicitly understand how micro texture analysis can improve face recognition. Results from the 2006 Face Recognition Vendor Test [28] demonstrated that high resolution face images are able to improve the matching accuracy of most commercial matchers, thus implying the usefulness of micro texture information.

In fingerprints, Level 3 features include micro information such as incipient ridges and pores, and irregular information such as scars, creases and other permanent details [29]. This information is typically used by latent fingerprint examiners. In the case of AFIS matching, higher resolution fingerprint images (1000 ppi) are necessary to use pore and ridge information to improve the matching accuracy, which is generally consistent with the proposed Level 3 face features: many moles and facial marks are not detectable at lower image resolutions. In the context of latent examination, the partial fingerprints available may require the use of Level 3 features to make a reliable determination of a subject's identity since there may not be a sufficient number of Level 2 features (minutiae) available. Similarly, forensic examination of face images may need to leverage face mark information to make a successful identify determination [30].

III. RELATIONSHIP OF FEATURE LEVELS TO HUMAN FACE RECOGNITION

There is a strong evidence to suggest that human face recognition is a largely holistic process [31]. In the proposed feature levels, the Level 1 features are holistic in nature but provide weak discrimination, and the Level 2 features are computed locally but are more discriminative. However, this feature scheme is still in agreement with the holistic human face processing model.

In human face recognition, it is believed that we do not operate directly on the reflectance of light from a face (i.e. the pixel intensity values). Instead, the primate visual system extracts visual features that are in turn processed holistically

[18], [32]. In biologically inspired machine object recognition systems, Gabor filters have been used to simulate these human visual features [32], [33]. Similar texture features such as local binary patterns and histograms of orientation gradients serve the same function, but they are more in tune with the discrete nature of digital images. Thus, though the various Level 2 features are generally computed locally across the face, their use in matching is generally in a holistic manner. After the local Level 2 features have been extracted from a face image, they can in turn be processed holistically. Such a decision on how Level 2 features should be used to identify or verify a subject are beyond the scope of the proposed feature taxonomy.

The feedforward nature of human face recognition [12], [32] agrees with the relationship among Level 1 and Level 2 facial features. In the early stages of processing, coarse reflectance information can be used to discard certain matches confidently. Later stages of recognition will extract filter responses similar to Gabor filters. For digital images, methods such as local binary patterns serve the same functionality as continuous sinusoidal functions, but in a manner more attuned to discrete signals.

IV. STUDIES ON FEATURE INDIVIDUALITY

Once a set of facial features has been agreed upon, the next step will be to conduct an analysis of the individuality of these facial features. As mentioned previously, such studies on the likelihood of observing a given facial feature are critical for legal proceedings and forensic examinations. Further, with a standardized set of facial features, commercial face matchers could output confidence measures based on these features and their uniqueness.

In this section, we discuss uniqueness studies that can be conducted with the same motivation as fingerprint studies [7]. These studies will generally require the assumption of fixed pose, illumination, expression, image quality, and other factors that cause degradation in face recognition performance. More specifically, these studies should be conducted on face

TABLE II
LIKELIHOOD OF OBSERVING CERTAIN DEMOGRAPHIC INFORMATION
(LEVEL 1 FEATURES) IN THE UNITED STATES [35].

Race	Gender		Age	
	Male	Female	15-64	< 64
Caucasian	75.0%	49.2%	0-14	20.1%
African American	12.4%	50.8%	15-64	66.9%
Asian	4.4%		< 64	13.0%
Other	4.9%			

images that adhere to the ISO/IEC 19794-5 standard for the interchange of biometric face images [34]. Division of Motor Vehicles (DMV's) in all the states and law enforcement mug shot systems are under increased pressure to adhere to these guidelines.

A. Level 1 Uniqueness

Most of the information regarding the likelihood of a person possessing some combination of human-derivable Level 1 facial features is already available. For example, Table II lists publicly available demographic information for the United States [35]. Using this information, we can calculate that the likelihood of a Caucasian Male, Age 15-64 is $0.75 * 0.492 * 0.669 = 0.247$. Explicitly using these soft biometric, Level 1 features have been shown to improve face recognition [26].

