Transparent face recognition in the home environment

G.M. Beumer, R.N.J. Veldhuis, A.M. Bazen
University of Twente
P.O. box 217 7500 AE Enschede, the Netherlands
Phone: +31 (0)53 489 2842 Fax: +31 (0)53 489 1060
E-mail: {g.m.beumer,r.n.j.veldhuis,a.m.bazen}@utwente.nl

Abstract—**The BASIS project is about the secure application of transparent biometrics in the home environment.**

Due to transparency and home-setting requirements there is variance in appearance of the subject. An other problem which needs attention is the extraction of features. The quality of the extracted features is not only depending on the proper preprocessing of the input data but also on the suitability of the extraction algorithm for this problem.

Possible approaches to address problems due to transparency requirements are the use of active appearance models in face recognition, smart segmentation, multi-camera solutions and tracking.

In this paper an inventory of problems and possible solution will be given.

Keywords— **Biometrics, face recognition, feature extraction, home environment**

I. Introduction

User convenience, or ease of use, is an important issue when considering the residential environment of 2010 . It is a factor besides security and privacy protection that will determine whether or not the possibilities offered by wide-band access to the home will be appreciated and used. *Biometric authentication, i.e. verifying the claimed identity of a person based on physiological characteristics or behavioural traits*, has the potential to contribute to both security and user convenience.

From a user-convenience point of view, biometric authentication has the advantage that it does not make use of pin codes or passwords that can be forgotten or tokens that can be lost. Another advantage that biometric authentication offers is the possibility of personalisation, because a device or service can recognise a user and adapt its settings to the users preferences. User convenience can be further increased, when biometric recognition is made transparent. This means that it does not require any specific user action, such as placing a finger on a sensor in order to present a fingerprint.

From a security point of view, biometric authentication offers the possibility to verify whether or not a user is physically present. However, it must be noted that biometric authentication is intrinsically characterised by false-acceptance and false rejection errors. Therefore, not all biometric recognition methods will be able to achieve the same level of security as for instance a pin code.

From a privacy point of view, the use of biometric authentication also introduces a new problem and raises user concerns. Namely, when used for privacy-sensitive applications, biometric data are a highly valuable asset. When such data are available to unauthorised persons, these data can potentially be used for impersonation purposes, defeating the security aspects that are supposed to be associated with biometric authentication. European privacy legislation provides for various protection regimes that cover biometric personal data, depending on their degree of vulnerability and the purpose of their processing. Initial results from studies, done in the context of the European project BIOVISION, show that there is a variety of user concerns, associated with loss of privacy, reuse of electronically stored fingerprints and written signatures and the fear that biometric data might reveal medical conditions. One of the most promising privacy enhancing solutions is anonymous biometric authentication. Biometric data are called anonymous when the data cannot be traced back to the user. This means that privacy sensitive information about physiological characteristics cannot be derived from the data.

The home is a challenging environment for the introduction of biometric authentication. First of all, it is a place where user convenience and personalisation are highly appreciated or even demanded. Biometric authentication, in particular transparent biometric authentication, seems the security mechanism to achieve this. Secondly, electronic banking and electronic voting will be typically done from the home. These applications require the privacy protection that anonymous biometric authentication can offer. Thirdly, the home environment poses some specific challenges that need to be addressed. For example, in contrast to access-control or banking applications, there is no professional system manager, who can assist with the enrolment and withdrawal of users, or who can set up and maintain biometric databases.

The goal of this project is *to investigate the possibilities* *of biometric authentication for securing the access to information and services in the personal environment, with a focus on user convenience and privacy protection*. The project will address (a) the problem of transparent biometric authentication as a means to enhance user convenience, (b) the problem of anonymous biometric authentication as a means to protect the users privacy, and (c) the specific problems of the use of biometric authentication in the home environment.

In this paper the use of face recognition in the home environment will be discussed. An inventory of some problems and possible solutions will be given.

II. Transparent biometrics

As said, biometrics is a way to identify a person by body characteristics or traits. Already a lot of biometric recognition methods are known such as fingerprint, face, iris, speaker, odour, gait, posture, grip recognition. Not all are as suitable for the home environment, due to patent restrictions, costs, performance, transparency requirements and others reasons.

