Chapter 4

FACE RECOGNITION AND ITS APPLICATIONS

Andrew W. Senior and Ruud M. Bolle
IBM T.J.Watson Research Center,
P.O. Box 704,
Yorktown Heights,
NY 10598, USA.
{aws, bolle}@us.ibm.com

Abstract Face recognition has long been a goal of computer vision, but only in recent years reliable automated face recognition has become a realistic target of biometrics research. New algorithms, and developments spurred by falling costs of cameras and by the increasing availability processing power have led to practical face recognition systems. These systems are increasingly being deployed in a wide range of practical applications, and future improvements promise to spread the use of face recognition further still. In this chapter, we review the field of face recognition, analysing its strengths and weaknesses and describe the applications where the technology is currently being deployed and where it shows future potential. We describe the IBM face recognition system and some of its application domains.

Keywords: Face recognition, robust biometrics

4.1. Introduction

Recognizing faces is something that people usually do effortlessly and without much conscious thought, yet it has remained a difficult problem in the area of computer vision, where some 20 years of research is just beginning to yield useful technological solutions. As a biometric technology, automated face recognition has a number of desirable properties that are driving research into practical techniques.

The problem of face recognition can be stated as ‘identifying an individual from images of the face’ and encompasses a number of variations other than the most familiar application of mug shot identification. One notable aspect of face recognition is the broad interdisciplinary nature of the interest in it:

within computer recognition and pattern recognition; biometrics and security; multimedia processing; psychology and neuroscience. It is a field of research notable for the necessity and the richness of interaction between computer scientists and psychologists.

The automatic recognition of human faces spans a variety of different technologies. At a highest level, the technologies are best distinguished by the input medium that is used, whether visible light, infra-red [29, 31] or 3-dimensional data [7] from stereo or other range-finding technologies. Thus far, the field has concentrated on still, visible-light, photographic images, often black and white, though much interest is now beginning to be shown in the recognition of faces in colour video. Each input medium that is used for face recognition brings robustness to certain conditions, *e.g.* infra-red face imaging is practically invariant to lighting conditions while 3-dimensional data in theory is invariant to head pose. Imaging in the visible light spectrum, however, will remain the preeminent domain for research and application of face recognition because of the vast quantity of legacy data and the ubiquity and cheapness of photographic capture equipment.

4.2. Face as a Biometric

Face recognition (see [6, 33] for recent surveys) has a number of strengths to recommend it over other biometric modalities in certain circumstances, and corresponding weaknesses that make it an inappropriate choice of biometric for other applications. Face recognition as a biometric derives a number of advantages from being the primary biometric that humans use to recognize one another. Some of the earliest identification tokens, *i.e.* portraits, use this biometric as an authentication pattern. Furthermore it is well-accepted and easily understood by people, and it is easy for a human operator to arbitrate machine decisions — in fact face images are often used as a human-verifiable backup to automated fingerprint recognition systems.

Because of its prevalence as an institutionalized and accepted guarantor of identity since the advent of photography, there are large legacy systems based on face images — such as police records, passports and driving licences — that are currently being automated. Video indexing is another example of legacy data for which face recognition, in conjunction with speaker identification [19], is a valuable tool.

Face recognition has the advantage of ubiquity and of being universal over other major biometrics, in that everyone has a face and everyone readily displays the face. (Whereas, for instance, fingerprints are captured with much more difficulty and a significant proportion of the population has fingerprints that can not be captured with quality sufficient for recognition.) Uniqueness, another desirable characteristic for a biometric, is hard to claim at current levels of accuracy. Since face shape, especially when young, is heavily influenced by genotype, identical twins are very hard to tell apart with this technology.

With some configuration and co-ordination of one or more cameras, it is be more or less possible to acquire face images without active participation of the subject. Such passive identification might be desirable for customization of user services and consumer devices, whether that be opening a house door as the owner walks up to it, or adjusting mirrors and car seats to the driver's presets when sitting down in their car.

