

# Compact Parallel Optical Correlator for Face Recognition, and Its Application

Kashiko Kodate and Eriko Watanabe  
*Faculty of Science, Japan Women's University*  
Japan

## 1. Introduction

With the ongoing progress in information technology, the need for an accurate personal identification system based on recognizing biological characteristics is increasing demand for this type of security technology, rather than conventional systems that use ID cards or pin numbers. Of all physical features, the face is the most familiar and recognizable, and using it for identification purposes avoids the need for physical contact, thereby also avoiding potential psychological or physical resistance, such as that encountered when trying to obtain fingerprints for example. Face recognition has been studied since the 1970s, with extensive research into and development of the digital processing of facial images. A range of software is already on the market. As a simple and compact recognition system that satisfies the required performance levels, we have implemented a hybrid system based on the optical recognition principle, using a Fourier transform lens.

In contrast to digital recognition, optical analog operations process two-dimensional images instantaneously in parallel using a lens-based Fourier transform function (Kodate Hashimoto, & Thapliya, 1999). In the 1960s, two methods were proposed; the VanderLugt Correlator proposed by VanderLugt ((a)Watanabe & Kodate, 2005), and the joint transform correlator (JTC) (Kodate Inaba & Watanabe, 2002) by Weaver and Goodman. The optical correlator by the Institut national d'Optique in Canada gained the New Product Award at the 1999 Conference on lasers and Electro-optics and the Quantum Electronics and Laser Science Conference. In Japan the conventional JTC was practically implemented by the Hamamatsu Photonics Company for fingerprint image processing. The process speed of optical correlators has steadily improved, however, operational speed, practicality and recognition rate were not fully tested against the fast-improving digitized system.

Against this background, the present group of authors has proposed a new scheme using a multi-beam, multi-channel parallel joint transform correlator (PJTC) as a means of making better use of spatial parallelism through the use of a diffraction-type multi-level zone-plate array (MLZP array) to extend a single channel JTC ((b)Watanabe & Kodate, 2005). Watanabe & Kodate, 2003). Features of the proposed system include extreme robustness, high recognition precision by its pre-processing and high reproducibility by its post-process. The compact and mobile versions are now assembled and named COPaC.

Specifications: 20x24x43 cm 3.6 kg, analysis time of 6.6 faces/s (Kodate Inaba & Watanabe, 2002).

In an attempt to downsize the hard disk, LD is replaced as a light source with a multi-light source module constructed by vertical cavity surface emitting laser (VCSEL) array and MLZPA. In spite of the constraints of the Fourier-transform type process, the speed of an optically controlled, liquid crystal spatial light modulator was accelerated to 30 ms. It is very difficult to achieve optical correlation speeds shorter than 30 ms.

In recent years, devices for optical information processing have been developed. Examples include Ferroelectrics liquid crystal special light modulator (FLC-SLM) and digital micromirror device (DMD) which can enhance the high-speed display (1 kHz-10 kHz), and herald the possibility of accelerating the systemic processing time. Furthermore, a novel holographic optical storage system that utilizes co-axial holography was demonstrated. The VLC is conducive to improvements in speed, given that it can dispense with the optically addressed SLM. In practice, however, comprehensive system design and implementation should be required to make the most of the theoretical potential. We implemented a fully automatic FARCO (dimensions: 33.0x30.5x17.8 cm<sup>3</sup>) ((a)Watanabe & Kodate, 2005), which achieved an operation speed of more than 4000 faces/s using four-channel processing. The correlation filter we used was more accurate than various correlation methods ((b)Watanabe & Kodate, 2005). Based on trial 1:N identification, FARCO achieved low error rates of 1% False acceptance rate (FAR) and 2.3% false rejection rate (FRR).

The recognition time of FARCO is limited to about 1,000frame/s due to the data transfer speed and storage capacity of the random access memory (RAM) used to store digital reference images. The time of data transfer speed is converted from the digital data to optical image data in the optical system. Using the ability of parallel transformation as optical holographic memory, the recognition rate can be vastly improved. In addition, a large optical storage capacity allows us to increase the size of the reference database.

