Face as an Index:
Knowing Who Is Who Using a PDA

Jie Yang, Xilin Chen, William Kunz, Harsh Kundra

April, 2002
CMU-CS-02-131

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213, USA

ABSTRACT

In this report, we present a PDA-based system for extending human memory or/and information retrieval using a human face as the lookup index. The system can help a user to remember names of people whom he/she has met before, and find useful information, such as names and research interests, about people whom he/she is interested in talking to. The system uses a captured face image as the lookup index to retrieve information from some available resource such as departmental directory, web sites, personal homepages, etc. We describe the development of a PDA-based face recognition system, and introduce algorithms for image preprocessing to enhance the quality of the image by sharpening focus, and normalizing both lighting condition and head rotation. We use a unified LDA/PCA algorithm for face recognition. We address design issues of the interface to assist in visualization and comprehension of retrieved information. We present user study and experiment results to demonstrate the feasibility of the proposed system.

This Research is partially supported by the National Science Foundation under Grant No. IIS-0121560 and Microsoft Research through a gift. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF, Microsoft Research, or any other party.

Keywords: Face recognition, PDA, PCA, LDA, user study, information retrieval

Table of Contents

1. INTRODUCTION

Many people struggle to remember names and faces of people whom they have met before. Many people also would like to know more about a person before approaching a new person in situations such as at a university or company cocktail party. In this report, we propose to use a PDA (Personal Digital Assistant) to help a user to remember names of people whom he/she has met before, and/or find useful information (such as names and research interests) about someone with whom he/she is interested in speaking. The proposed system uses a human face as an index to retrieve relevant information from available resources such as a departmental directory, web sites, and personal homepages. For example, there is a face database online that contains basic information (name, phone number, office location, personal homepage, etc.) of current faculty, staff, and students at the School of Computer Science at Carnegie Mellon University. The proposed system can take advantage of such a source to retrieve relevant information for a user in a variety of situations.

The palm-size PDA is becoming ubiquitous among professionals. With continuous increases in computing power, memory, and accessories, a PDA can now provide more potential applications in our daily lives. Several palm size PDAs, such as Compaq iPAQ and HP Jornada, are now available with powerful CPU's (over 200 ~ 400MHz) and large amounts of memory (32 ~ 64 MB). These devices can provide essential resources to support a face recognition task. Furthermore, with wireless communication to the Internet, a PDA can retrieve relevant information at a user's request. A PDA with a camera and face recognition software can be used in a subtle and transparent way that will allow someone to use his/her PDA to recall a persons name and any relevant information about someone, thus forever doing away with the embarrassing request, "Excuse me, could you tell me your name again?"

The main technology in this system is automatic recognition of human faces. Automatic face recognition has been an active research area in the last two decades. The progress in this area can be found in review papers (Chellappa 1995, Samal 1992) and the proceedings of the last four international conferences on Automatic Face and Gesture Recognition. In general, face recognition methods fall into two different categories:

holistic template matching and geometric local feature-based schemes. Each type of system has certain advantages and disadvantages, even though both types of systems have some successful examples. Therefore, we should carefully select different approaches based on the specific requirements of a given task. Taking into consideration various requirements, limitations, and applications of a PDA-based face recognition task, we have decided to use a holistic template based matching scheme. For example, we may have different number training samples, from one to many, for different people. We have developed efficient algorithms to deal with such problems in our previous research (Yang 2000).

The previous work in system development has been mostly focused on using a fixed camera or a video surveillance camera. However, some recent publications have focused on face recognition systems on a wearable platform (Clarkson 2001, Mann 1996, Pentland 1998). The system was built on the "WearCam" platform, which permits both hands to be free while the user wears at least one camera on the body in some manner and receives feedback from the head-mounted display. This kind of system requires special hardware. In addition, the image quality is generally poor since it requires a wide-angle lens (or fish eye lens), and it is difficult to position the camera without using hands. Faces are generally only recognizable under bright lighting conditions and from less than 10 feet away (Clarkson 2001).

In this report, we present a PDA-based face recognition system. The system takes advantage of human feedback to aid in processing. We discuss some of the associated challenges of developing a PDA-based face recognition system. We describe a prototype system built from an off the shelf PDA, and we introduce algorithms for image preprocessing to enhance the quality of the image by sharpening focus, and normalizing both lighting condition and head rotation. We use a unified LDA/PCA algorithm for face recognition. The algorithm maximizes the LDA criterion directly without a separate PCA step, which eliminates the possibility of losing discriminative information due to a separate PCA step. We demonstrate effectiveness of these algorithms and the feasibility of this system through experiments. In the rest of this report, we describe the development of a PDA-based face recognition system. We discuss design of the interface for visualization of retrieved information. Finally, we present the user study and results of the experiments to demonstrate feasibility of the proposed system.

2. FACE RECOGNITION USING A PDA

Unfortunately, although most automatic face recognition systems used today work very well under constrained conditions, they may fail under unconstrained conditions that can vary vastly. Many factors can cause a face recognition system to perform poorly. Our experiments indicate that face localization and alignment are crucial for a face recognition system. The task is very difficult for a system under highly variable conditions, but not for a human. With the flexibility of a PDA, a user can capture an image from the best position and help select facial features appropriately which can be used by the system to normalize faces. The system can then focus on the recognition task. As such, we have shifted some duties from computers to human users: the systems no longer attempt to do everything themselves, but instead, take advantages of human capabilities, which in turn is then capable of assisting the user by recognizing the face. The underlying assumptions to this approach and a possible solution come from incorporating the human operative into the process, which concept is different from a fully automatic face recognition system and a "WearCam" type wearable face recognition system.

