MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Face Recognition: Where We Are and Where To Go From Here

Michael Jones

TR2009-023 June 2009

Abstract

Face recognition has become a very active research field. Despite the existence of commercial face recognition systems, there are still important challenges for further research. The problems of variable lighting, pose, facial expression, aging, and inaccurate alignment continue to cause larger than desired error rates. This paper discusses the current state of the art in face recognition and suggests some promising directions for future research.

Journal of the Institute of Electrical Engineering of Japan

Paper

Face Recognition: Where We Are and Where To Go From Here

Michael J. Jones* Non-member

Face recognition has become a very active research field. Despite the existence of commercial face recognition systems, there are still important challenges for further research. The problems of variable lighting, pose, facial expression, aging, and inaccurate alignment continue to cause larger than desired error rates. This paper discusses the current state of the art in face recognition and suggests some promising directions for future research.

Keywords: Face recognition, access control, surveillance

1. Introduction

This paper will try to assess the current state of the art in face recognition and suggest some promising directions for future research. Assessing the current state of the art is a difficult task. The amount of work going on in the field of face recognition is huge and growing. A Google Scholar search for papers with the words "face recognition" yielded 1,100 papers in 2000 and 3,190 papers in 2007. When the large amount of unpublished work from face recognition companies and industrial laboratories is also considered, the amount of work is overwhelming. The task of trying to summarize all of this work is sure to ignore many important papers. Despite this, if we step back, we can get a good sense of the state of the art and some of the trends in the field to suggest promising directions for future research. This paper is not intended to be a survey of the field. Zhao et al. (35) and Bowyer et al. (5) provide good recent surveys. Instead, we will try to give a sense of the current state of the technology and some of the major problems that require further research.

2. Face Recognition Scenarios

There is no single state-of-the-art face recognition system. The main reason for this is because there are many different face recognition applications that each have different requirements and constraints. These include access control systems, recognition from surveillance cameras, recognition in photo collections, and recognition in consumer devices. The latter includes devices such as cars or televisions that recognize users to (for example) customize user preferences. Each of these applications has a different set of constraints, such as whether special sensors (for instance, 3D sensors) could be used.

Besides the different applications there are also different modes for face recognition. One is *verification*, in which the user discloses his identity (using an ID card, for instance) and the system must verify this identity by comparing a face image just acquired against one stored in a database. This is called 1-to-1 matching. The newly acquired face image is often called a *probe face* or a *query face*. The faces stored in the database are called *gallery faces* or *enrolled faces*. The other basic mode is *identification*. In this mode, no identity is given, so a newly acquired face image must be compared against every face in the database to determine whether there is a match. This is called 1-to-N matching.

To measure the accuracy of a face recognition system, there are two common methods. One is to plot cumulative match scores, and the other is to plot false acceptance rate versus false rejection rate. To compute the cumulative match score, it is first assumed that each probe face has a corresponding gallery face. Then the question is how often the correct match is in the top k scores. A particular face recognition system is said to have a rank-k recognition rate of $X\%$ if the correct match is within the top k matches $X\%$ of the time. A graph of cumulative match scores plots rank-k recognition rates versus k. The "rank-1 recognition rate" (how often the correct match is the top score) is often just called the "recognition rate".

The other common way to measure recognition accuracy is to compute the false acceptance rate (FAR) and false rejection rate (FRR). This is becoming the more common measure in rigorous tests. It does not make any assumptions about the probe set being a subset of the gallery set. FAR and FRR are computed over a large set of face image pairs. A face image pair consists either of two images of two different people or of two images of the same person (presumably taken at different times or under different conditions). The false acceptance rate is the percentage of face image pairs of different people that are mistakenly classified as the same person. The false rejection rate is the percentage of face image pairs of the same person that are mistakenly classified as different people. The verification rate is equal to 1 - FRR and is the percentage of same face pairs that are correctly classified. Since most face recognition systems return a similarity score for a pair of face images, there is a threshold which determines whether a face pair is classified as same or different. This threshold can be used to trade off between the false acceptance rate and the