Transparency means that in order to be recognised a person does have to perform any explicit action. Thus any biometric which does require user action such as fingerprint, grip and iris recognition are, at least with current technology, unsuitable because the person has to present an finger, hand or eye to a sensor. Face, posture, gait recognition are examples of biometrics which can be applied in a transparent way.

A. Face recognition

Currently our research focuses on face recognition. This is because in our opinion it offers possibility to be adapted to transparency and does not involve patented technology. Face recognition lends itself well for transparent use because it is based on cameras that can in a later stage be applied to an other biometric such as gait or posture recognition.

III. Problems

In this section several of the problems arising in the application of transparent face recognition in the home environment and their causes will be discussed.

A. Variability

Variability is the fact that two images of a person taken for identification can differ due to numerous reasons.

As a consequence of of the transparency users in the house will not perform any action to be recognised but just follow their daily routine. Also, in the house the environmental conditions cannot be controlled as in a laboratory. In most face-recognition systems there is a controlled situation whit controlled lighting conditions, a fixed frontal pose, neutral expression. In a transparent environment this is not the case and the conditions are far from ideal, which leads to a high variability.

Pose The pose of a person is how the person is facing the camera. This is often not in a frontal or prescribed way. This means that the images on which the recognition is to be based will be frontal, profile, from the back of the head or from any other angle. Also it is obvious that the distance to a camera will give big variability in scale.

Lightning Apart from pose there are also differences in the lightning conditions. This is caused not only by the difference in position of cameras in a multi camera system. The conditions in the house itself may vary where one can think of sunlight coming through the window or the switching on or off the lights in the house.

Occlusion When an image of a person is a unobstructed frontal shot all his facial features are visible. However, part of the face can be hidden behind the head itself. Also, parts could be hidden behind objects such as furniture or other persons in the room. The fact that not all of the face is visible means that certain features need to be guessed.

Expression People are living and interacting with emotions showing from their faces. This makes that images taken from person in the home environment will contain images of people with different expressions. This can cause problems if the system is not trained well enough to be able to cope with the expressions. These variations are in general fast and could change over seconds.

Slow changes ”People change” is a well known saying. This is also true for their faces. It can be anything from people starting to grow a beard of aging effects. In face recognition this is not a good thing. When the system is trained to recognise people and the people change the recognition might start to fail or stop working properly.

The variability encountered can be roughly separated into two groups; intrinsic and extrinsic variability. Intrinsic variability is variability coming from the person itself. They can be fast, such as expression, or slow as aging. Extrinsic variability are basically variabilities which are caused by the position of the camera or other outside influences.

B. Face detection and image quality

When a face has to be identified it first has to be detected in the image. In addition it is known that the performance of the recognition depends for a large part on the registration method. In order to do proper registration features in the face have to be found. In a home situation, the non-uniform background, large subject-camera distances and variability in poses make the detection of a face and it's features more difficult and less reliable.

C. Limited enrollment

A recognition system has three basic phases in which it processes images. The first is the training phase in which it learns the differences between classes in a general way. A face recognition system can be trained with a number of pictures of a number of people. Only the general data and parameters of the feature extraction are kept. The second phase is the enrollment phase. Here the systems learn the individuals it has to recognise later. Of these persons the extracted data is kept. The final phase is the recognition. An image is compared to the data stored during the enrollment phase. Because there is no professional system manager the enrollment phase must be simple and minimal. Only one or a few images taken in one session will be stored in the database during enrollment. These images will be taken in the home environment and not in a laboratory under well defined conditions. This will make the images less suitable for enrollment and may reduce the performance and it will make the problems discussed before even worse.

IV. Solutions

In this section solutions for the problems discussed in the previous section will be proposed. Some solutions might be suitable to handle more then one problem while some problems are possible to solve with more than one of the proposed solutions.

A. Multiple images

When the quality of the images is not good enough, more images can be used to enhance the overall system performance. There are two moments where the information can be combined; before or after the classification. First the images could be used to create a better image of the person. After that the recognition takes place. An other approach is to perform recognition on each of the captured images and combine the scores into one final recognition score.

There are several ways to use multiple images of the same person to enhance the performance. The first, is to use more than one image over time. This is making use of image sequences which come from the camera. The other one is to use multiple cameras to capture multiple images at the same time at different angles. A combination is also possible.