2.1 CHALLENGES

A PDA-based face recognition system not only shares some common challenges that the previous face recognition systems have met, but also has some particular problems. Many challenges of a face recognition system come from substantial variations in appearance that faces undergo with changing illumination, orientation, scale, and facial expressions. Although a PDA with the attached camera could capture high quality pictures of a face under a normal condition, it may capture a poor quality image when a person is far away or the lighting conditions are poor. Problems with focus were the most common challenges in our experiments. Furthermore, depending on the rotations of the head and the position of a person relative to the lighting sources, the illumination of the face changes dramatically. This means we can observe the full range of shade variations even though the overall lighting conditions in the room remain constant over the time. Moreover, people constantly move their heads and change their facial expressions. Given the dynamic nature of the environment where we use a PDA-based face recognition system, the image that we capture can be significantly different from the image in the database. Last but not least, a PDA has limited resources including:

- **Limited computing power:** Computing power of PDAs is much less powerful than that of desktop or laptop computers. Currently, the most powerful CPU for the Palm OS based system is 66MHz DragonBall CPU, while the most powerful CPU for WIN CE based system is a 206MHz StrongARM CPU. In addition, all CPUs of current PDAs don't have hardware supporting float point computation in order to reduce power consumption and chip size. Float point computations on a PDA are implemented by software through a float emulation library. For a floating task, a 200MHz integer CPU may be even slower than a 50MHz CPU with float point computing component.

- **Limited memory:** For a variety of reasons, all PDAs are configured with limited memory. A typical Palm OS based system is with 8-16MB memory, and a WIN CE based system is with 16-64MB memory. This space is for storage and application programs. The memory size is crucial for a computer vision application because some algorithms require a huge space.

- **Limited size of display:** All PDAs use a small display. Most of the Palm OS PDAs use 160 x 160 displayer (SONY is the only exception, which uses the 320x320 or 320x480 displayer), and all of the Palm PC and Pocket PC use 320x240 display, which is the quarter of standard VGA. A small sized display imposes more constraints in user interface design.

We will address these challenges in both algorithm development and system implementation.

2.2 APPROACHES

We shall now address the challenges in a PDA-based face recognition system in both image preprocessing and recognition algorithms.

2.2.1 Image Preprocessing

Almost all the cameras attached to a PDA are very simple and therefore cannot guarantee high quality images. Blurry images are one of the most common problems. The problem comes mainly from the structure of the simple camera, which does not have auto focus capabilities. To compensate for such a disadvantage, it uses a small size diaphragm to approximate a pinhole camera and sets its focus at the most commonly used distance. However, the camera will be out of focus when an object is far beyond the focus range, because it is impossible to implement a true pinhole camera. Figure 1 depicts this case. In addition, a small size diaphragm requires longer exposure time during which unexpected motion can happen. All these factors can cause a blurred image.

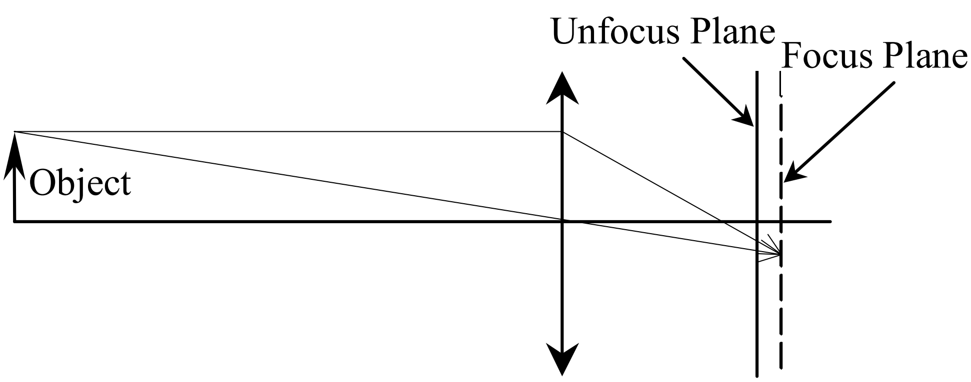


Figure 1 The near-pinhole camera model

A blurred image can be sharpened using image processing algorithms. In fact, a blurred image can be considered the result of a clear image passing through a linear shift-invariant system; i.e.,

$$g(x, y) = f(x, y) \otimes h(x, y), \qquad (1)$$

where $f(x, y)$ is the clear image, $h(x, y)$ is the kernel function, $\otimes$ is convolution operator, and $g(x, y)$ is the blurred image.