* Mitsubishi Electric Research Laboratories

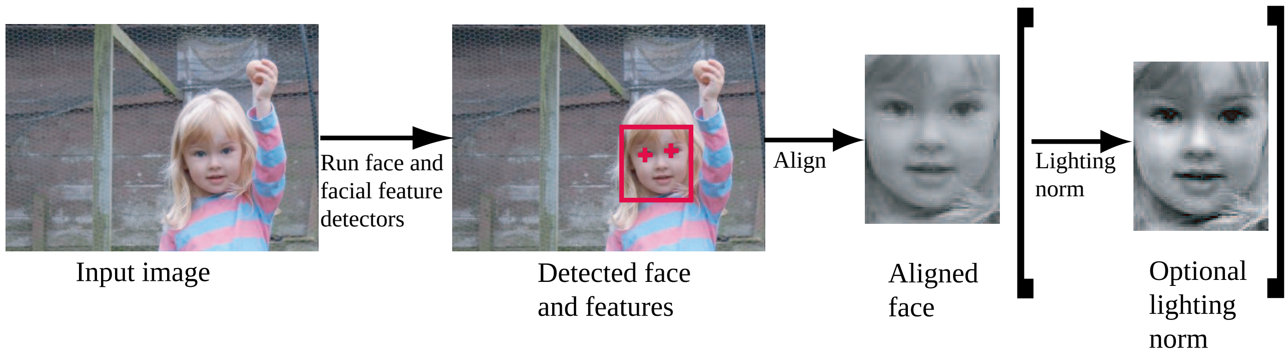

Fig. 1. The preprocessing pipeline

false rejection rate. Plotting the FRR and FAR numbers for all possible thresholds, yields the receiver operating characteristic (ROC) curve.

Both verification and identification share the same basic building block: given two faces images, return a similarity score. To compute a similarity score between two faces, most face recognition systems require a number of preprocessing steps. These are illustrated in Figure 1. The first step is to find any faces in the input image using a face detector. Next, two or more facial features (such as the corners of the eyes) are found in each face; these are used to normalize the face for in-plane rotation, scale, and translation. Cropped face images are usually created. The cropped faces may be lighting normalized to help reduce the effect of lighting variations. Finally, many systems compute some kind of feature vector given the normalized face, and these feature vectors are compared to yield a similarity score.

Even if we consider only a particular application, there are dozens of different face recognition algorithms (if not more), and most of these have not been tested on the same test sets. It is meaningless to state the accuracy of a face recognizer without also stating the test set used for evaluation. This is because the details of the test set, such as how much variation in pose or illumination was included, have a very large effect on accuracy. As a result it is very difficult to get a handle on the state of the art. Nonetheless most researchers would agree that for the simplest scenario of a cooperative access control system with controlled lighting, the problem is basically solved. That is, there are many face recognition algorithms that perform very well. For example, in the recent Face Recognition Grand Challenge (FRGC) test conducted by the US National Institute of Standards and Technology [(25)], one of the scenarios tested was that of cooperative subjects in a controlled lighting environment (although small expression variations of smiling and neutral were present). The median verification rate of the 17 algorithms that were tested was 91% at 0.1% false acceptance rate. The best system reported 99% verification rate at 0.1% false acceptance rate.

The problem is that this is not a realistic scenario for any application. Even for the simplest application, that of an access control system in which subjects are required to cooperate, there is no guarantee that it can be installed in a location with completely controlled illumination. Nearby outside windows may cause large changes in illumination, and even typical overhead office lights cause variable illumination as a subject moves relative to the lights. Furthermore, variations in pose of the same user from one use of the system to the next cannot be avoided. Finally, if the system is expected to work over a period of many years, then aging is also an issue.

This leads us to the fundamental problems confronting face recognition systems today: lighting, pose, facial expression, aging, and alignment. The lighting problem comes from the fact that different illuminations can cause very large changes to the image of a person's face. Similarly, changes in head pose, facial expression, and age can also lead to very different face images for the same person. Furthermore, all face recognition systems need to align the input face in some way. For a 2D face recognition system, this is typically done by finding two or more facial feature points to do the alignment normalization shown in Figure 1. If the facial feature points are not found accurately, this can cause large changes to the output face image. All of these potential variations to the face image make the face recognition problem difficult. Any new research in face recognition should address at least one of these problems in order to make a real contribution.

Next, we will examine current approaches to each of these problems in order to get a sense for where we are and where we might go from here.