Surveillance systems rely on passive acquisition by capturing the face image without the cooperation or knowledge of the person being imaged. Face recognition also has the advantage that the acquisition devices are cheap and are becoming a commodity (though this is not true for non-visible wavelength devices and some of the more sophisticated face recognition technologies based on 3-dimensional data).

The main drawbacks to face recognition are its current relatively low accuracy (compared to the proven performance of fingerprint and iris recognition) and the relative ease with which many systems can be defeated (Section 4.2.1). Finally, there are many attributes leading to the variability of images of a single face that add to the complexity of the recognition problem if they can not be avoided by careful design of the capture situation. Inadequate constraint or handling of such variability inevitably leads to failures in recognition.

These include:

- **Physical changes:** facial expression change; aging; personal appearance (make-up, glasses, facial hair, hairstyle, disguise).
- **Acquisition geometry changes:** change in scale, location and in-plane rotation of the face (facing the camera) as well as rotation in depth (facing the camera obliquely, or presentation of a profile, not full-frontal face).
- **Imaging changes:** lighting variation; camera variations; channel characteristics (especially in broadcast, or compressed images).

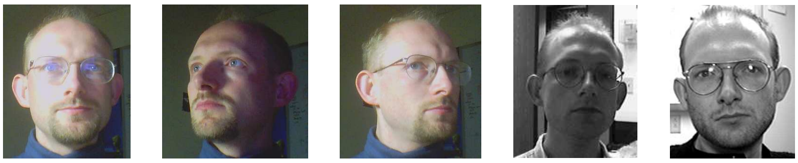

Figure 4.1. Sample variations of a single face: in pose, facial appearance, age, lighting and expression.

No current system can claim to handle all of these problems well. In particular there has been little research on making face recognition robust to the effects of aging the faces. In general, constraints on the application scenario and capture situation are used to limit the amount of invariance of face image sample that needs to be afforded algorithmically.

The main challenges of face recognition today are handling rotation in depth and broad lighting changes, together with personal appearance changes. Even under good conditions, however, accuracy needs to be improved.

4.2.1 Robustness and Fraud

All biometric recognition systems are susceptible to accidental errors of two types which both must be minimized: False Accept (FA) errors where a random impostor is accepted as a legitimate users and False Reject (FR) errors where a legitimate user is denied access. Designers of biometric systems must also be very conscious of how the system will behave when deliberately attacked. Naturally much of biometric system design falls into the more traditional categories of physical, procedural and electronic security — preventing an attacker from circumventing the recognition system or preventing false enrollment of biometric identities into a system's database, for example. That is, purposeful and successful attempts at creating a false accept error by general means of security attacks. Nevertheless, there are a number of security attack types that are specific to biometrics.

It is very easy to change one's facial appearance to make one look very different, and so to prevent identification, *i.e.* cause a false rejection. This is particularly important in a 'non-cooperative' application where the biometric is being used to prevent a single person from obtaining a privilege (such as a vote or driving licence) more than once. While underlying bone structure is extremely difficult to change, it is also hard to measure, and all face recognition systems rely on more superficial, changeable characteristics (Section 4.3.3) making them defeasible for determined individuals.

It is also possible for some people to impersonate others with a high degree of similarity (an important vulnerability in 'cooperative' applications like physical access control). Photographs, rubber masks, video replay all allow impostor attacks — the deliberate engineering of a false acceptance error. Detection of such fake biometrics data is only superficially handled by commercial systems, though this is improving. A couple of years ago, few systems had a test to detect authenticity (rejecting objects that looked too flat to be faces rather than photographs), but a recent PC Magazine test [21] found that both systems tested could distinguish a real person from a photograph. More sophisticated shape algorithms could be devised, and elastic deformation can be used to prevent simple photograph replay attacks. (One system allows the option of requiring a change in facial expression during verification.) With computing power more abundant, the technology for detecting fake biometrics will keep improving.