We can obtain a clear image from the blurred image using a deconvolution operation, if $h(x, y)$ is known. A practical method is to use a Gaussian model to approximate

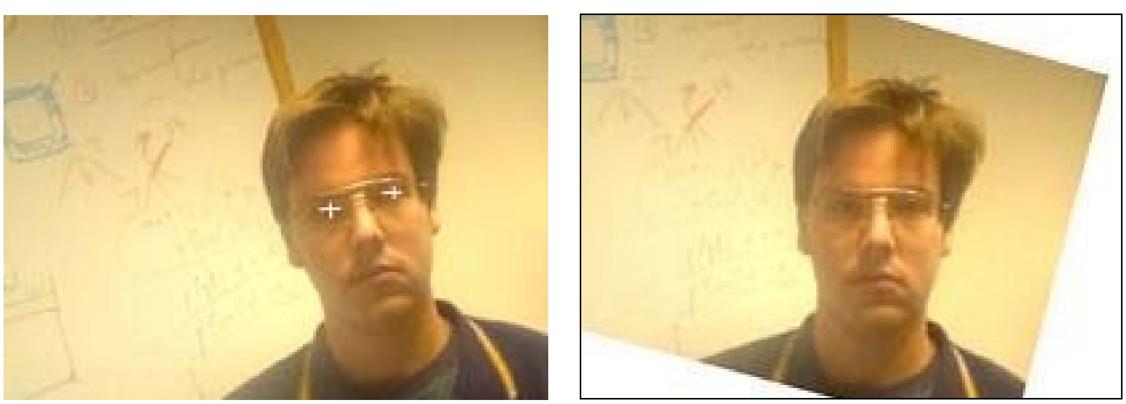
$h(x, y)$. The variance of the model can be estimated from training samples from the camera. Considering the limited resources on a PDA, we used an iteration method (Gao 1997) to implement the deconvolution. The problem then becomes how to solve the following iteration equations when g and h are known,

$$\begin{cases} f_0(x,y) = T[g(x,y)] \\ f_i(x,y) = \lambda g(x,y) + q(x,y) \otimes T[f_{i-1}(x,y)] \quad (i > 0) \end{cases}, \qquad (2)$$

where T are some constrained operators, and

$$q(x,y) = \delta(x,y) - \lambda h(x,y), \; \delta(x,y) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}.$$

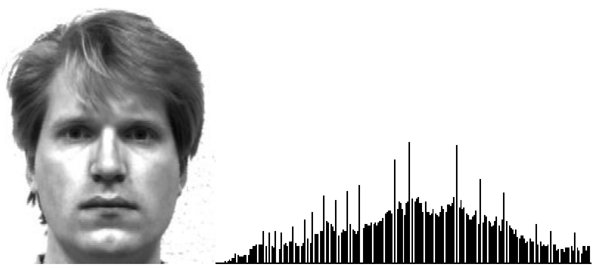
We use normalization methods to deal with variations in scale, head rotation, and lighting. In a PDA-based application, a user has a certain freedom to find an optimal position to capture an image. For example, a user can try to capture a front face instead of a side face. However, we cannot always capture every person's head in an upright position. Furthermore, there is no guarantee that a user can hold the PDA perfectly horizontal. This will result in a head rotation such as the one on the image plane as shown in Figure 2 (a). Such a rotation will decrease the performance of a face recognition system. The alignment can be ascertained through facial features. Figure 2(b) shows the result of normalizing an image by adjusting the scale and rotation angle of the image shown in Figure 2(a) using the locations of the irises.



(a) (b)

Figure 2 An example of rotated head and its recovery

In order to handle changes in lighting conditions, a face image captured under a normal lighting condition is selected and its histogram becomes the standard histogram. For each additional image, we normalize its histogram to match the standardized histogram. After normalization, all input images will have roughly the same mean and variance in their intensity histograms. Figure 3(a) is an example of a light-reference face selected from Yale's face database and its accompanying histogram in Figure 3 (b).



(a) (b)

Figure 3 An example of a reference face and its histogram

2.2.2 Confidence Measure

The confidence measure is another important factor for a face recognition system. Unfortunately, there is no formal way to judge the recognition results. For a face recognition system, the image quality is the primary factor that influences the result. We evaluate the recognition result using a quality estimation model. The model considers several aspects of an image, such as entropy, noise, mean value, and contrast. The entropy measures the complexity of an image, but complexity alone is not a sufficient measurement. For example, a random synthesized image is complex but meaningless to people. Such a case can be detected by measuring noise in an image. We can use the second order derivative to measure noise in the image based on the assumption that the image is piecewise in most areas and the edge is notable. We can further measure the dynamic scope of the image intensity from the mean and the contrast of the image. We can then use the following criterion to measure the quality of an image:

$$Quality = \begin{cases} 10\log\dfrac{E^2}{\mu_n} & \text{if } \mu < \mu_{\min} \text{ or } \mu > \mu_{\max} \text{ or } Con < Th_{Con} \\ 20\log\dfrac{E^2}{\mu_n} & \text{otherwise} \end{cases}, \qquad (3)$$

where E is entropy, μ_n is the noise mean, μ is the mean of the image, and Con is the contrast. The $\mu_{\min}$, $\mu_{\max}$ and Th_{Con} are the thresholds which can be set experimentally.

We can measure the quality of an image using Equation (3). Figure 4 is a noise rich image that was taken under a poor lighting condition. Equation (3) estimates values of 18.6940 and 33.7625 for Figure 4 and Figure 3 (a), respectively.