When appearance-based Level 1 features are computed using eigen decomposition, the individuality of a face in the eigenspace is dependent on the distance threshold τ and the diagonal matrix of eigenvalues $D \in \mathbb{R}^{d,d}$. For a particular point x (which corresponds to a face image projected into the eigenspace), its likelihood is computed as

$$P(x) = \frac{1}{(2\pi)^{k/2} |D|^{1/2}} e^{-\frac{1}{2} x^T D^{-1} x} \quad (1)$$

To determine the likelihood of a false association occurring with the template x , the integral of the sphere with radius τ centered at x must be computed over the density in Eq. 1. This uniqueness measure relies on an empirical selection of τ , where τ is generally chosen in order to satisfy a particular true accept rate. For example, in one widely accepted fingerprint individuality model, the distance threshold for corresponding minutiae was chosen such that a 0.975 true accept rate was achieved [7].

Using the proposed Level 1 uniqueness model it is observed that the uniqueness of a face is not uniform. Instead, depending on where a template face resides in the face space, the likelihood of a false association changes in proportion with the Gaussian density in Eq. 1. More specifically, false associations are more likely to occur with faces that reside closer to the origin.

B. Level 2 Uniqueness

Studies of individuality (or uniqueness) of Level 2 features are the most important [5]. This is especially the case for the Level 2 features that are computable by both humans and

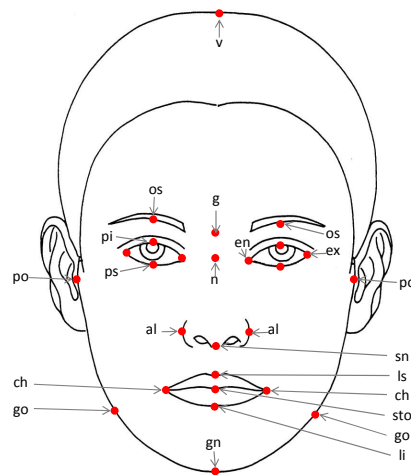


Fig. 5. Anthropometric landmarks that are extractable from frontal, 2D face images. From these 26 landmarks, a collection of 31 facial ratios have been found in medical literature that contain statistical data of their variation with respect to a person's age and sex [36]. Future work involves studying the uniqueness of these ratios (along with other Level 2 facial features). These anthropometric ratios are of particular interest because they can be generated by humans and explained in legal courts.

computers. Without such an analysis, prosecutors and expert testimony cannot provide judges and juries any quantitative information regarding the likelihood that another person could exhibit the same set of facial features.

Figure 5 shows the location of 26 anthropometric facial landmarks found in medical literature that can be reliably located in 2D, frontal face images. From these 26 facial landmarks, 31 ratios have been found that contains mean and standard deviations across age and gender [36]. These statistics have been previously used by orthoplastic surgeons, however they will also improve our future studies on the uniqueness of the face using anthropometry. Again, while these anthropometric features have not received much attention in automated face recognition, they (i) can be generated by humans, and (ii) have precedent in criminal courts.

Similar studies should be conducted on the uniqueness of machine computable features. Observing the distribution of local binary pattern descriptors, SIFT descriptors, and Gabor wavelet responses (for example) computed at specified locations on the face (e.g. center of the eye, corner of the mouth, left nostril, etc.) could augment automated face recognition methods by providing a prior distribution for each feature response. Such studies may validate the use of machine computed features for legal testimony as well.

C. Level 3 Uniqueness

Statistical analysis of the individuality of facial marks involves both the categorization of the types of facial marks present (e.g. moles, freckles, birth marks, etc.), and the observation of their distribution across the face. Park and Jain [26] defined a set of face marks and manually labeled them in a set of 1,225 face images. The distribution of these marks was higher in certain face regions than others, indicating that certain face marks are more unique than