3. Illumination

Comparing two faces with different illuminations is one of the fundamental problems for face recognition systems. It is well known that the differences in pixel values caused by different lighting conditions can be greater than the differences between two different people under the same lighting condition [(1)]. See Figure 2 for examples of how a face can vary under different illuminations. In their seminal work, Belhumeur and Kriegman [(3)] developed a theoretical answer to the question of what set of images an object (face) can realize under all possible illuminations (see also (13) (14) (29)). They showed that it is a convex cone in $\mathcal{R}^n$ where n is the number of pixels.

Fig. 2. Examples of a face under various illuminations

Fig. 3. Examples of a face in various poses

They coined the term *illumination cone* to describe this set of images. Under a Lambertian reflectance assumption and assuming a convex object, they showed that it can be computed from as few as three images. Under more realistic assumptions (non-Lambertian reflectance, non-convex shape) they presented evidence that the illumination cone lies near a low-dimensional subspace. The implication is that under fixed pose, a face under any illumination is well approximated by a convex combination of a small set of basis face images. In practice, the problem with using this important result is that multiple illuminations of the same face in the same pose may not be available so that the illumination cone cannot be computed. Subsequent work has tried to address these problems. For example, Basri and Jacobs [(2)] analyzed the properties of the illumination cone and showed analytically how to compute a low dimensional subspace to approximate the illumination cone. The basis functions of the low dimensional subspace are represented using spherical harmonics, which is the analogue to Fourier analysis on the surface of a sphere.

Another issue with the illumination cone model is that to handle non-convex objects which have cast shadows (shadows cast by a nose for instance), the model must be explicitely extended to accurately handle these. Georghiades et al. [(13)] used ray tracing along with the 3D surface computed during the computation of the illumination cone to correctly model cast shadows. This makes the framework much more computationally expensive. More computationally efficient ways of handling cast shadows are desirable.

There are many other important schemes for dealing with illumination changes which there is not enough space to discuss, but the illumination cone and related methods are the most fundamental.

Although the theory of how faces change due to changing illumination is well developed, converting this theory to practical systems still requires innovation. It also depends on the application. In applications where extra effort can be spent during enrollment, a different solution may be feasible than in applications where a single image is all that is available for enrollment. Also, making the solution to varying illumination work across changing pose is another important problem to solve.

4. Pose

Comparing faces under varying pose is another fundamental challenge for face recognition systems. Figure 3 shows how a face image can change under modest pose variation. Unlike the case of illumination variation, it does not appear that all images of a face under varying pose can be so neatly characterized. This does not necessarily mean that images of a face under varying pose do not lie on a relatively low-dimensional manifold in image space. However, this manifold does not seem to submit to a simple mathematical characterization.

Solutions to the pose problem depend greatly on the particular application. One approach is view-based in which a set of pose-specific classifiers or similarity measures is used to compare faces in a similar pose. Pentland et al. [(23)] takes this approach, for example. The problem with this approach is that is requires multiple views of each gallery face and it requires good pose estimation of the query (probe) face to know which pose-specific classifier and gallery image to use for comparison.

Another approach is to use a 2D model, such as a morphable model [(16)] or an active appearance model [(8)]. These models, which use very similar formulations, both separate a set of example images (such as faces varying in pose) into texture and shape components. To do this, all of the example images must be put into correspondence with a reference image. Then the variations in shape (positions of the facial features) and the variations in texture (brightness values of the pixels) are represented as linear combinations of basis shapes and textures. These basis shapes and textures are usually found using principal component analysis. Because these are 2D models, they have difficulties with self occlusions on the face caused by pose changes. However, a modest amount of pose variation can be handled. Given a new input face image, the model is fit to the image using an iterative optimization scheme by finding the shape and texture parameters that yield a model image that most closely matches the input image. The model parameters can be mapped into those that affect mainly pose and those that affect other variations. After matching, the parameters that affect pose can be ignored, and the parameters that affect only identity are used for recognition [10]. Another possibility is to render a new model image after changing the pose parameters to represent a frontal pose and in this way produce a pose-normalized version of the input face for recognition. The main problems with the morphable model and active appearance models are in the matching of the model to the input image. This optimization process can get stuck in bad local minima. The process is only robust if the model can be initialized to a starting point close to the input face. Furthermore, creating a model that includes a large set of different faces with all the variations required (such as illumination and pose) is a tedious process.