A.1 Image sequences

The knowledge of where a face was in the previous image can provide important clues for where to look for it in current image. It could be that special video tracking algorithms could provide even more advantages to plain image processing. The detection can be improved using tracking.

A.2 Multiple camera

The use of multiple camera's shows advantages in even simple recognition algorithms. More cameras means more images. When a person is visible on multiple cameras at the same time the chance of obtaining a good recognition result will improve. Even in the very simple case where identification is done in parallel and independent for each camera. The probability that the right decision is made increases. A challenge lies in combining the input of the cameras before the recognition is done.

Using multiple cameras the deployment of 3D technology becomes possible. It can be used to do 3D recognition but also to remove obstructions in front of the face.

Expectations are that more can be gained using advances technologies to combine the cameras.

B. Invariance models

In order to cope with extrinsic variability and fast intrinsic variability in a face a possible approach is to take the variance out of the face. Making a face free of variance in shape will make recognition more reliable. One should however be aware that making faces shape free reduces within-class as well as between-class variability. This might lead to performance degeneration. Also finding the shape of the face can help improve the recognition compared to recognition purely on the texture.

B.1 Active Appearence Models

Promising work on the compensation for fast intrinsic variability and some extrinsic variabilities such as pose and small occlusions is done in the field of the Active Shape Models and Active Appearance Models (AAM) [1]. Active appearance models contain not only information on the gray scale information of the model but also on the shape.

When shape parameters are known the face can be warped to an expression free patch. This means that the face is now is free of variance as expression. This new texture can now be used to determine who the image belongs to.

C. Smart segmentation

A different approach to the problem of occlusion could be to simply split the face in different areas like the right eye, left eye, nose, mouth, left ear, right ear and perhaps more. Then try to identify a person on each of these segments using the technology which also is used for 'whole face' recognition. The outcome of these identifications should be combined into a reliable estimate of the identity.

D. Colour

Most face recognition systems use grayscale images to work with. In black and white images there is only the intensity on a 0-255 grayscale to work with. Working on colour images may give a few advantages. There simply is more information available in a colour image which could help improving the recognition [2]. Besides that the colour information may help in locating suitable face candidates [3] but it depends more on lightning conditions.

In colour images a few problems arise which are not present in black and white image processing. One of the problems is how to handle the colour data. Once could simply use the red, green and blue (RGB) values which give the intensities three basic colours. When one looks at the colour space as a three-dimensional space with red, green and blue on the axis the colour space will be a cube making the correct calculation of distances difficult. For Hue Saturation and intensity (HSI) the space would look like a cone. Again the distance between two colours is not easily calculated and uniform throughout the colour space. The Lab colour space [4] however is shaped like a sphere and the difference between the calculated colour values is equivalent to the visual distinctness between the colours.

E. Feature extraction enhancement

The feature extraction should be done with existing technologies. There are of course still possibilities to improve the recognition. Traditionally a person is judged against n classes. It is however also possible to create n feature extractors which separate one class from all other classes. It should be investigated if n is small enough to implement this and if this does improve the results.

An image of n pixels can be seen as a single point in a n-dimensional space. Each dimension is a feature. Computing the statistics of many points in the n-dimensional space (or images) requires a lot of time. Due to time restrictions it is necessary to reduce the number of features. This of course with out loosing (to much) relevant information.

Traditional PCA and LDA systems have one transformation matrix for all the classes. It could be good idea to use a class-depending transformation matrices.

Each feature extraction can address a two class problem; the person and all other persons. There is the class for which extraction is optimised and all other classes (persons) on one big heap, which we call the *other-class*. The extraction maximises the distance between the class and the other-class over the within class variance of the class.

A new image is processed using all feature extractions. There it is tested to belong to the class or the other-class. Ideally the outcome is only once 'class' and 'other-class' for all other occasions.

This is made possible by online training and online updating.