A novel holographic optical storage system that utilizes co-axial holography has recently been demonstrated (Horimai & Tan, 2006). This process can produce a practical and small holographic optical storage system more easily than conventional off-Coaxial holographic system. At present, the system seems to be most promising for ultra high density volumetric optical storage

In this chapter, we present the compact optical parallel correlator for face recognition, and its application. This is a product of lengthy processes, and long-term efforts to overcome several obstacles such as attaining precision in the optical correlation system, operational speed and downsizing. Crucial technologies applied in the system include our unique and extremely precise phase filter and high-speed optical devices that we have been perfecting over many years. Combined with these, a novel system structure and algorithm were proposed and tested in rigorous experiments. Section 2 addresses the basic principle of optical pattern recognition by optical Fourier transform and the importance of phase information in face image. In Section 3, the concept of an optical correlation system for facial recognition and dedicated algorithm is presented. Section 4 provides the design of the correlation filter and an evaluation and comparison of the correlation filters for FARCO, and also tests for evaluating the recognition system, and experimental results of 1:N identifications and so on. A highly precise algorithm using multiple database images for FARCO is introduced in Section 5. A constructed 3 dimensional (3-D) face model is discussed in Section 6. In Section 7, a high security facial recognition system using a cellular phone is presented. A super high-speed optical correlator that integrates an optical correlation technology used in FARCO (fast face-recognition optical correlator) and co-axial

holographic storage system is proposed. Section 9 discusses the future works and Section 10 concludes the chapter.

## 2. The Basic Principle of Optical Correlation

The optical implementation of pattern recognition can be accomplished with either Fourier domain complex matched filtering or spatial domain filtering. Correlators that use Fourier domain matched filtering are commonly known as VLC's. The basic distinctions between them are that the VLC depends on Fourier-domain spatial filter synthesis (e.g., Fourier hologram). In other words the complex spatial detection of the VanderLugt arrangement is input scene independent. The basic optical setup of the VLC type of correlator is depicted in Fig. 1. A prefabricated Fourier-domain matched filter  $H(u,v)$  is needed in the VLC. This section describes the basic principle optical pattern recognition by optical Fourier transform. We also address the importance of phase information in the face image.

Optical correlation is one pattern-matching method used for finding whether there is a correlation between images in the database and input images, using analogue calculation based on the parallelism of light. In this method, all the information of the images can be searched at the speed of light, to determine whether the images belong to the same person. The principle of optical setup was designed for Fourier-domain matched filter, which was applied to our face recognition system. Let us consider a thin convex lens of focal length  $f$ , illuminated by a laser beam. Output facial images set in the real space are Fourier-transformed in the spatial frequency plane. The complex conjugate of this Fourier-transformed image is stored in a matched filter, which is then to be positioned at the optical devices (Fig.1 (b)) and Fourier-transformed again.

On the output plane, the correlation term and convolution term will be displayed. For a more detailed description of this setup, please refer to precedent literature (Goodman & Moeller, 2004. Hecht, 1998).

This section describes the basic principle optical pattern recognition by optical Fourier transform. We also address the importance of phase information in the face image.

### 2.1 Importance of Phase Information

As shown in Fig. 2, 2-dimensional images in real space have amplitude information, which can be singled out by Fourier transformation into amplitude and phase information in spectrum space.

The phase image is defined using absolute values of distribution. Where  $\mathcal{F}$  is Fourier operator, where  $\varphi_F(u, v)$  is phase distribution.

$$F(u,v) = |F(u,v)| \exp[-j\varphi_F(u,v)] \quad (1)$$

Images, of which only phase was Fourier-transformed, can be defined by the following equation.

$$f\varphi(x,y) = \mathcal{F}^{-1}\{|F(u,v)| \exp[-j\varphi_F(u,v)]\} \quad (2)$$

The intensity of the object is represented by amplitude  $F(u,v)$ , while its phase information is displayed by  $\varphi_F(u,v)$ .