More recently, approaches based on 3D face models have gotten a lot of attention for dealing with pose. There are various ways 3D models can be used which again depend heavily on the application scenario. One solution is a 3D-from-2D approach that matches a 3D model to a 2D face image and then pose normalizes by rendering a frontal view of the face [4]. This requires a method of filling in missing texture for parts of the face that were occluded in the non-frontal view. The main difficulty with this approach is robustly and quickly matching the 3D model to the 2D image. This is especially difficult when there are other sources of variability such as lighting and facial expression. Another variation of this approach is not to render a pose-normalized frontal face image, but rather directly compare the fitted 3D models. This, however, still leaves the problem of how to compare 3D face models. We briefly discuss approaches to this problem below.

Yet another approach that has seen a lot of activity recently is to use a 3D sensor to acquire a 3D face. Face recognition is then performed by comparing 3D faces. There are various types of 3D sensors such as stereo cameras, structured light systems and laser rangefinders. One major problem with such systems is that the acquired 3D shape often has significant artifacts such as holes or spikes. The problem can be especially bad under difficult lighting conditions such as when there are large specular regions [5].

There are various ways to compare 3D faces. See (5) for a review. Many techniques involve an iterative closest point algorithm to match 3D shapes or local curvature measurements at different feature points [5]. Most require a technique either for estimating pose or for finding correspondences between two 3D models, which is not a trivial problem.

Of the 3D approaches, only the 3D-from-2D approach is potentially applicable in all of the same application scenarios as 2D recognition. Pure 3D approaches do not apply, for example, to recognition from photo collections or from existing surveillance cameras. For access control scenarios, the added cost of a 3D sensor may be justified by the improved robustness to pose. Results from the Face Recognition Grand Challenge test in 2005 for comparing 3D faces with 3D faces showed very good accuracy, with recognition rates around 97% for the best system. However, this test did not include variations to lighting so, it is hard to extrapolate these results to real-world access control scenarios.

For applications in which a 3D sensor is feasible, 3D face acquistion will most likely lead to a robust solution to the pose problem. The main barrier appears to be more robust 3D sensors that are not so adversely affected by harsh lighting conditions. For other applications in which 3D sensors are not an option, some form of 3D-from-2D approach appears to be the most promising for pose robustness. More research is required on the best way to do this.

Fig. 4. Examples of a face with various expressions

5. Expression

Comparing faces with different facial expressions is another problem for some face recognition applications. For access control systems in which the subject must cooperate, facial expression is not a major issue. For non-cooperative face recognition approaches, however, such as recognizing people in family photos, it is an important problem. See Figure 4 for an example of a face with many different expressions. There are various approaches to dealing with expression that are based on the idea of building a model of the changes that faces undergo through expression variations. Models are learned through examples. For instance, a morphable model (2D or 3D) [4][16] or active appearance model [8] can be used to capture expression variations (as well as identity, lighting, and pose variations). As with pose, the morphable model or active appearance model is fit to the input face, and then either the face can be expression normalized to yield an image of a face under neutral expression (analogous to pose and lighting normalization), or some “expression-free” model parameters are extracted and used directly as the feature vector for recognition. The problem with this approach is that fitting a morphable model or active appearance model is computationally expensive and not very robust, especially in the presence of lighting and pose variations.

Another approach is multilinear models or tensorfaces, which is an extension of singular value decomposition (SVD) to multi-dimensional tensors [19] [32]. Given a set of images for each person that varies over expression as well as possibly over lighting and pose, different subspaces can be learned for each mode of variation. A new face is matched to the tensorface model by projecting the new face into the tensorface space and finding the closest gallery face in that space. The drawback of this approach is that it requires a large set of example images for each person with all of the desired variations.

Some systems take the very simple approach of ignoring the mouth region which is the most variable in the presence of expression changes. This may be adequate for some applications.

6. Aging

The fact that your face changes significantly as you age means that the face is a fundamentally flawed biometric. See Figure 5 for an example of the significant changes a face undergoes as it ages. For other biometrics such as irises and fingerprints, aging is a much less serious problem. Despite this fact, the face is still often prefered as a biometric, mainly because it is so easy to capture.