The combination with other biometrics — particularly lip motion verification or speaker ID [23] reduces the exposure to impersonation attacks, but further measures are necessary to prevent video replay attacks where a pre-recorded sequence of the authorized individual is somehow injected into the system. Well established in speaker identification literature [2], prompted-text or text-independent verification can avoid a simple replay attack, at the cost of a more intrusive, complex and expensive system, but the advances in trainable speech and face synthesis algorithms [11, 15] furnish attacks on even these sophisticated systems.

4.3. The Technology of Face Recognition

In this section we briefly review some of the technologies that have been used for face recognition. In general, face recognition systems proceed by detecting the face in an image, with the effect of estimating and normalizing for translation, scale and in-plane rotation. Given a normalized image, the features, either global or local, are extracted and condensed in a compact face representation which can then be stored in a database or a smartcard and compared with face representations derived at later times.

4.3.1 Related Fields

Face recognition is closely related to many other domains, and shares a rich common literature with many of them. Primarily, face recognition relies upon face detection described in Section 4.3.2. For recognition of faces in video, face tracking is necessary, potentially in three dimensions with estimation of the head pose [18]. This naturally leads to estimation of the person's focus of attention [9, 32] and estimation of gaze [20] which are important in human-computer interaction for understanding intention, particularly in conversational interfaces. Correspondingly there is much work on person tracking [27] and activity understanding [37] which are important guides for face tracking and for which face recognition is a valuable source of information. Recent studies have also begun to focus on facial expression analysis either to infer affective state [30] or for driving character animations particularly in MPEG-4 compression [26]. The recognition of visual speech (*i.e.* lip-reading, particularly for the enhancement of acoustic speech recognition) is also a burgeoning face image processing area [1].

4.3.2 Face Detection

Naturally, before recognizing a face, it must be located in the image. In some cooperative systems, face detection is obviated by constraining the user. Most systems use a combination of skin-tone and face texture to determine the location of a face and use an image pyramid to allow faces of varying sizes to be detected. Increasingly, systems are being developed to detect faces that are not full-frontal [13]. Cues such as movement and person detection can be used [38] to localize faces for recognition. Typically translation, scale and in-plane rotation for the face are estimated simultaneously, along with rotation-in-depth when this is considered.

4.3.3 Face Recognition

There is a great diversity in the way facial appearance is interpreted for recognition by an automatic system. Currently a number of different systems are under development, and which is most appropriate may depend on the application domain. A major difference in approaches is whether to represent the appearance of the face, or the geometry. Brunelli and Poggio [5] have compared these two approaches, but ultimately most systems today use a combination of both appearance and geometry. Geometry is difficult to measure with any accuracy, particularly from a single still image, but provides more robustness against disguises and aging. Appearance information is readily obtained from a face image, but is more subject to superficial variation, particularly from pose and expression changes. In practice for most purposes, even appearance-based systems must estimate some geometrical parameters in order to derive a 'shape-free' representation that is independent of expression and pose artefacts [8, 12]. This is achieved by finding facial landmarks and warping the face to a canonical neutral pose and expression. Facial features are also important for geometric approaches and for anchoring local representations.

Face appearance representation schemes can be divided into local and global, depending on whether the face is represented as a whole, or as a series of small regions. Most global approaches are based on a principal components representation of the face image intensities. This representation scheme was devised first for face image compression purposes [17] and subsequently used for recognition purposes [39]. The latter coined the term *eigenfaces* for this type of representation. A face image is represented as a vector of intensities and this vector is then approximated as a sum of basis vectors (eigenfaces) computed by principal component analysis from a database of face images. These principal components represent the typical variations seen between faces and provide a concise encapsulation of the appearance of a sample face image, and a basis for its comparison with other face images. This principal components representation is, like for example the Fourier transform, a decorrelating transform to an alternative basis where good representations of the salient characteristics of an image can be created from only a few low-order coefficients despite discarding many of the higher-order terms.

Other researchers have taken the approach of local representations [42, 25, 36]. Local representations have the advantage that only part of the representation is corrupted by local changes on the face. Thus, donning sunglasses only affects the local features near the eyes, but it may still be possible to recognize someone from features derived from around the nose and mouth. However, as mentioned above, inherently local representations are harder to estimate and there is a trade-off between feature estimation precision and feature size (locality of the representation).