Figure 4 A noise rich image

2.2.3 Recognition Algorithm

As mentioned in the previous section, a holistic template matching-based method is more suitable for applications performed by a PDA-based face recognition system. Among various approaches, techniques based on Principal Components Analysis (PCA) (Kirby 1990), popularly called *eigenfaces* (Turk 1991, Pentland 1994), have played a fundamental role in dimensionality reduction and have demonstrated excellent performance in some limited scenarios. Many different methods are being developed for face recognition, such as the Euclidean distance (Turk 1991), Bayesian (Moghaddam 1997) and Linear Discriminant Analysis (LDA) (Belhumeur 1997, Etemad 1997, Liu 1998, Swets 1996, Zhao 1998). Unlike PCA that encodes information in an orthogonal linear space, LDA encodes discriminatory information in a linearly separable space such that bases are not necessarily orthogonal. Researchers have demonstrated that LDA based algorithms outperform PCA algorithm for many different tasks (Belhumeur 1997, Zhao 1998). However, the standard LDA algorithm has difficulty processing high dimensional image data. PCA is often used for projecting an image into a lower dimensional space, or so-called "face space," and then LDA is performed to maximize the discriminatory power to compute similarity/dissimilarity scores. In those approaches, PCA plays the role of dimensionality reduction and forms a PCA subspace. Relevant information might be lost due to an inappropriate choice of dimensionality in the PCA step (Zhao 1999). However, LDA can be used not only for classification, but also for dimensionality reduction. We employ a unified LDA/PCA algorithm (Yang 2000) for PDA-based face recognition. The algorithm maximizes the LDA criterion directly without a separate PCA step. This eliminates the possibility of losing discriminatory information due to a separate PCA step. The algorithm is equivalent to the PCA approach in the special case where the total scatter matrix is used in the numerator of Fisher's criterion, and each inner-class variance for each person is zero. This algorithm provides a flexible way for handling a variable number of training samples for each person.

The basic idea of LDA is to find a linear transformation such that feature clusters are most separable after the transformation. This can be achieved through scatter matrix analysis (Fukunaga 1990). For an *M*-class problem, the between-class and within-class scatter matrices S_b and S_w are defined as:

$$S_b = \sum_{i=1}^{M} \Pr(C_i)(\mu_i - \mu)(\mu_i - \mu)^T = \Phi_b \Phi_b^T, \tag{4}$$

$$S_w = \sum_{i=1}^{M} \Pr(C_i)\Sigma_i = \Phi_w \Phi_w^T, \tag{5}$$

where $Pr(C_i)$ is the prior probability of class C_i; μ is the overall mean vector; and Σ is the average scatter of the sample vectors of different classes C_i around their representative mean vector μ_i.

The class separability can be measured by established criterion. A commonly used criterion is the ratio of the determinant of the between-class scatter matrix of the projected samples to the within-class scatter matrix of the projected samples:

$$J(A) = \arg\max_A \frac{\left|AS_b A^T\right|}{\left|AS_w A^T\right|}, \tag{6}$$

where A is an $m \times n$ matrix with ($m \leq n$). The most frequently used LDA algorithm in practice is based on simultaneous diagonalization (Fukunaga 1990). The basic idea of the algorithm is to find a matrix A that can simultaneously diagonalize both S_w and S_b. The simultaneous diagonalization algorithm involves matrix inversion. To our knowledge, most algorithms require that the within-class scatter matrix be S_w non-singular, because the algorithms diagonalize S_w first. Such a procedure breaks down when the within-class scatter matrix S_w becomes singular. This can happen when the number of training samples is smaller than the dimension of the sample vector. This is the case for most face recognition tasks. For example, a small sized image of 64x64 pixels turns into a 4096-dimensional vector. The solution to this problem is to perform two projections (Belhumeur 1997, Etemad 1997, Swets 1996, Zhao 1999):

- Perform PCA to project the n-dimensional image space onto a lower dimensional sub-space;
- Perform discriminant projection using LDA.

The null space of S_w may contain useful information if the projection of S_b is not zero in that direction, but the null space of S_b can be safely discarded. Almost all the LDA algorithms diagonalize S_w first. This requires S_w to be non-singular because the procedure involves inversion. However, the simultaneous diagonalization algorithm can start from either matrix of two symmetric matrices. In other words, we can diagonalize S_b first instead of S_w. If we begin diagonalization from S_b, we need to keep S_b non-singular. It will not lose any useful information if we remove the null space from S_b. This leads to a direct LDA algorithm to obtain an exact solution without a separate dimensionality reduction step.