The changes that a face undergoes as it ages are similar across different people. The morphological changes as well as textural changes (wrinkles) are similar from person to person. This makes it possible to model the effects of aging and incorporate such models into the face recognition system.

Compared to the illumination, pose, and expression problems, there has been relatively little work on making face recognition robust to variations in age. The main reason for this is probably due to the fact that it is difficult to collect datasets of face images that include the same person taken at many different ages. The datasets that do exist also have significant variations in pose and illumination that make it difficult to study age in isolation. However, many of the techniques that can handle the other forms of variation can also handle age variations given the right training data.

In some applications, aging may not be a significant problem because enrolled faces may be periodically updated so that the age difference between enrolled face and query face is never very large. In other applications, however, this may not be possible.

One of the few papers to deal directly with the problem of age in face recognition is the work of Lanitis et al. [18]. They used an active appearance model that included examples of the same person at different ages to model faces. After matching this model to an input face, they mapped the model parameters to a set of age-normalized model parameters using a mapping learned from the training examples. Using their age normalization method they showed a modest amount of improvement in recognition.

Another paper that looked at the problem of age in face recognition is Ramanathan and Chellappa [27]. They used a Bayesian intra/extra model to model age

Fig. 5. Examples of the same person at various ages

variations [20]. They found that, as expected, the similarity score between two images of the same person taken at different ages decreases as the age difference increases. When the age difference was four years or less, the face verification accuracy was good for their test set, but for greater age differences accuracy was poor.

Aging may prove the most difficult variation to deal with because of the difficulty of collecting large amounts of appropriate training data and because the facial changes can be very significant if the age difference is large.

7. Alignment

The alignment problem is a very important practical issue in face recognition systems that is often ignored in conference and journal papers. The problem is to align or geometrically normalize and crop a face in an input image to yield a fixed sized output image with the face in a known position. This is usually done by first detecting a face, finding a number of facial feature points within the detected face region, and then computing a four parameter similarity transform (scale, in-plane rotation, x and y translation) to map the input face to a fixed sized output image. Other transformations such as affine may be used instead, but a similarity transform is usually best because it does not distort the geometry of the face. To solve for the four parameters of the similarity transform, only two facial feature points are needed, but more are often found to improve robustness.

Today there are robust face detection algorithms such as (33) and (30) that can be used for finding frontal, upright faces in cluttered scenes. There are also multi-view face detection algorithms for finding non-frontal or non-upright faces [15] [17] [30]. These algorithms are sufficient for finding faces for most applications, and thus face detection is no longer a major problem for face recognition.

The bounding box found by a face detection algorithm is not in general precise enough to use for alignment. It finds a rough bounding box around the face but is not intended to find the precise scale, rotation, and location of the face. For this reason, more precise facial feature detectors are needed for good alignment.

There are many different algorithms for finding facial features. Some are based on learning a classifier (for example Support Vector Machines (SVM), neural networks, boosted classifiers, or Bayesian classifiers [12]) that detects a single facial feature given an image patch or some features of the image patch (such as a Gabor features [34]). The image patch typically covers some area around the feature. The training examples need to be carefully manually aligned so that the resulting classifier is precise. As with face detection, such feature detectors work best when they are constrained to work over a narrow range of facial poses.

Other algorithms attempt to use the geometry of a group of facial features to constrain the search for features. Burl et al. [7] is an early example of such an algorithm.

Once again, a 2D morphable model [16] or active appearance model [8] could be used to localize facial features, although the robustness of the matching process is still an issue as is the relative computational expense. As a refinement stage after another method has been used for initialization, these methods may be more useful [9].

There has been very little comparative evaluation of various facial feature localization algorithms so it is difficult to determine the best approaches. A detailed comparison of different algorithms would be an important contribution to the research community. However, it is safe to say that there is further room for improvement to design feature detectors that are fast, precise, and robust to lighting, pose, and expression changes.

8. Basic approaches to handling facial variations

As mentioned earlier, there have been thousands of research papers that propose various approaches to making a face recognizer that is robust to at least some of the types of facial variations. After examining the various approaches, we can see that many of them share some basic ideas. One central idea is to build a general model of how faces may vary. This is usually done by learning a model from example face images. Once a model is created, there are three basic ways to use it:

- Match the model to the face and then normalize to get a new face image with a canonical illumination, pose, expression, or age.
- Match the model to the face and then use for recognition only the parameters that model identity changes.
- Use the model to generate synthetic examples of the face under all variations.