Matching. Having processed a face and extracted the features, these are stored or transmitted as a facial code (face template), which can be as small as 84 bytes (Visionics). For each representation type, a distance or similarity measure is defined that allows 'similar' faces to be determined. Much of the art in biometrics is in the design of a model of the biometric data and, given a scheme for extracting the model parameters as a representation of the data, in creating a similarity measure that correctly discriminates between samples from the same person and samples from different people. As with any biometric system, some threshold on similarity must be chosen above which two face images are deemed to be of the same person. Altering the threshold gives different False Accept and False Rejection Rates (Section 4.2.1) — trading the one off against the other depending on the security level required. This is a trade-off between convenience and security: user-friendly matchers have a low false reject rate, while secure matchers have a low false accept rate.

4.3.4 Performance

The Face Recognition Technology (FERET) tests from Jonathan Phillips [28] provided an early benchmark of face recognition technologies. Phillips has continued the evaluation of face systems for US government agencies in the Face Recognition Vendor Tests [4]. This report provides an excellent independent evaluation of three state-of-the-art systems with concrete performance figures. The report highlights the limitations of current technology — while under ideal conditions performance is excellent, under conditions of changing illumination, expression, resolution, distance or aging, performance falls off, in some cases dramatically. Current face recognition systems are not very robust yet against deviations from the ideal face image acquisition but there is continual performance improvement.

4.4. Privacy Issues

With the widespread deployment of security cameras, and the increasing financial and technological feasibility of automating this surveillance, public fears have also increased about the potential for invasion of privacy that this technology can bring about. Notable deployments of face recognition in the London borough of Newham, in Tampa Florida [41] and at the 2001 Super bowl [40] have raised the spectre of intrusive applications of face recognition. It is now starting to become easy and cheap to connect a face recognition system to a blanket video surveillance system with great potential for crime prevention, but also bringing undreamt-of powers of control to totalitarian regimes, and the erosion of civil liberties by an ever-wakeful, omniscient 'big brother' [24] capable of tracking the activities of its citizens from cradle to grave.

Technology will have answers to assuage these fears: Cryptography will go a long way toward privacy-guarding; and rigorous rights management, to limit access to the information, will prevent privacy violations by unauthorized individuals. Automatic identity-masking controls may make these technologies in theory less privacy-intrusive than human visual surveillance systems in that an automatic surveillance system can prevent voyeurism by only allowing people access to the video when a security incident has been detected. However, it seems that this technology is a tool as any other, and only legislation, self-regulation and social pressure will guide its use to beneficial rather than oppressive aims. Inevitably, in a pluralist world, there will be applications that tend to the latter.

4.5. Application Domain

Many applications for face recognition have been envisaged, and some of them have been hinted at above. Commercial applications have so far only scratched the surface of the potential. Installations so far are limited in their ability to handle pose, age and lighting variations, but as technologies to handle these effects are developed, huge opportunities for deployment exist in many domains.

Access Control. Face verification, matching a face against a single enrolled exemplar, is well within the capabilities of current Personal Computer hardware. Since PC cameras have become widespread, their use for face-based PC logon has become feasible, though take-up seems to be very limited. Increased ease-of-use over password protection is hard to argue with today's somewhat unreliable and unpredictable systems, and for few domains is there motivation to progress beyond the combinations of password and physical security that protect most enterprise computers. As biometric systems tend to be third party, software add-ons the systems do not yet have full access to the greater hardware security guarantees afforded by boot-time and hard disk passwords. Visionics' face-based screen lock is one example, bundled with PC cameras. Naturally such PC-based verification systems can be extended to control authorization for single-sign-on to multiple networked services, for access to encrypted documents and transaction authorization, though again uptake of the technology has been slow.