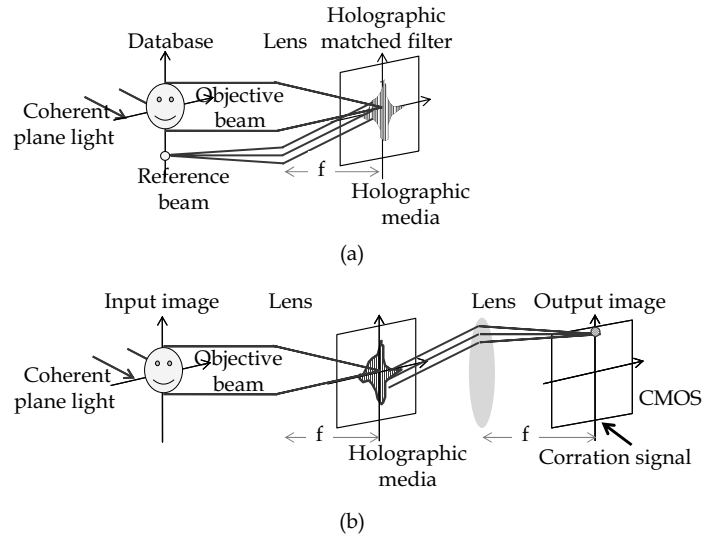


Figure 1. Concept for a matched filtering correlation in FARCO system: (a) Recording a matched filter, (b) Filtering (Optical correlation) process

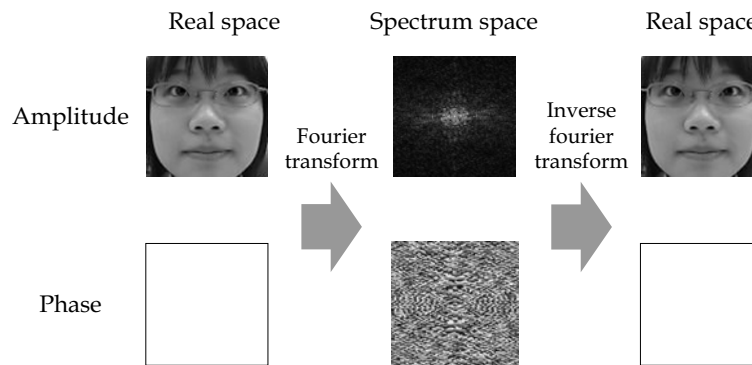


Figure 2. Phase and amplitude information of 2-D images

Figure 2 exhibits both amplitude and phase information, using the facial images in real space. As exemplified by Fig. 2, the phase information is more important than the amplitude information.

Here are Fourier-transformed images of a human being (A) and monkey (B) (Fig.3). Successively, if the amplitude image of (A) in the frequency space multiplied by phase image of (B) is inversely Fourier-transformed, the facial image of (B) is reconstructed. Similarly, by combining phase image (A) and amplitude image of (B), the facial image of (A) was reproduced. This proves that the phase information has contained information of the whole face. Recognition requires this phase information.

The facial image shown on the bottom left is the inverse Fourier transform of the phase spectrum from person (B) combined with the amplitude spectrum of person (A). This image shown on the bottom right is the inverse Fourier transform of the phase spectrum from

person (A) combined with the amplitude spectrum of person (B). Note that the phase spectrum is more dominant than the amplitude spectrum to produce the processed images.