Direct LDA Algorithm for Face Recognition (Yang 2000)

1. Remove the null space from S_b and diagonalize S_b .

Do an eigen-analysis of $\boldsymbol{\Phi}_b^T\ \boldsymbol{\Phi}_b$ (an M x M matrix). Then sort eigenvectors in decreasing order of their corresponding eigenvalues. Map each eigenvector x of $\boldsymbol{\Phi}_b^T\ \boldsymbol{\Phi}_b$ onto $v = \boldsymbol{\Phi}_b x$, which is the eigenvector of S_b. Normalize the v's and write them down side by side to get V, such that

$$V^T S_b V = \Lambda , \tag{7}$$

where $V^TV=I$, Λ is diagonal matrix sorted in decreasing order. Discard those with eigenvalues sufficiently close to 0 (below ε). Let Y be the first m columns of V, which gives us

$$Y^T S_b Y = D_b, \tag{8}$$

2. Diagonalize S_w. Let $Z = YD_b^{-\frac{1}{2}}$, we have

$$(YD_b^{-\frac{1}{2}})^T S_b (YD_b^{-\frac{1}{2}}) = Z^T S_b Z = I, \tag{9}$$

   Diagonalize Z^TS_wZ by eigen analysis:

$$U^T Z^T S_w ZU = D_w, \tag{10}$$

where $U^TU = I$, D_w may have zeros in its diagonal. Since the objective is to maximize the ratio of between-scatter against within-scatter, those eigenvectors corresponding to the smallest eigenvalues of D_w are the most discriminative dimensions. We can optionally pick only the most discriminative several dimensions. In fact, we can sort the diagonal elements of D_w in a decreasing order and discard some eigenvectors with small eigenvalues.

3. The LDA transformation is:

$$A = (ZU)^T, \tag{11}$$

Matrix A diagonalizes both the numerator and the denominator of Fisher's criterion:

$$AS_w A^T = D_w,\ AS_b A = I.$$

Finally, we can manipulate the data into a more spherical shape, which is done with the transformation:

$$X^* \leftarrow D_w^{-\frac{1}{2}} AX. \tag{12}$$

Figure 5 HP Jornada pocket PC with HP pocket camera

3. A PROTOTYPE SYSTEM

Several palm size PDAs, such as Compaq iPAQ and HP Jornada, are now available with powerful CPU's (over 200MHz) and large amounts of memory (32 ~ 64 MB). These devices are powerful enough to support a PDA-based face recognition system. We have developed a prototype system using a HP Jornada Pocket PC with a HP pocket camera as shown in Figure 5. The camera can be simply plugged into the CF type I slot of the pocket PC. It can capture a 640x480 resolution image with a swivel lens that can be rotated 180 degrees. The system consists of two parts: an interface and a face recognition module. The interface lets a user select a face and facial features for face alignment. The face recognition module contains image preprocessing and face recognition algorithms discussed in the previous section. The hairstyle can play an important role in a global matching algorithm such as PCA and LDA. In order to eliminate influence of the hairstyle, the system removes hair by a facial area mask based on locations of irises.

3.1 SYSTEM DEVELOPMENT

The system runs on Windows CE environment. We developed the system using Microsoft embedded development tools on a desktop computer and then download the executable code to a pocket PC.

Simple usage involves a user capturing an image of the subject. Because of the non-intrusive nature of a PDA, this can be done unbeknownst to the subject. The user then circles the face of the person to be recognized. This brings up a zoomed in version of the person's face where the user then points to the subject's pupils. Having the user identify the face and irises has many advantages. A human has near 100% accuracy in recognizing contours of the face and location of the eyes, tasks for which even the best software heuristics are still far from perfect. In this current prototype system, the face is decomposed using the method explained earlier and compared to the face database, which is stored on the PDA. Depending on the number of dimensions kept, the PDA is capable of storing thousands of faces. The training of the face database was performed on a desktop. The resultant scatter matrix Y* is the only resulting information that needs to be exported to the PDA. A simple form of compression is used on this purely numerical data. This allows thousands of faces to be represented by about 1 MB of a data file, which is well within the bounds of current PDA storage capabilities. Future incarnations of this project might involve reading the dimensional data from a centralized database via a wireless connection. This would permit a nearly infinite number of faces in the database.

Currently, after the system successfully identifies the face, it reports the name of the person with a confidence level in the recognition. Upon further query, the system can also report *n*-best matching faces.

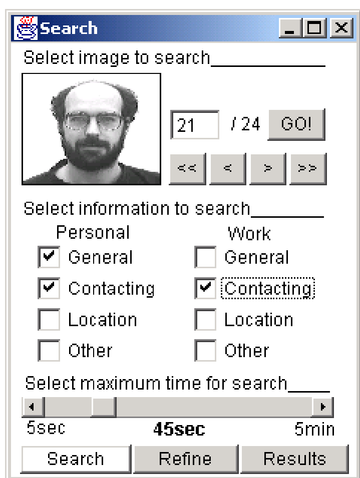


Figure 6 The interface for PDA-based face recognition system

3.2 USER INTERFACE DESIGN

In order to better visualize the retrieved information, we have designed a user interface to allow a user to specify the amount of information that he/she would like to retrieve. We first did a paper design and interviewed 10 people about various interface architectures.

After a few iterations of interviews, we have implemented a JAVA-based interface based on a favorite design by a majority of interviewed people as shown in Figure 6

4. USER STUDIES

In order to determine the level of user acceptance to the proposed system, we have performed two different user studies: a user survey and a user acceptance test. In the first study, we designed an online survey to collect data from users about their usage of camera equipment and their attitude towards photography in public. In the second study, we tested user acceptance by taking pictures, using a PDA and attached camera in public places. We report the results from these two studies below.