Face verification is being used in kiosk applications, notably in Mr. Payroll's (now Innoventry) cheque-cashing kiosk with no human supervision. Innoventry claims to have one million enrolled customers. Automated Teller Machines, already often equipped with a camera, have also been an obvious candidate for face recognition systems (*e.g.* Viisage's FacePIN), but development seems not to have got beyond pilot schemes. Banks have been very conservative in deploying biometrics as they risk losing far more through customers disaffected by being falsely rejected than they might gain in fraud prevention. Customers themselves are reluctant to incur burdensome additional security measures when their personal liability is already limited by law. For better acceptance, robust passive acquisition systems with very low false rejection probabilities are necessary.

Physical access control is another domain where face recognition is attractive (*e.g.* Cognitec's FaceVACS, Miros' TrueFace) and here it can even be used in combination with other biometrics. BioId [23] is a system which combines face recognition with speaker identification and lip motion.

Identification Systems. Two US States (Massachusetts and Connecticut [3]) are testing face recognition for the policing of Welfare benefits. This is an identification task, where any new applicant being enrolled must be compared against the entire database of previously enrolled claimants, to ensure that they are not claiming under more than one identity. Unfortunately face recognition is not currently able to reliably identify one person among the millions enrolled in a single state's database, so demographics (zip code, age, name *etc.*) are used to narrow the search (thus limiting its effectiveness), and human intervention is required to review the false alarms that such a system will produce. Here a more accurate system such as fingerprint or iris-based person recognition is more technologically appropriate, but face recognition is chosen because it is more acceptable and less intrusive. In Connecticut, face recognition is the secondary biometric added to an existing fingerprint identification system. Several US States, including Illinois, have also instituted face recognition for ensuring that people do not obtain multiple driving licenses.

Surveillance. The application domain where most interest in face recognition is being shown is probably surveillance. Video is the medium of choice for surveillance because of the richness and type of information that it contains and naturally, for applications that require identification, face recognition is the best biometric for video data. though gait or lip motion recognition have some potential. Face recognition can be applied without the subject's active participation, and indeed without the subject's knowledge. Automated face recognition can be applied 'live' to search for a watch-list of 'interesting' people, or after the fact using surveillance footage of a crime to search through a database of suspects.

The deployment of face-recognition surveillance systems has already begun (Section 4.4), though the technology is not accurate enough yet [14]. The US government is investing in improving this technology [10] and while useful levels of recognition accuracy may take some time to achieve, technologies such as multiple steerable zoom cameras, non-visible wavelengths and advanced signal processing are likely to bring about super-human perception in the data-gathering side of surveillance systems.

Pervasive Computing. Another domain where face recognition is expected to become very important, although it is not yet commercially feasible, is in the area of pervasive or ubiquitous computing. Many people are envisaging the pervasive deployment of information devices. Computing devices, many already equipped with sensors, are already found throughout our cars and in many appliances in our homes, though they will become ever more widespread. All of these devices are just now beginning to be networked together. We can envisage a future where many everyday objects have some computational power, allowing them to adapt their behaviour — to time, user, user control and a host of other factors. The communications infrastructures permitting such devices to communicate to one another are being defined and developed (*e.g.* Bluetooth, IEEE 802.11). So while it is easy to see that the devices will be able to have a well-understood picture of the virtual world with information being shared among many devices, it is less clear what kind of information these devices will have about the real physical world.

Most devices today have a simple user interface with inputs controlled only by active commands on the part of the user. Some simple devices can sense the environment, but it will be increasingly important for such pervasive, networked computing devices to know about the physical world and the people within their region of interest. Only by making the pervasive infrastructure *'human aware'* can we really reap the benefits of productivity, control and ease-of-use that pervasive computing promises. One of the most important parts of human-awareness is knowing the identity of the users close to a device, and while there are other biometrics that can contribute to such knowledge, face recognition is the most appropriate because of its passive nature.