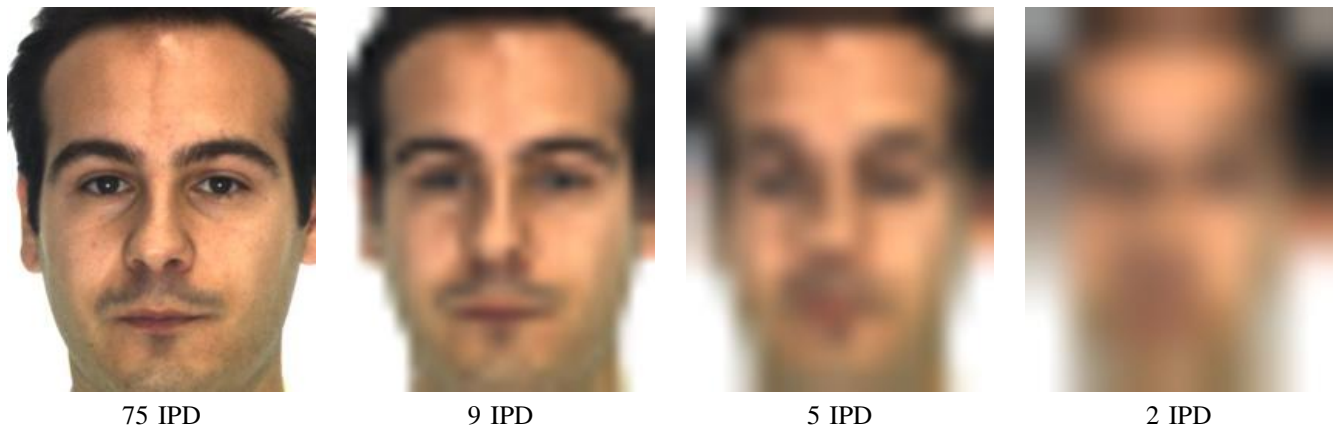


Fig. 6. An example of a 75 IPD face image downsampled to lower IPDs, and interpolated back to the original size. When using the appearance-based Level 1 features, our experiments demonstrated only minor changes in accuracy as the resolution was reduced to as low as 5 IPD. This supports the proposed Level 1 features for use in the early stage of a cascaded recognition model (i.e. when using a false accept rate in the order of 10.0%). At low false accept rates, the Level 2 feature accuracies all deteriorated by 9 IPD.

others. Similar studies should be conducted across larger datasets. Again, such a uniqueness study is important to establish the individuality of a given set of facial marks on a subject’s face.

Further studies on all three feature levels should involve their persistence over time. Such studies on the persistence and stability of each facial feature will further substantiate or counter claims made about a person’s identity with respect to age discrepancies. For example, such studies may indicate that measurements of the nose are unreliable beyond 10 years due to cartilage growth.

V. EXPERIMENTS

We will now attempt to experimentally validate the claim that appearance-based methods in face recognition belong as Level 1 features, and local image descriptors belong as Level 2 features. Level 1 features are to be used in early recognition stages to rapidly eliminate match candidates that have a grossly different appearance and can be extracted from low resolution images. Level 2 features are used for late stage face recognition and must extract more detailed image information. Thus, Level 1 features should have stable accuracies using low resolution images while operating at a high false accept rate (FAR). On the other hand, Level 2 features will require higher resolution face images, and must demonstrate high accuracy at lower false accept rates.

Our test dataset consisted of 876 subjects, each with one mated probe and gallery image, where: 117 subjects were from the AR dataset [24], 294 subjects were from the XM2VTS dataset [37], 193 subjects from the FERET dataset [38], and 272 subjects were from a private data set collected at the University of Notre Dame. The number of subjects is slightly less than the original datasets because certain subjects from each dataset were removed that failed to enroll with a commercial face recognition system. Each image was initially normalized by fixing the interpupillary pixel distance (IPD) to 75 pixels, performing an in plane rotation to set the angle between the eyes to 0° , and cropping the image to

250x200 pixels. Figure 3 shows an example of face images after this normalization process. After normalizing each face image, we reduced the IPD by a factor of 2^k , $k = 1 \dots 5$ by subsampling the face image by a factor of 2^k . After the resolution had been lowered by a factor of 2^k we perform bicubic interpolation to convert the image to its original resolution (see Figure 6). At each scale we test for statistically significant changes in the true accept rate at false accept rates of 10.0%, and 1.0%. Mean and standard deviation statistics of the accuracies of the different feature levels at each IPD were generated using 1,000 bootstrap samples from the test subjects.

Level 1 features were represented using PCA and LDA features. The PCA and LDA projections were computed using the CSU Face Identification Evaluation System [39]. These subspaces were trained using 14,604 face images from 400 subjects (~ 35 images/subject) from the FRGC 2.0 dataset [4].

We used local binary pattern feature descriptors and SIFT descriptors as Level 2 features. Both descriptors were extracted using window sizes of 32 pixels from uniformly spaced locations on the normalized face images. The Chi-Squared dissimilarity metric was used to generate the match scores, as originally proposed in [14].