4.1 ONLINE SURVEY

We conducted a study of how people use their camera equipment and collected data through an online survey. The survey is divided into 4 sections. In the first section, we try to discover the different kinds of devices that people have. In the second section, we try to find out how these devices are used. In the third and fourth sections, we try to find out how often and by which means people publish the photographs that they take. In addition, there are questions in these sections to find out about users' attitudes toward public photography. There are a total of 24 questions in the survey and the pilot tests showed that it took under 10 minutes to complete it. Some sample questions from the survey are as follows:

- Do you own a PDA (Personal Digital Assistant)?
- Have you ever-photographed celebrities and/or other rich and famous people?
- Have you ever accidentally photographed people who you do not personally know (For example, you may be photographing a sunset on a beach and happen to capture other people there)?
- Do you publish your images on the web?

We collected data from a total of 157 people ranging in age from 19 to 69 years old. Most of these people are students and researchers at the university. Here are some interesting results from the survey:

- Nearly everyone (151 out of 157) had access to a camera; 103 out of 149 people had used it used at least once in the previous month.
- While many people owned a PDA (70 out of 157) only 2 of them used PDA cameras.

- 99 out of 157 had a personal website and 85 publish their images on the web.

In the survey, more than half the people ($p<2.29e^{-29}$ for sign test, n+=142, n-=12) reported that they had unintentionally photographed other people (as in background etc.).

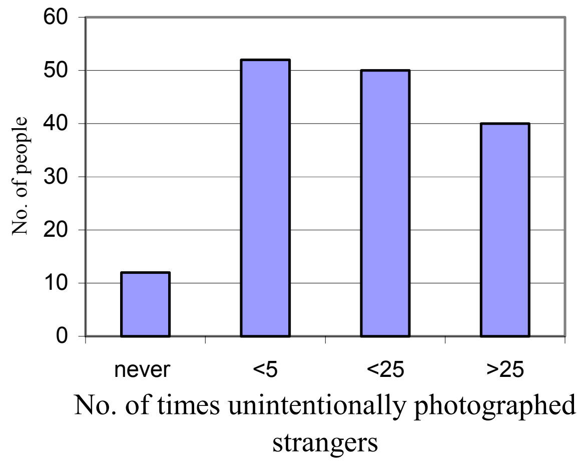
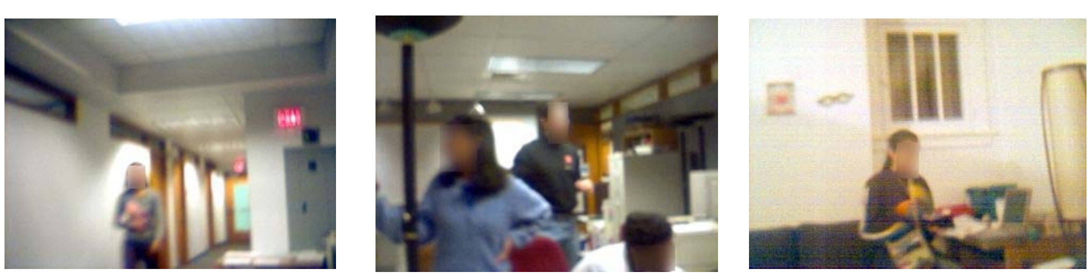


Figure 7 Number of people vs. Number of times they have unintentionally taken photographs of strangers.

From the data we learned that more people were concerned than pleased or angry when someone would take their picture without their knowledge or consent.

4.2 USER ACCEPTANCE TEST

To study how people react when they are photographed in public, we took 205 photographs of people in different places such as shopping malls, building lobbies, conference rooms, cafeterias, restaurants, stairwells, corridors, public streets, and department parties. We used an HP Jornada Pocket PC with an HP Pocket Camera. All these photographs were taken from a distance of 20 feet or less. We blurred people's faces immediately to remove their identities, as shown Figure 8.



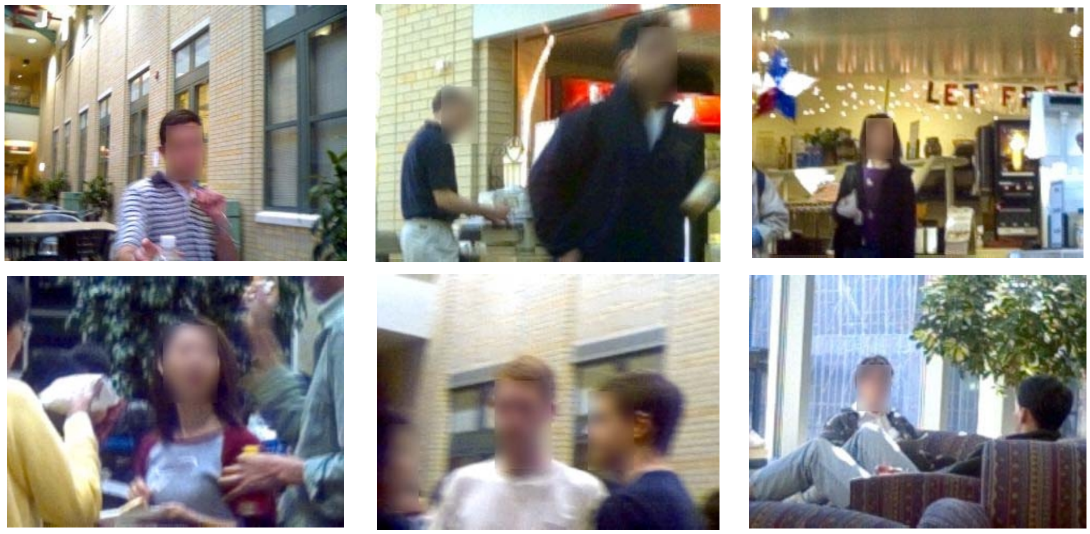


Figure 8 Sample images

During the test, some people noticed the green blinking light on the front of the camera that lights up when a photograph is taken, but their reaction varied from indifference to mild curiosity. Out of a total of 205 pictures taken, only one curious person came to us to ask us what we were doing.