Another possibility related to the third approach above is to gather a set of example face images covering the desired range of facial variations for each enrolled person. The set of images for one person that cover all facial variations (whether acquired or synthesized) is called the face space for that person. Once the face space for a person is known, then matching a probe face becomes computing the distance of the probe to the face space. There are different ways that this can be done (see for example (21)).

A face recognition system may use a mix of these approaches for different variations. Also, the input and model may be either 2D or 3D. Clearly, it is not practical to acquire actual images of a face under many poses, illuminations, expressions, and ages for every enrolled person. Thus some kind of general model of how faces vary under different illuminations, poses, etc. is needed. Formulating a model that is easy to build, covers all of the desired variations and can be matched robustly is the difficult challenge.

9. Where To Go From Here

Given the discussion above, it is clear that new work in face recognition should address one or more of the five basic problems: lighting, pose, expression, age and alignment. One way to do this is to devise new models that either are better able to capture the important sources of facial variations or are more efficient or more robust to fit to an input face.

One promising direction for future research is the idea of using an "enhanced" enrollment procedure such as collecting multiple images with lighting, pose, and expression variations for applications that allow it (for example some access control applications and some surveillance camera applications). The motivation for this approach comes from research on human face recognition performance (in other words the ability of people to recognize faces). One of the most striking findings of this research is that people are much better at recognizing faces of familiar people. It appears that people build very good models of familiar faces and so can recognize such faces well even from very low resolution images [6] [31]. A computer algorithm could similarly take advantage of a good model of each face built at enrollment time. A single query image may then be compared to the detailed model for each enrolled face. Some work along these lines has been done (for example, algorithms that compare an enrolled 3D model to a 2D query face image) but there is room for progress here.

Today the resolution of still cameras is very high, and the resolution of video cameras is increasing. Taking advantage of this high resolution information is another promising direction for improving face recognition accuracy. There are various ways that high resolution imagery can be exploited. One straightforward way is to use it to get better facial feature localization and thus better alignment. Another possibility is to use skin markings such as freckles, moles, and scars as unique identifiers. Some work along this line has already been done by Pierrard and Vetter [26], who analyzed moles to recognize faces. It may also be possible to analyze the very fine ridges and pores in skin for computing similarity scores analogous to what is done in fingerprint recognition.

Another promising direction for future research is to use ideas from the rapidly growing area of computational photography. Computational photography attempts to modify the camera hardware in a way that makes it possible to compute images which cannot be acquired from an ordinary digital camera. The basic idea behind these approaches is to modify the camera hardware in order to make the problem easier than it would be using a traditional camera. One possibility is to deal with illumination by rapidly taking multiple images under different illuminations (for example with and without flash) and use these to compute an image under a canonical illumination [11] [24]. Computational photography approaches can also be used to handle motion blur and focus problems [22][28]. For some face recognition applications, modifying the camera hardware used to capture faces is not an option. In applications where it is an option, computational photography has the potential of lessening some of the problems confronting face recognition.

Another trend in face recognition research is the use of 3D sensors and the use of video to improve face recognition accuracy. Both of these ideas have the potential to improve the state of the art for scenarios in which they are applicable.

Finally, another important trend in face recognition research is the use of very difficult test sets. In early work, face recognition algorithms were usually tested on small test sets taken under controlled conditions. While these early systems reported very high recognition rates, they did not perform well in practice under typical difficult conditions. Today, there are some difficult publicly available test sets that are helping to push researchers to tackle the hard problems. For example, the Face Recognition Grand Challenge test set from the US National Institute of Standards and Technology [25] has proven to be a difficult test set especially because of significant variations in illumination. It is also quite large, with over 16,000 images in some of the testing subsets. Another very difficult test set is Labeled Faces in the Wild from the University of Massachusetts Amherst. This test set consists mostly of celebrities gathered from the World Wide Web. The photos include uncontrolled lighting conditions, various poses (although not too far from frontal), various expressions, and occlusions. By focusing on such difficult test sets, researchers will be forced to confront the difficult problems for face recognition.