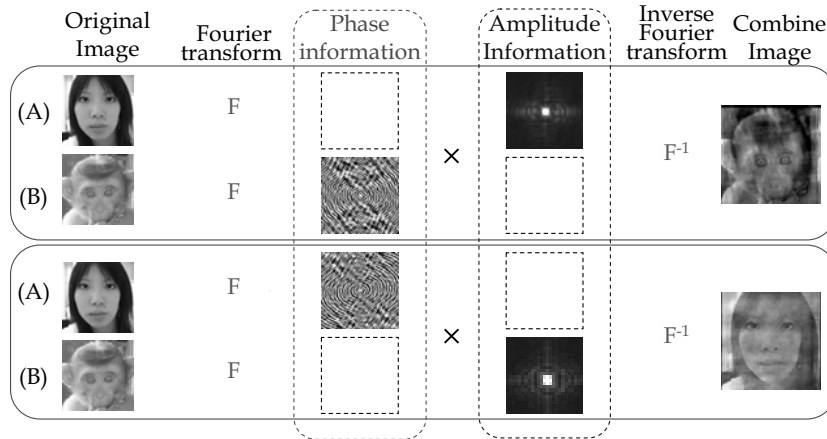


Figure 3. Phase spectrum of facial images

### 3. Algorithm for Our Facial Recognition System

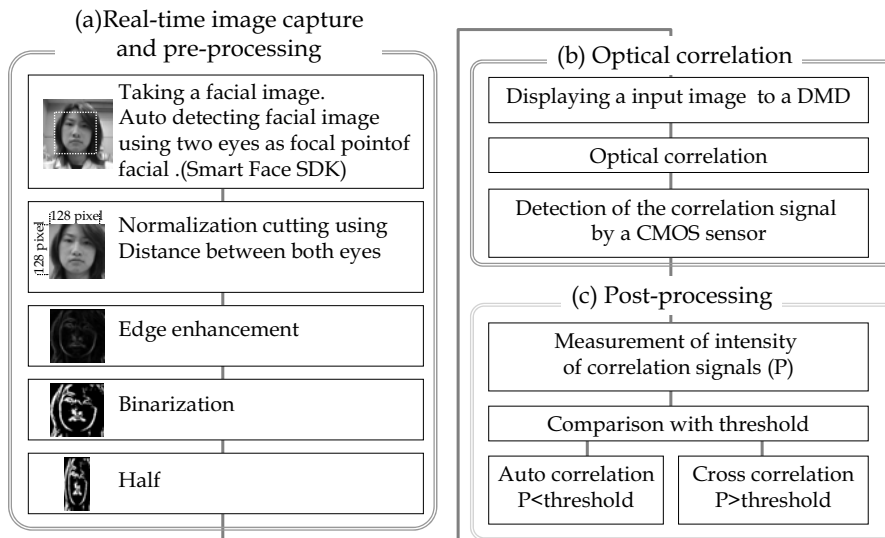


Figure 4. Flow-chart illustrating our hybrid facial recognition system: (a) Real-time image capture and pre-processing, (b) Optical correlation. (c) Post-processing

An algorithm for the FARCO is presented in Fig.4 Under this system, pre- and post-processes using a personal computer (PC) are highly conducive to enhancing the S/N ratio and robustness. This Section describes the concept of the system for facial recognition and the dedicated algorithm.

### 3.1 Signal - to - Noise Ratio Enhancements by Pre-processing – Extraction, Normalization and Angle Adjustments

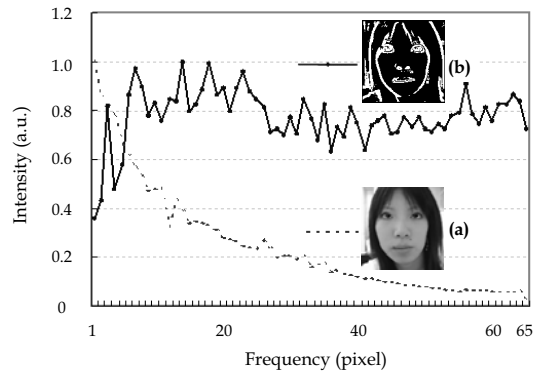


Figure 5. Fourier Spectrum of facial images: (a) Gray scale image (b) Edge-enhancement and binarization image