Figure 7 compares the true accept rates of the Level 1 features and the Level 2 features as a function of the image resolution. In Figure 7(a), it is seen that at a high FAR (10.0%), no significant change in the true accept rate (TAR) occurs for the when the IPD is as low as 5 pixels for the LDA features, and 9 pixels for the PCA features. Figure 6 shows an example of the face images interpolated back to their original size after being lowered to different IPDs. For the high FAR of 10.0%, when the IPD was lowered to 2 pixels the TAR of both Level 1 features was only reduced by less than 10.0%. Conversely, the Level 2 features exhibit a major decrease in accuracy at this resolution.

Figure 7(b) shows the TAR at a lower FAR of 1.0%. This scenario simulates the later stages of a cascaded recognition

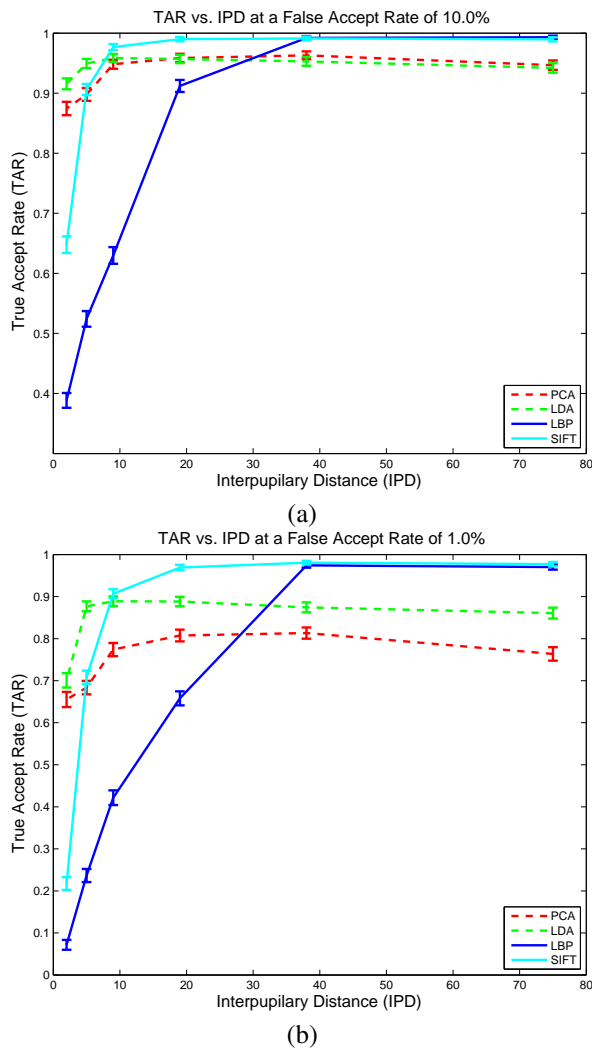


Fig. 7. Recognition accuracy of the Level 1 appearance-based features PCA and LDA (dashed lines), and the Level 2 descriptors LBP and SIFT (solid lines) as a function of the image resolution. The results above are an example of the effectiveness of Level 1 features in the early stages of recognition (high FAR rate), even with very low resolution images. Later stages of recognition (low FAR rate) need more detailed Level 2 features. This model in agreement with a widely accepted model for human face recognition [12].

model. The Level 2 LBP features are observed to offer high performance in this late stage recognition scenario when operating on acceptable quality face images (IPD > 15 pixels). The difference in accuracy between the Level 1 and Level 2 features are more pronounced when operating at this FAR, though the Level 1 features are still robust to decreases in image resolution.

The benefit of a cascaded approach to face recognition is that the Level 1 features are of much lower dimensionality than the Level 2 features. Discarding portions of the gallery by only computing similarity measures using the smaller number of Level 1 features, instead of the larger number of Level 2 features, improves matching speeds and follows the proposed feed forward model of human recognition [12]. For example, our experiments used a roughly 500-dimensional Level 1 PCA representation, whereas the Level 2 LBP feature

TABLE III

THE TRUE ACCEPT RATES (%) FOR RECOGNITION USING ONLY LEVEL 2 FEATURES, AND A TWO-STAGE CASCADED MATCHER.