10. Conclusions

Despite the large amount of work that has been done, face recognition remains an unsolved problem for which there is plenty of room for new ideas. New ideas are needed to improve the robustness of face recognition systems to changes in illumination, pose, expression, and age and to improve the robustness of alignment to these variations. Despite these serious difficulties with using the face as a biometric, it will continue to be a very desirable biometric because of its convenience, especially compared to fingerprints and irises.

Promising directions for future research include techniques for building better models using an enhanced enrollment process, taking advantage of higher resolution images, using ideas from computational photography, using video or 3D sensors and focusing on more difficult test sets. Thus, there are many opportunities for new researchers to contribute new ideas to the field.

References

(1) Y. Adini, Y. Moses and S. Ullman: "Face recognition: The problem of compensating for changes in illumination direction," *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol 19, pp. 721-732, 1997.

(2) R. Basri and D.W. Jacobs: "Lambertian reflectance and linear subspaces," *IEEE Trans. Patt. Anal. and Mach. Intell.*., Vol 25, No. 2, pp. 218-233, 2003.

(3) P.N. Belhumeur and D.J. Kriegman: "What is the Set of Images of an Object under All Possible Illumination Conditions?," *Intl. J. of Computer Vision*, Vol. 28, No. 3, pp. 1-16, 1998.

(4) V. Blanz and T. Vetter: "Face Recognition Based on Fitting a 3D Morphable Model," *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol. 25, No. 9, pp. 1063-1074, 2003.

(5) K.W. Bowyer, K. Chang, P. Flynn: "A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition," itshape Computer Vision and Image Understanding, Vol 101, pp. 1-15, 2006.

(6) V. Bruce, Z. Henderson, K. Greenwood, P.J.B. Hancock, A.M. Burton, P.I. Miller: Verification of face identities from images captured on video," *J. Experimental Psychol.: Applied*, Vol. 5-4, pp. 339-360, 1999.

(7) M. Burl, T. Leung and P. Perona: "Face Localization Via Shape Statistics," *First International Workshop on Automatic Face and Gesture Recognition*, 1995.

(8) T.F. Cootes, G.J. Edwards and C.J. Taylor: "Active Appearance Models," *IEEE Trans. Patt. Anal. and Mach. Intell.*, Vol 23, No. 6, pp. 681-685, 2001.

(9) D. Cristinacci, T. Cootes and I. Scott: "A Multi-Stage Approach to Facial Feature Detection," *Proc. of British Mach. Vision Conf.*, Vol. 1, pp. 277-286, 2004.

(10) G.J. Edwards, A. Lanitis, C.J. Taylor and T.F. Cootes: "Statistical Models of Face Images - Improving Specificity," *Image and Vision Computing*, Vol. 16, No. 3, pp. 203-211, 1998.

(11) E. Eisemann and F. Durand: "Flash Photography Enhancement via Intrinsic Relighting," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, Vol. 23, No. 3, pp. 673-678, 2004.

(12) M. Everingham and A. Zisserman: "Regression and classification approaches to eye localization in face images," *Intl. Conf. on Automatic Face and Gesture Recognition*, pp. 441-448, 2006

(13) A.S. Georghiades, D.J. Kriegman and P.N. Belhumeur: "From Few to Many: Illumination Cone Models for Face Recognition Under Variable Lighting and Pose," *IEEE Trans. Patt. Anal. and Mach. Intell.*, Vol 23, No. 6, pp. 643-660, 2001.

(14) P. Hallinan: "A low-dimensional representation of human faces for arbitrary lighting conditions," *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pp. 995-999, 1994.

(15) C. Huang, H. Ai, Y. Li and S. Lao: "High-Performance Rotation Invariant Multiview Face Detection," *IEEE Trans. Patt. Anal. and Mach. Intell.*, Vol. 29, No. 4, pp. 671-686, 2007.

(16) M.J. Jones and T. Poggio: "Multidimensional Morphable Models: A Framework for Representing and Matching Object Classes," *Intl. J. of Comp. Vis.*, Vol. 29. No. 2, pp. 107-131, 1998.

(17) M.J. Jones and P. Viola: "Fast Multi-view Face Detection," Mitsubishi Electric Research Labs Technical Report TR2003-96, pp. 1-10, 2003.