There are many examples of pervasive face recognition tasks: Some devices such as Personal Digital Assistants (PDAs) may already contain cameras for other purposes, and in good illumination conditions will be able to identify their users. A domestic message centre may have user personalization that depends on identification driven by a built-in camera. Some pervasive computing environments may need to know about users when not directly interacting with a device, and may be made ‘human aware’ by a network of cameras able to track the people in the space and identify each person, as well as have some understanding of the person’s activities. Thus a video conference room could steer the camera and generate a labelled transcript of the conference; an automatic lobby might inform workers of specific visitors; and mobile workers could be located and kept in touch by a system that could identify them and redirect phone calls.

4.6. The IBM Face Recognition System

In recent years we have developed a face recognition system at IBM Research for use in a variety of projects across a number of application domains. The system is more fully described elsewhere [34, 35, 9, 22, 1] but here we present a brief overview of the approach and the application domains.

The system consists of four modules: face detection and tracking; facial feature finding; face representation; and matching. These are carried out in turn on any still image or video frame presented for recognition.

4.6.1 Face Detection

Face detection scans an image pyramid to detect faces regardless of scale and location, and uses a filtering hierarchy procedure to filter out locations that do not represent faces with successively more accurate face classifiers. A variety of face classifiers is used varying from the fast, but less accurate, Fisher’s linear discriminant to a mixture of Gaussians model which is slower but is more correctly able to determine if an image region is a face or not. For colour images, the first stage of the filtering hierarchy is a skin-tone detector.

4.6.2 Feature Finding

The next stage of the system finds 29 standard features (such as corners of eyes, nose, mouth and eyebrows, some of which are shown in figure 4.2) on the face for use in anchoring the representation. Based on the location, scale and orientation of the detected face, the system uses anthropometric data gathered from a training set to predict the approximate location of the principal features (eyes, nose and mouth). The system works in a hierarchical manner to locate first these larger features, and then to locate smaller sub-features (such as the corners of eyes, nose and mouth) relative to them. Detectors (a combination of linear discriminant and Distance from feature space similar to the face detector) trained on a database of labelled face features are applied over a region close to the prediction to determine the feature’s actual location, indicated by the maximum response for the detector in the search region.

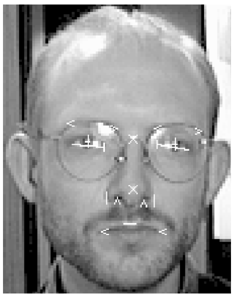

Figure 4.2. Principal facial features (in white) located by the system.

The procedure is repeated, predicting the sub-features' locations relative to the principal features and localizing them with trained detectors operating on a larger scale image. Finally the feature locations are verified with collocation statistics to reject any mislocated features, and additional anchor points are generated by geometric combinations of the visually located anchors.

4.6.3 Recognition

Recognition is carried out by finding a local representation of the facial appearance at each of the anchor points. The representation scheme used here is a vector of Gabor wavelet responses [43]. A range of 40 Gabor wavelets, with varying scale and orientations, is used to represent the local image appearance around each of the anchor features. This produces a 40-element vector, **a**, for each of the feature locations. The set of 29 vectors comprises the representation of the person's face to be stored in the face database.

Matching is carried out by comparing these features pairwise using the following similarity measure (each feature from one face with the corresponding feature from another face).

$$\mathcal{S}(\mathbf{a}, \mathbf{a}') = \frac{\sum_j a_j a'_j}{\sqrt{\sum_j a_j^2 \sum_j (a'_j)^2}} \tag{4.1}$$

Each such comparison gives a similarity score. Combining all these scores gives an overall match score used to determine if the face images represent the same person. Multiple representations from successive images in a video sequence can be aggregated into a distribution capturing the facial variation, and these distributions can be compared using statistical distance measures to give a similarity score based on many frames of data.

4.6.4 Applications

The system has been designed to be generally applicable to a variety of applications, and as such accepts colour or black and white images both still and video. It has been used as a black-and-white mug shot identification system; with PC-attached cameras for computer logon from a smart-card stored database; and on broadcast video for indexing from a database of enrolled TV presenters [35]. Components of the system have also been used in a number of other projects such as audio-visual speech recognition (visual lip reading to enhance acoustic speech recognition) [1] and user intention determination (using visual cues to understand the user, particularly to whom speech is being addressed) [9].