	FAR = 1.0%	FAR = 0.1%
LBP	97.51 ± 0.34	95.11 ± 0.46
PCA → LBP	97.33 ± 0.34	95.01 ± 0.48
LDA → LBP	96.68 ± 0.44	94.42 ± 0.53
SIFT	98.02 ± 0.26	94.61 ± 0.66
PCA → SIFT	97.81 ± 0.28	94.51 ± 0.71
LDA → SIFT	97.16 ± 0.28	94.22 ± 0.70

Note: $X \rightarrow Y$ indicates using X features in the first stage and Y features in the second stage.

contain 9,086 components.

To demonstrate the benefit of cascading Level 1 and Level 2 features, an additional experiment was conducted. In this experiment, 219 of the 876 test subjects were used for secondary training to determine the threshold τ that achieved FAR of 50.0% using the Level 1 features. The remaining 657 subjects were used to generate the results. The second stage classifier, which used Level 2 features, only considered gallery subjects whose distance was less than τ in the first cascade stage. 1,000 bootstrap samples were used to generate the mean and standard deviation accuracies.

Table III compares the results of (i) matching using only Level 2 features, and (ii) cascaded matching with (PCA or LDA) Level 1 features in the first stage, and (LBP or SIFT) Level 2 features in the second stage. No significant change in the accuracy is noticed using the cascaded approach. Yet, using the cascaded approach, half of the gallery can be matched using feature vectors an order of magnitude smaller than the non-cascaded approach.

VI. CONCLUSIONS

This paper proposed a taxonomy of facial features that: (i) may be computed by both humans and machines or machines only, (ii) follows the same structure as features defined for fingerprints, and (most importantly) (iii) facilitates future studies on the uniqueness and individuality of facial features. Our organization of facial features is shown to be in agreement with the generally accepted models of human face recognition [12]: specifically, face recognition is performed in a feed forward model, that uses coarse appearance information in the initial layers (Level 1 features), and utilizes facial features similar to filter responses (Level 2 features) in subsequent layers.

Experiments were conducted to support the agreement of the feature taxonomy with a well studied model of human face recognition. These experiments demonstrated the effectiveness of Level 1 features in the early stages of recognition, and Level 2 features for the final stages of recognition.

Our objectives for proposing this taxonomy is to open a dialog regarding a precise organization of facial feature specifications. Once a standardization of such features has been accepted, studies on the individuality of these features

can be conducted. Such an analysis will offer the same benefit as similar ongoing studies on fingerprint individuality.