Facial images were taken automatically by a digital video camera. For automated extraction of input facial images, we can get the two eye points of these facial images using face detecting softwares (development kit by Toshiba Co. or Takumi Co.). By utilizing this software, we were able to detect four points in a facial image (the position of the eyes and nostrils). The size of the extracted image was normalized to 128x128pixel by the center of gravity. For input images taken at an angle, affine transformation was used to adjust the image and normalization, fixing on the position of the eyes. Following on from this, edge enhancing with a Sobel filter, and binarizing, i.e., defining the white area as 20%, equalized the volume of transmitted light in the image. Fig. 5 shows the original facial image, a pre-processed image, and two Fourier power spectra. Grey-scale images have concentrated information in the lower spatial sphere. Pre-processing disperses the image feature components up to the higher spatial frequency, as indicated in Fig. 5. The efficiency of optical utilization is increased in the spatial frequency domain in Figure 5(b). Correlation results with regard to the two images show that the S/N ratio of the pre-processed image increased by 155%, proving its validity. With this increase, the edges of facial features are extracted as the color data fades. Hence, factors such as make-up and skin color have no effects on recognition performance. Glasses without frame or thin frame were also negligible, since their edges are hardly captured. Accordingly, robustness improves. We have shown previously that the binarization of the input images with appropriate adjustment of brightness is effective in improving the quality of the correlation signal ((a)Watanabe & Kodate, 2005).

### 3.2 Individual Authentication and Threshold-value Determination in Post-process

A biometric recognition system can operate under two different modes: 1:N identification or 1:1 verification. Here the 1:N identification mode is applied, under which one's biometric pattern is calculated from his/her biometric features and examined against a database containing N images. Although the other method 1:1 can also be adopted, given the high speed of our system, the 1:N identification system was chosen.

The False Acceptance Rate denotes the probability that a biometric system will incorrectly identify an individual, or will fail to reject an impostor. For a positive (verification) system, it can be estimated from: (the number of false acceptances) / (the number of impostor verification attempts). Conversely, False Rejection Rate (FRR) is the probability that a biometric system will not identify an individual, or will fail to accept a person in the database. For a positive (verification) system, it can be estimated from: (the number of false rejections) / (the number of times it fails to verify).

In practical application, the threshold value has to be customized by the application. The threshold value varies with its security level; depending on whether the system is designed to reject an unregistered person or permitting at least one registered person. We have to decide the optimum threshold value using the appropriate number of database images based on the biometrics guideline for each application. In this paper, the threshold value is fixed where FRR and FAR are lowest.

We determined the sample number was according to the definition of one of the accuracy evaluation authorities, National Biometric Test Center (Japan). The error margin  $P$  in the sample number  $N$  was given by the following equation, under the reliability 95%.

$$N = 3/P \quad (3)$$

For example, where  $P=0.01$ , 300 samples (i.e. persons) are required. The sample number determined was 300 persons; the facial images of 240 women (20-50 years old) were taken on the same day, while images of 60 men were pre-processed. One hundred and fifty were registered, while the second half was unregistered. The correlation value (where  $i$  is the entry number of test image and  $j$  is the reference number) is stored in a memory and the maximum value,  $P_{1mCoaxial}$  searched. Then, all the correlation values are normalized by  $P_{imax}$ . Here, we define the comparison value, which is used as threshold for 1:N identification,  $C_i$  by Equation(4). (Inaba Kodate & Watanabe, 2003)

$$C_i = \Sigma P_{ij}/P_{imax}-1/(N - 1) \quad (4)$$

Using the devised correlation filters, simulation experiments were conducted under the assumption of 1:N identification.

#### 4. Fast Face-Recognition Optical Correlator

We have developed a face recognition system named FARCO, which has three different configurations depending on its recognition rate shown in Fig.6. This is an improvement on one-to-one ID recognition system, which requires little calculation time. The FARCO system is a hybrid optical correlator that integrates an optical correlation technology and digital database. FARCO can be applied to several hundreds of images, with its operation speed of 1000 to 5000 faces per second.