5. EXPERIMENTAL RESULTS

We have performed many different experiments to study image processing and face recognition algorithms, and to demonstrate the feasibility of these processes on current hardware. We used publicly available databases for algorithm evaluations to make it comparable to other research. We used images taken by the hardware of our prototype system to demonstrate the feasibility of the system, and we will make those images available to interested researchers.

5.1 IMAGE PROCESSING

We have tested the effectiveness of the image processing algorithms. Figure 9 is an image taken by the HP pocket camera. The image is blurred because it is unfocused. We applied Equation (2) to enhance the quality of the image. Figure 10 shows the results of the enhanced image after 30 passes with $\sigma = 0.5$ and $\sigma = 0.3$ separately. The improvements are obvious.

Figure 9 An example of a blurred image

(a) σ = 0.5 (b) σ = 0.3

Figure 10 The enhanced images after 30 iterations

5.2 RECOGNITION

We have evaluated recognition accuracy of the recognition algorithm using two different databases: the ORL (Olivetti-Oracle Research Lab) database (ORL Website) and the Yale Face Database. The ORL dataset consists of 400 frontal faces: 10 tightly-cropped images of 40 individuals with variations in pose, illumination, facial expression (open/closed eyes, smiling/not smiling), and accessories (with/without glasses). We reduced the number of dimensions to 39 and tested recognition accuracy versus number of training samples. We varied the number of training samples from 1 to 9. For each test, *n (n = 1,2,…,9)* randomly-selected images for each individual in the dataset were placed in the training set, and the remaining images were used for testing. Ten runs for each of *n* samples were performed with different, random partitions between training and testing images. The best recognition rate for the new algorithm is about 95% when *n=9*, which is compatible with the best result obtained by other researchers on the same test set using different algorithms.

The new LDA algorithm has the capability of handling a small size of training data. We have proved that the algorithm is equivalent to the eigenface (PCA) approach in the special case where each person has only one sample in the training set. The feasibility of the new algorithm has been demonstrated by experimental results (Yang 2000). We used the Yale face database to test the robustness of the system against variation. The Yale face database contains images from 15 different people. Each person has 11 pictures with variations in lighting condition, emotion, and with/without glasses. We used the normal face as the training sample for each person and the rest of 10 pictures as the testing samples. Figure 11 depicts the results. The vertical axis of Figure 11 is number of people who have been recognized correctly. The horizontal axis of the Figure 11 indicates changes of emotions, lighting conditions, and with/without glasses. The results indicate that the system is robust against variations in emotion changes. However, it performs poorly for side lighting source changes. This is an inherent problem of a holistic template matching-based method, because the algorithm encodes the intensity of the whole face.

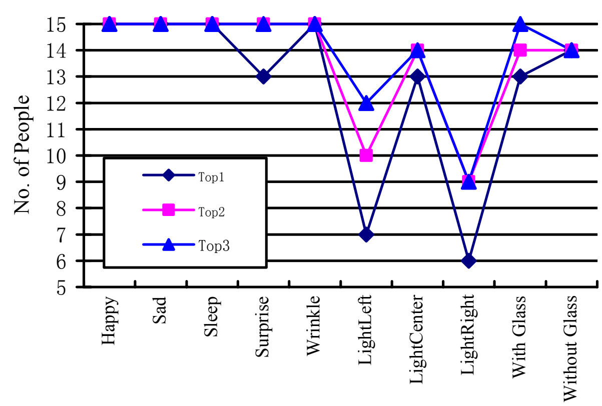


Figure 11 Experiments of the Yale face database

5.3 POCKET CAMERA

We have tested the face recognition system using the images taken from the HP pocket camera used for the prototype system. Considering the applications of the system, we performed two specific tests: effective distance of the camera and recognition accuracy. To test the effective distance of the system, we took pictures at different distances and then performed the recognition task. We took images at 10, 15, 20, 25, and 30 feet from the subject, both indoors and outdoors. The system performed very well for images within 20 feet, and worked fine for images taken at 25 feet. But the performance decreased significantly after 25 feet. This can be illustrated numerically using Equation (3). Table 1 shows the evaluation results of a set of images at different distances.