4.7. Conclusions

Face recognition is a technology just reaching sufficient maturity for it to experience a rapid growth in its practical applications. Much research effort around the world is being applied to expanding the accuracy and capabilities of this biometric domain, with a consequent broadening of its application in the near future. Verification systems for physical and electronic access security are available today, but the future holds the promise and the threat of passive customization and automated surveillance systems enabled by face recognition.

References

[1] S. Basu, C. Neti, N. Rajput, A. Senior, L. Subramaniam, and A. Verma. Audio-visual Large Vocabulary Continuous Speech Recognition in the Broadcast Domain. In Multimedia Signal Processing, 1999.

[2] H. S. M. Beigi, S. H. Maes, U. V. Chaudhari, and J. S. Sorensen. IBM Model-based and Frame-by-frame Speaker Recognition. In Speaker Recognition and its Commercial and Forensic Appications, Avignon, April 1998.

[3] Biometrics in Human Services User Group.
URL: http://www.dss.state.ct.us/digital.htm.

[4] Duane M. Blackburn, Mike Bone, and P. Jonathon Phillips. Facial Recognition Vendor Test 2000 Evaluation Report. Technical Report, Department of Defence Counterdrug Technology Development Program Office, February 2001. http://www.dodcounterdrug.com/facialrecognition/DLs/FRVT_2000.pdf.

[5] Roberto Brunelli and Tomaso Poggio. Face Recognition: Features versus Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10):1042–1052, October 1993.

[6] Rama Chellappa, Charles L. Wilson, and Saad Sirohey. Human and Machine Recognition of Faces: A Survey. Proceedings of the IEEE, 83(5):705–740, May 1995.

[7] Chin-Seng Chua, Feng Han, and Yeong-Khing Ho. 3D Human Face Recognition using Point Signature. In International Conference on Face and Gesture Recognition, pages 233–238, 2000.

[8] Ian Craw and Peter Cameron. Face Recognition by Computer. In David Hogg and Roger Boyle, editors, Proceedings of the British Machine Vision Conference, pages 498–507. Springer Verlag, September 1992.

[9] P. de Cuetos, C. Neti, and A. Senior. Audio-visual intent to Speak Detection for Human-computer Interaction. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2000.

[10] Defense Advanced Research Projects Agency. Human Identification at a Distance, BAA00-29 edition, Feb 2000.
URL: http://www.darpa.mil/iso2/HID/BAA0029_PIP.htm.

[11] Robert Donovan. Trainable Speech Synthesis. PhD thesis, Cambridge University Engineering Department, 1996.

[12] G. J. Edwards, C. J. Taylor, and T. F. Cootes. Interpreting Faces using Active Appearance Models. In International Conference on Face and Gesture Recognition, number 3, pages 300–305, April 1998.

[13] Raphael Feraud, Olivier Bernier, Jean Emmanuael Viallet, and Michel Collobert. A Fast and Accurate Face Detector for Indexation of Face Images. In International Conference on Face and Gesture Recognition. IEEE, March 2000.

[14] Lee Gomes. Can Facial Recognition Help Snag Terrorists? The Wall Street Journal, September 21 2001.

[15] H.P. Graf. Sample-based Synthesis of Talking Heads. In Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, pages 3–7, July 2001.

[16] http://www.innoventry.com.

[17] M. Kirby and L. Sirovich. Application of the Karhunen-Loève Procedure for the Characterization of Human Faces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(1):103–108, 1990.

[18] M. La Cascia, S. Sclaroff, and V. Athitsos. Fast, Reliable Head Tracking under Varying Illumination: An Approach Based on Registration of Texture-mapped 3D Models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(4):322–336, April 2000.

[19] B. Maison, C. Neti, and A. Senior. Audio-visual Speaker Recognition for Video Broadcast News: Some Fusion Techniques. In Multi-media Signal Processing, 1999.