In order to correspond to greater demand, the S-FARCO (Super fast face-recognition optical correlator) system is used. The S-FARCO is equipped with a holographic optical memory, which could store and process information optically. This enables optical correlation without decoding information in the database, which greatly reduces processing time (more details in Section 8). This Section will present the correlation filter, its optical setup and experimental results focus on FARCO.

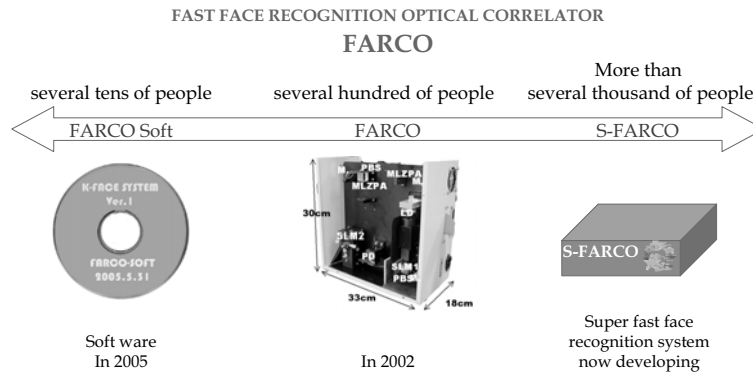


Figure 6. Three different ramifications of developed Optical Correlator FARCO

#### 4.1 Design of a Correlation Filter

This Section presents a novel filtering correlation for face recognition is introduced, using phase information with emphasis on the Fourier domain. Comparing it with various correlation methods, we evaluate our own filtering correlation method.

The performance of the filtering correlation was evaluated through one-to-N identification with a database of 300 front facial images. The size of the database was formulated according to the guideline of biometrics authentication. The facial images for input and reference are taken by the cellular phone (DoCoMo D506is). Students take their own facial images three times with the cellular phone. The database contains ten facial images for each person. There are three types for the experimental condition: (1) classical correlation, (2) phase-only correlation, and (3) filtering correlation. ((b)Watanabe & Kodate, 2005)

##### (1) classical correlation

The general  $f$  and  $h$  correlation equation is given as follows Equation (5).

$$g(x,y) = \iint f(x',y')h(x'-x,y'-y)dx'dy' \quad (5)$$

In this paper, we call Equation (5) the classical correlation.

##### (2) phase-only correlation

It performs the correlation between two signals,  $f(x,y)$  and  $h(x,y)$ , using the Fourier plane relationship

$$g(x,y) = \mathcal{F} [F(u,v)H^*(u,v)] \quad (6)$$

in which \* denotes conjugate.  $F$ , the Fourier transform of the operator, while is the Fourier transform of one signal  $f(x,y)$ ,  $H^*(u,v)$  is the correlation filter corresponding to the other signal, and  $u$  or  $v$  stands for the two vector components of the spatial frequency domain. The classical matched filter for a signal  $h(x,y)$  was defined as  $H^*(u,v)$ . In polar form it can be represented as follows:

$$H^*(u,v) = |H(u,v)| \exp\{-i\varphi(u,v)\} = |H(u,v)| [\cos\{\varphi(u,v)\} - i\sin\{\varphi(u,v)\}] \quad (7)$$

The phase-only filter, which acts in a similar way to the matched filter, is derived from Equation (5). By setting every amplitude at the number equal to 1 or alternatively by multiplying it by  $1/H(u,v)$ , we obtained the phase only filter:

$$H_p(u,v) = \exp\{-i\varphi(u,v)\} \quad (8)$$

where  $p$  stands for phase. Classical correlation has a high correlation signal without a correct peak position. In the case of phase-only correlation, the maximum value is obtained only at the peak position.; at all other points, smaller values approximate zero. This correlation is highly precise. (Horner & Gianino, 1984. Bartelt, 1985)