Table 1 The evaluation results of a set of images

	Mean	Contrast	Noise	Entropy	**Quality**
10 feet	120	66	7.5499	4.5719	**20.3664**
15 feet	110	68	8.1118	4.5048	**18.3396**
20 feet	122	52	7.8167	4.2939	**17.1624**
25 feet	128	52	7.4440	4.2104	**17.3541**

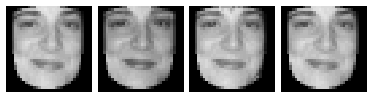
To test recognition accuracy, we collected data from our lab using this system. We collected a total 116 pictures for 24 people under various lighting conditions and in different poses. We randomly selected one image from each person for training and the rest for testing. The recognition accuracy was 81.25%, if the system considers only top one choice. In a PDA-based based face recognition task, a user can help to make decision from an *n*-best list provided by the system.

We have also tested flexibility of the system. A good example is the application of the recognition system to the missing federal intern Chandra Levy. In order to investigate the case, Washington police composed pictures showing what she might look like. We used the images found in the Internet as training images (Figure 12) and the images published by the police (Figure 13) as the test images. Figure 14 shows the masked faces from Figure 12 and 13. It is clear that the mask has effectively removed any influences of the hairstyle. We added the training set to our database and then tested the system. As the result, all test images have been recognized correctly by the system. This example suggests another application of the proposed system for law enforcement.

Figure 12 Chandra Levy pictures found from Internet

Figure 13 The composed pictures by Washington police

(a) The training samples



(b) The testing images

Figure 14 The masked faces

6. CONCLUSIONS

We have proposed to use the human face as a lookup index for extending human memory and/or information retrieval. Technological developments in palm-size PDA's have made it possible to take face recognition systems in a new direction. A PDA-based face recognition system can take advantage of having a human-in-the-loop and enhances human capability of recognizing other people. We have presented techniques for developing a PDA-based face recognition system and demonstrated the feasibility of such a system through user studies and experiments. The applications of the system include human-computer interaction, information retrieval, and law enforcement.

7. ACKNOWLEDGMENTS

The authors would like to thank our colleagues in the Interactive Systems Labs for their inspiring discussions, support, and help on data collection. We would like to thank Elaine M. Newton for her valuable comments on the draft of this report, and Rahul Culas for his help on some of user tests.

8. REFERENCES

[Belhumeur 1997] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection *IEEE PAMI*, 19(7):711-720, 1997.

[Chellappa 1995] R. Chellappa, C. Wilson, and S. Sirohey. Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83(5):705-740, 1995.

[Clarkson 2001] B. Clarkson and A. Pentlan, Predicting Daily Behavior via Wearable Sensors, working paper, MIT Media Lab, July, 2001.

[Etemad 1997] K. Etemad and R. Chellappa. Discriminant analysis for recognition of human face images. Journal of the Optical Society of America A, 14(8):1724-1733, 1997.

[Fukunaga 1990] K. Fukunaga. Introduction to Statistical Pattern Recognition (Second Edition). New York: Academic Press, 1990.

[Gao 1997] W. Gao and X. Chen, A stochastic approach for blurred image restoration and optical flow compensation on field image sequence. Journal of Computer Science and Technology, 12(5):385-399, 1997.

[Kirby 1990] M. Kirby and L. Sirovich. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE PAMI*, 12(1):103-108, January 1990.

[Liu 1998] C. Liu and H. Wechsler. Enhanced fisher linear discriminant models for face recognition. Proceedings of Fourteenth International Conference on Pattern Recognition, vol. 2, pages 1368-1372, Brisbane, Qld., Australia, 1998.

[Mann 1996] S. Mann. Wearable, tetherless computer--mediated reality: Wearcam as a wearable face-recognizer, and other applications for the disabled. TR 361, MIT Media Lab, Cambridge, MA, USA, 1996.

[Moghaddam 1997] B. Moghaddam and A. Pentland. Probabilistic visual learning for object recognition. *IEEE PAMI*, 19(7):696-710, 1997.

[ORL Website] ORL web site: *http://www.cam-orl.co.uk*

[Pentland 1998] A. Pentland, Wearable intelligence, *Scientific American Presents*, 9(4), pp. 90-95, 1998.

[Pentland 1994] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. *Proceedings of the 1994 Conference on Computer Vision and Pattern Recognition*, pages 84--91, Seattle, WA, 1994.

[Samal 1992] A. Samal and P. Iyengar. Automatic recognition and analysis of human faces and facial expressions: A survey. *Pattern Recognition*, 25(1):65-77, 1992.

[Swets 1996] D. Swets and J.Weng. Using discriminant eigenfeatures for image retrieval. *IEEE PAMI*, 18(8):831--836, August 1996.

[Turk 1991] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):72--86, 1991.

[Yale Face Database] Yale face database:
http://cvc.yale.edu/projects/yalefaces/yalefaces.html

[Yang 2000] J. Yang, H. Yu, W. Kunz, An efficient LDA algorithm for face recognition," *International Conference on Automation, Robotics, and Computer Vision (ICARCV'2000)*, Singapore, December, 2000.

[Zhao 1998] W. Zhao, R. Chellappa, and N. Nandhakumar. Empirical performance analysis of linear discriminant classifiers. *Proceedings of the 1998 Conference on Computer Vision and Pattern Recognition*, pages 164-169, Santa Barbara, CA, 1998.

[Zhao 1999] W. Zhao, R. Chellappa, and P. Philips. Subspace Linear Discriminant Analysis for Face Recognition. Technical Report CAR-TR-914. Center for Automation Research, University of Maryland, 1999.