[20] Y. Matsumoto and A. Zelinsky. An Algorithm for Real-time Stereo Vision Implementation of Head Pose and Gaze Direction Measurement. In IEEE International Conference on Face and Gesture, page 499, 2000.

[21] Glenn Menin. Performance Tests: Fingerprint Biometrics. PC Magazine, June 12 2001.

[22] Chalapathy Neti and Andrew W. Senior. Audio-visual Speaker Recognition for Broadcast News. In DARPA Hub 4 Workshop, pages 139–142, March 1999.

[23] Ana Orubeondo. A New Face for Security. InfoWorld.com, May 2001.

[24] George Orwell. 1984. 1948.

[25] P. S. Penev and J.J. Atick. Local Feature Analysis: A General Statistical Theory for Object Representation. Network: Computation in Neural Systems, 7(3):477–500, 1996.

[26] E. Petajan. The Communication of Virtual Human Faces using mpeg-4 Tools. In International Symposium on Circuits and Systems, volume 1, pages 307–310, 2000.

[27] Second International Workshop on Performance and Evaluation of Tracking and Surveillance. IEEE, December 2001.

[28] P. Jonathon Phillips, Hyeonjoon Moon, Patrick Rauss, and Syed A. Rizvi. The FERET September 1996 Database and Evaluation Procedure. In Josef Bigün, Gérard Chollet, and Gunilla Borgefors, editors, Audio- and Video-based Biometric Person Authentication, volume 1206 of Lecture Notes in Computer Science, pages 395–402. Springer, March 1997.

[29] P. Jonathon Phillips, Patrick J. Rauss, and Sandor Z. Der. FERET (Face Recognition Technology) Recognition Algorithm Development and Test Results. Technical Report ARL–TR–995, Army Research Laboratory, October 1996.

[30] Rosalind W. Picard. Affective Computing. MIT Press, 2000.

[31] F. Prokoski. History, Current Status, and Future of Infrared Identification. In Proceedings of IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications, pages 5–14, June 2000. Facial Thermogram.

[32] James M. Rehg, Kevin P. Murphy, and Paul W. Fieguth. Vision-based Speaker-detection using Bayesian Networks. In Proceedings of Computer Vision and Pattern Recognition, volume 2, pages 110–116, 1999.

[33] A. Samal and P.A. Iyengar. Automatic Recognition and Analysis of Human Faces and Facial Rxpressions: A Survey. Pattern Recognition, 25(1):65–77, 1992.

[34] Andrew W. Senior. Face and Feature Finding for a Face Recognition System. In Second International Conference on Audio- and Video-based Biometric Person Authentication, pages 154–159, March 1999.

[35] Andrew W. Senior. Recognizing Faces in Broadcast Video. In IEEE International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, pages 105–110, September 1999.

[36] Daniel L. Swets and John (Juyang) Weng. Using Discriminant Eigenfeatures for Image Retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(8):831–836, August 1996.

[37] T. Tan, editor. Second IEEE International Workshop on Visual Surveillance. IEEE, 1999.

[38] Jochen Triesch and Christoph von der Malsburg. Self-organized Integration of Adaptive Visual Cues for Face Tracking. In International Conference on Face and Gesture Recognition, pages 102–107. IEEE, March 2000.

[39] M. Turk and A. Pentland. Eigenfaces for Recognition. Journal of Cognitive Neuro Science, 3(1):71–86, 1991.

[40] Press releases http://www.viisage.com, January 20 2001.

[41] Press releases http://www.visionics.com, June 2001.

[42] Laurenz Wiskott, Jean-Marc Fellous, and Norbert Krüger. Face Recognition by Elastic Bunch Graph Matching. Technical Report IR-INI 96–08, Buhr-Universität Bochum, Institut für Neuroinformatik, April 1996.

[43] Laurenz Wiskott and Christoph von der Malsburg. Recognizing Faces by Dynamic Link Matching. In Proceedings of the International Conference on Artificial Neural Networks, pages 347–352, 1995.