### (3) Filtering correlation

We optimized correlation filter with emphasis on Fourier domain taking the following points into account: (i) carrier-spatial frequency should be contained within the minimum frequency range of facial characteristics. ((b)Watanabe & Kodate, 2005)

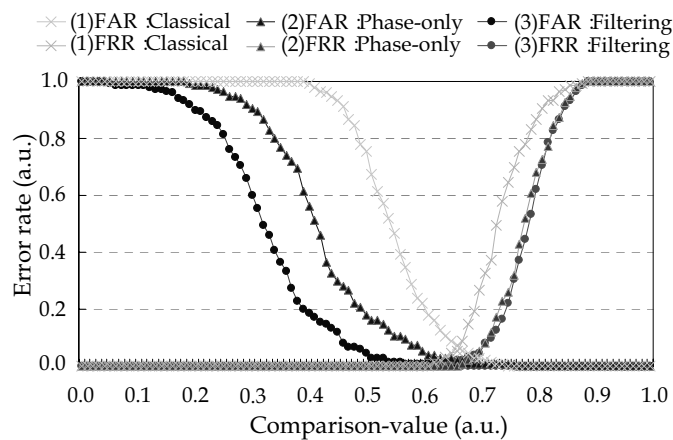


Figure 7. Error rate of three kinds of correlation

In those three different types of correlation methods, experimental error rates are shown in Fig.7 and Table 1. If the intensity exceeded a threshold value, the input image would be regarded as a match with a registered person. Error rates divided by the total number of cases were given by the False Rejection Rate (FRR) and False Acceptance Rate (FAR). With the threshold value set at optimum value (arbitrary units), the FAR and FRR are shown Figure 7. Error rates are plotted on the vertical Coaxial and comparison value on the horizontal Coaxial. As the results show in Table 1, EER has improved by 0.7%. Where FAR is 0%, FRR has improved by 2.0%.The results clearly shown that the designed correlation filter has remarkably high precision.

This filter, based on optical Fourier transform, facilitates parallel operation without changing its configuration. (Watanabe & Kodate et al., 2006)

	FRR[%]	FAR[%]	EER[%]
Classical correlation	62.7	0.0	7.3
Phase-only correlation	42.7	0.0	1.3
Filtering correlation	2.0	0.0	0.7

Table 1. Results for three kinds of correlation

#### 4.2 Optical Setup of FARCO

The optical devices installed into the FARCO are displayed in Table 2. The facial image display SLM in the database is composed of an FLC-SLM, with the capacity of high-speed display (2500 frame/s). The FLC-SLM, featuring a reverse display, was constructed with an LD of wavelength 635nm as a light source, and driven by a pulse, flickering at the positive values. Nematic Liquid Crystal SLM (NLC-SLM) is used as a spatial optical modulator for the display of the correlation filter. FARCO is employed the filter using filtering correlation as shown in §4.1. Moreover, the MLZP (Orihara Klaus & Kodate et al., 2001) is used for optical parallel processing. As a receiving device, a photo diode, divided into four parts, was used. The diode can simultaneously receive four signals at the speed of 20 MHz. An isolator was installed as a reflection-proof instrument. The optical setup of the FARCO is shown in Fig. 8. ((a)Watanabe & Kodate, 2005)

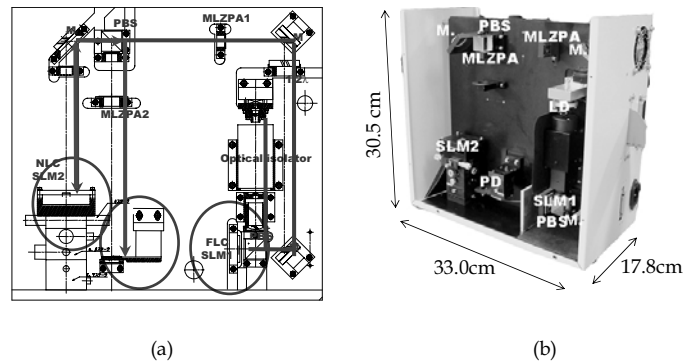


Figure 8. Fabricated FARCO. (a) Optical set-up. M.: Mirror, H.M.: Half mirror, P. : Polarizer, A. : Analyzer, L. : Lens, PBS : Polarized Beam Splitter, C.L. : Collimate Lens, SLM : Spatial Light Modulator, MLZPA. : Multi Level Zone Plate Array, LD: Laser diode, PD: Photo diode