(18) A. Lanitis, C.J. Taylor and T.F. Cootes: "Toward Automatic Simulation of Aging Effects on Face Images," *IEEE Trans. Patt. Anal. and Mach. Intell.*, Vol. 24, No. 4, pp. 442-455, 2002.

(19) J. Lee, B. Moghaddam, H. Pfister and R. Machiraju: "A Bilinear Illumination Model for Robust Face Recognition," *IEEE Intl. Conf. on Comp. Vis.*, Vol. 2,No. 2, pp. 1177-1184, 2005.

(20) B. Moghaddam and A. Pentland: "Probabilistic Visual Learning for Object Representation," *IEEE Trans. Patt. Anal. and Mach. Intell.*, Vol. 19, No. 7, pp. 696-710, 1997

(21) H. Murase and S. Nayar: "Visual Learning and Recognition of 3-D Objects from Appearance," *Intl. J. of Comp. Vis.*, Vol. 14, pp. 5-24, 1995.

(22) R. Ng, M. Levoy, M. Brdif, G. Duval, M. Horowitz and P. Hanrahan: "Light Field Photography with a Hand-Held Plenoptic Camera," Stanford Tech Report CTSR 2005-02, 2005.

(23) A. Pentland, B. Moghaddam and T. Starner: "View-based and Modular Eigenspaces for Face Recognition," *IEEE Conf. on Comp. Vis. and Patt. Rec.*, pp. 84-91, 1994.

(24) G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe

and K. Toyama: "Digital Photography with Flash and No-flash Image Pairs," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, Vol. 23, No. 3, pp. 664-672, 2004.
(25) P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer and W. Worek: "Preliminary Face Recognition Grand Challenge Results," *Intl. Conf. on Automatic Face and Gesture Recognition*, 2006.
(26) J.S. Pierrard and T. Vetter: "Skin Detail Analysis for Face Recognition," *IEEE Conf. on Comp. Vis. and Patt. Rec.*, pp. 1-7, 2007.
(27) N. Ramanathan and R. Chellappa: "Face Verification across Age Progression," *IEEE Conf. on Comp. Vis. and Patt. Rec.*, Vol. 2, pp. 462-469, 2005.
(28) R. Raskar, A. Agrawal and J. Tumblin: "Coded Exposure Photography: Motion Deblurring using Fluttered Shutter," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, Vol. 25, No. 3, pp. 795-804, 2006.
(29) A. Shashua: "Geometry and Photometry in 3D Visual Recognition," PhD thesis, MIT, 1992.
(30) H. Schneiderman and T. Kanade: "A statistical method for 3D object detection applied to faces and cars," *IEEE Conf. on Comp. Vis. and Patt. Rec.*, Vol 1, pp. 746-751, 2000.
(31) P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell: "Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About," *Proc. of the IEEE*, Vol 19, No. 11, pp. 1948-1962, 2006.
(32) M.A.O. Vasilescu and D. Terzopoulos: "Multilinear Analysis of Image Ensembles: TensorFaces," *Proc. European Conference on Computer Vision*, pp. 447-460, 2002.
(33) P. Viola and M.J. Jones: "Robust Real-Time Face Detection," *Intl. J. of Comp. Vis.*, Vol. 57, No. 2, pp. 137-154, 2004.
(34) L. Wiskott, J.M. Fellous, N. Kuiger and C. von der Malsburg: "Face Recognition by Elastic Bunch Graph Matching," *IEEE Trans. Patt. Anal. and Mach. Intell.*, Vol. 19, No. 7, pp. 775-779, 1997.
(35) W. Zhao, R. Chellappa, P.J. Phillips and A. Rosenfeld: "Face Recognition: A Literature Survey," *ACM Computing Surveys*, Vol. 35, No. 4, pp. 399-459, 2003.

Michael Jones (Non-member) Michael Jones was born in Dallas, Texas on September 14, 1967. He received his bachelors degree in computer science from Rice University in 1990. He received his Ph.D. from the Artificial Intelligence Laboratory at M.I.T. in 1997. He is currently a senior research scientist at Mitsubishi Electric Research Laboratories in Cambridge, Massachusetts where his main research focus has been on face detection and recognition.