E.1 recursive SVD

In BASIS the system is meant to update it self for two reasons. First of all to keep the setup and enrollment procedure as simple as possible. In the beginning the system will start to recognise a person with a less then optimal error rate. If the system can learn by itself the faces it has to recognise better the performance will improve, yielding better EER and therefore lower user frustration and thus higher user convenience. Secondly the system has to be able to adapt to slow intrinsic variability. To cope with this it is necessary to do online recalculations of the transformation matrix which is used to transform an image onto the feature space. If the system takes an image of a person which is good enough to be enrolled then the new transformation matrix can be calculated using the old transformation matrix and the new image alone. To calculate the transformation matrix, SVDs are used. In order to calculate the new transformation matrix a recursive SVD (RSVD) algorithm [5] is to be used. How to apply a RSVD is explained in appendix A.

The usefulness, in terms of speed and robustness of this and other RSVD algorithms should be investigated.

F. Selflearning systems

Selflearning systems may solve problems caused by a limited enrollment phase.

In order to cope with the slow variability the system should be able to update enrolled data. After the recognition is confirmed the image could be used to update the system. This has of course advantages and risks, such as stability of the system. How the proper operation of the system can be ensured when a wrong image is enrolled or how this can be prevented is to be investigated. Also there is the risk of overtraining the system to a level where it cannot detect if someone is enrolled or not. It will simply assign the not enrolled person to a class increassing error rates.

V. Planning

We plan to realise a sequence of demos in which the working of biometrics in the home environment are explored. The first demo is a simple recognition system with controlled conditions. The second one will include two cameras. It will find face and features by itself. After that updating algorithms and automated enrollment of users will be added. The final system should be able to grow to a good working recognition system by using online updating from only a few enrollment images. These demos will serve as a guideline for the research to be done.

VI. Conclusions

In this paper the problems arising when deploying face recognition in the home environment have been discussed. Also have some possible solutions and technologies to implement them been discussed defining research to be done. Also a plan to implement this in demos has been made.

References

[1] T.F.Cootes and C.J.Taylor, "Statistical models of apearence for computer vision." internet, Oktober 2001. http://isbe.man.ac.uk/.

[2] E. D. A. Albiol, L. Torres, "Optimum color spaces for skin detection.," in *IEEE International Conference on Image Processinng*, October 2001.

[3] T. C. Filipe Tomaz and H. Shahbazkia, "Improved automatic skin detection in color images," in *VIIIth digital image computing: techniques and applications* (S. C. T. H. O. S. A. T., ed.), pp. 419–427, December 2003.

[4] G. S. W. Stiles, *Color science: concepts and methods, quantitavive data and formulae*. John Wiley and sons, 2nd ed., 1982.

[5] M. Gu and S. C. Eisenstat, "A stable and fast algorithm for the singular value decomposition," 1994.

Appendix

I. Recursive SVD algorithm

This theory is based on the updating of the SVD as described by Gu and Eisenstat [5]. Let X be a matrix containing measurement vectors. There are n_m measurements vectors each containing n_p values. The SVD of X looks like:

$$X = USV^T \tag{1}$$

where $X, U \in R^{n_p \times n_m}$ and $S \in R^{n_m \times n_m}$. Now assume that a measurement is added to the matrix X

$$X_{new} = [X\ x]. \tag{2}$$

Without proof is stated how the result of an SVD on the new matrix can be approximated as

$$X_{new} \approx \hat{U}\hat{S}\hat{V}^T. \tag{3}$$

where $\hat{U}$, $\hat{S}$ and $\hat{V}$ are the estimates of U, S and V after adding a measurement. First compute a

$$a = \frac{x - U(U^T x)}{||x - U(U^T x)||} \tag{4}$$

and u,s and v by performing an SVD:

$$\begin{bmatrix} S & U^T x \\ 0 & a^T x \end{bmatrix} = usv^T. \tag{5}$$

Now the approximation of $\hat{U}$,$\hat{S}$ and $\hat{V}$ are given by:

$$\hat{U} = [U\ a]u \tag{6}$$
$$\hat{S} = s \tag{7}$$
$$\hat{V} = \begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix} v \tag{8}$$
$$\tag{9}$$

This algorithm also works with $\hat{U}$, $\hat{S}$ as inputs. After the increment $X_{new}, \hat{U} \in R^{n_p \times n_m+1}$ and $\hat{S} \in R^{n_m+1 \times n_m+1}$. Feature reduction works as on regular SVD, keeping the relevant parts of $\hat{U}$, $\hat{S}$.