REFERENCES

- [1] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, "Handbook of fingerprint recognition." Springer Professional Computing, 2009, pp. 39–40.
- [2] D. Y. Tsao, W. A. Freiwald, R. B. Tootell, and M. S. Livingstone, "A cortical region consisting entirely of face-selective cells," *Science*, vol. 311, no. 5761, pp. 670–674, Feb 2006.
- [3] N. Kanwisher, J. McDermott, and M. M. Chun, "The fusiform face area: a module in human extrastriate cortex specialized for face perception," *Journal of Neuroscience*, vol. 17, no. 11, p. 43024311, 1997.
- [4] P. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *Proc. of IEEE Conference on Computer Vision & Pattern Recognition*, 2005.
- [5] N. Spaun, "Facial comparisons by subject matter experts: Their role in biometrics and their training," in *Int. Conference on Advances in Biometrics*, 2009.
- [6] G. Edmond, K. Biber, R. Kemp, and G. Porter, "Laws looking glass: Expert identification evidence derived from photographic and video images," *Current Issues in Criminal Justice*, vol. 20, no. 3, pp. 337–377, 2009.
- [7] S. Pankanti, S. Prabhakar, and A. K. Jain, "On the individuality of fingerprints," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, pp. 1010–1025, 2002.
- [8] N. R. Council, *Strengthening Forensic Science in the United States: A Path Forward*. National Academies Press, 2009.
- [9] M. P. Evison and R. W. Vorder Bruegge, Eds., *Computer-aided forensic facial comparison*. CRC Press, 2010, ch. 1,10.
- [10] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.
- [11] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, pp. 711–720, 1997.
- [12] V. Bruce and A. Young, "Understanding face recognition," *British Journal of Psychology*, vol. 77, no. 3, 1986.
- [13] L. Wiskott, J.-M. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," in *Proc. IEEE Int. Conference on Image Processing*, vol. 1, p. 129, 1997.
- [14] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
- [15] B. Klare, Z. Li, and A. K. Jain, "Matching forensic sketches to mugshot photos," *IEEE Trans. Pattern Analysis and Machine Intelligence (to appear)*, 2010.
- [16] E. Meyers and L. Wolf, "Using biologically inspired features for face processing," *Int. Journal of Computer Vision*, vol. 76, no. 1, pp. 93–104, 2008.
- [17] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.
- [18] M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," *Nature Neuroscience*, vol. 2, no. 11, p. 10191025, 1999.
- [19] T. Sakai, M. Nagao, and T. Kanade, "Computer analysis and classification of photographs of human faces," in *Proc. First USA-JAPAN Computer Conference*, 1972, pp. 55–62.
- [20] A. Bertillon, *The Bertillon System of Identification*, R. McClaghry, Ed., Chicago, IL, 1896.
- [21] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and simile classifiers for face verification," in *Proc. IEEE Int. Conference on Computer Vision*, 2009.
- [22] Z. Sun, A. Paulino, J. Feng, Z. Chai, T. Tan, and A. K. Jain, "A study of multibiometric traits of identical twins," in *Proc of SPIE, Biometric Technology for Human Identification VII*, 2010.
- [23] U. Park, Y. Tong, and A. Jain, "Age-invariant face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence (to appear)*, vol. 32, no. 5, pp. 947–954, may 2010.
- [24] A. Martinez and R. Benavente, "The AR face database," in *CVC Technical Report 24*, 1998.
- [25] H. T. F. Rhodes, *Alphonse Bertillon, Father of Scientific Detection*. Abelard-Schuman, New York, 1956.
- [26] U. Park and A. K. Jain, "Face matching and retrieval using soft biometrics," *IEEE Trans. on Information Forensics and Security (TIFS) (To Appear)*, 2010.
- [27] G. Z. et al., "A major quantitative-trait locus for mole density is linked to the familial melanoma gene cdkn2a: A maximum-likelihood combined linkage and association analysis in twins and their sibs," *The American Journal of Human Genetics*, vol. 65, no. 2, pp. 483–492, 1999.
- [28] P. Phillips, W. Scruggs, A. OToole, P. Flynn, K. Bowyer, C. Schott, and M. Sharpe, "Frvt 2006 and ice 2006 large-scale results," in *NISTIR 7408*, 2007.
- [29] A. K. Jain, Y. Chen, and M. Demirkus, "Pores and ridges: High-resolution fingerprint matching using level 3 features," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 15–27, 2007.
- [30] N. A. Spaun, "Forensic biometrics from images and video at the federal bureau of investigation," in *Int. Conference on Biometrics: Theory, Applications and Systems*, 2007, p. 13.
- [31] P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell, "Face recognition by humans: Nineteen results all computer vision researchers should know about," *Proc. IEEE*, vol. 94, no. 11, pp. 1948–1962, 2006.
- [32] E. Meyers and L. Wolf, "Using biologically inspired features for face processing," *Int. Journal of Computer Vision*, vol. 76, no. 1, pp. 93–104, 2008.
- [33] T. Serre, L. Wolf, and T. Poggio, "Object recognition with features inspired by visual cortex," in *Proc. of IEEE Conference on Computer Vision & Pattern Recognition*, 2005, pp. 994 – 1000.
- [34] International Organization for Standardization / International Electrotechnical Commission, "ISO/IEC 19794: Information technology biometric data interchange formats. Part 5: Face image data," 2005.
- [35] C. I. Agency, "Cia world factbook." Skyhorse Publishing, 2008.
- [36] L. Farkas and I. Munro, *Anthropometric Facial Proportions in Medicine*. Charles C. Thomas Publisher, 1986.
- [37] K. Messer, J. Matas, J. Kittler, and K. Jonsson, "XM2VTSDB: The extended M2VTS database," in *Proc. of Audio and Video-based Biometric Person Authentication*, 1999.
- [38] P. Phillips, H. Moon, P. Rauss, and S. Rizvi, "The FERET evaluation methodology for face-recognition algorithms," in *Proc. of IEEE Conference on Computer Vision & Pattern Recognition*, 1997.
- [39] J. R. Beveridge, D. Bolme, B. A. Draper, and M. Teixeira, "The csu face identification evaluation system," *Machine Vision and Applications*, vol. 16, pp. 128–138, 2005.