Light source	Wave length	635nm(LD)
	Power	30mw
Ferroelectrics LC-SLM For Databbase (Displaytech)	Pixel number	1280x768pixels
	Pixel pitch	13.2mm
	Operation speed	2.5kHz
Nematic LC-SLM For Matched filter (Boulder nonlinear systems)	Pixel number	512x512pixels
	Pixel pitch	15mm
	Feature	2pi@780nm
MLZPA	Channel number	1 or 4
	Focal length	300 or200mm
	Aperture size	3.26mm
	Phase levels	8
Detector	Operation speed	20MHz
	Active area	10x10mm <sup>2</sup>

Table 2. Specifications of the optical devices for FARCO

### 4.3 Flowchart of FARCO System

A flow chart for optical correlation with the FARCO is presented in Fig. 9. Firstly, pre-processing was carried out within less than 200ms, covering the extraction of facial images through to the calculation of the correlation filter. On receiving final signals from the pre-process stage, NLC-SLM started up and showed the correlation filter obtained on the SLM in the correlator (see (2)-1 in Fig. 9). Initiated by the start-up signal from the NLC-SLM, the light pulse source begins to operate ((2)- 2, 3). Subsequently, the moment images in the database, stored in advance in an FLC-SLM board RAM of a control computer, are set for calculation, the instant correlation begins as each image is shown at the speed of 1000frame/s. Correlation signals through the filter were captured by a detector, and the recorded intensity is classified as a resemblance level ((2)- 4, 5). Each loop per optical correlation required 1 ms, although four parallel optical correlations with a parallel Fourier transform device, MLZP, only necessitate 4000 faces/s. In the final stage, post-processing the intensity values of these optical signals yielded final outcomes, i.e. recognition rates. In the case of the one-to-N identification experiment,  $C_i$  was calculated after N-loops, while values from every loop in the memory were recorded. Being used as an internal LAN on the university campus, this face recognition system is accessible to anyone with a camera and input software.

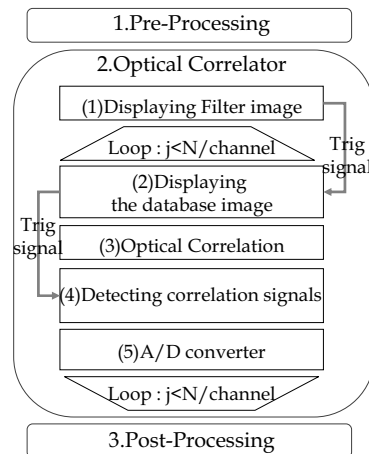


Figure 9. Flow chart of FARCO system

### 4.4 Experimental Results with FARCO

#### 4.4.1 N identification with 4 channels

1:N identification experiments were conducted to examine the recognition performance of the FARCO, using the sample of 300 individual facial images. The FARCO system processes 1000 images per second by one channel. Thus it takes 0.15 seconds for the database with 150 registered person images. The ROC curve, acquired by the experiments with 4 channels, is shown in Fig. 10. This curve represents the FAR of 1.3% and the FRR of 2.6%, which is the lowest error rate of all with the FARCO. Relying on FAR values, the level of security system applicability varies. At the FAR below 1%, PC login and entry to communal houses are possible. With 1% FAR, FRR becomes 3.6%, which is within possible range of availability, in

compliance with the biometrics evaluation. Results of recognition experiments on a database containing over 4000 images are presented in Fig 11. This database has 80 people with 50 multiplex images per person. Operationalization at 4 ch enabled correlation with 4000 images.

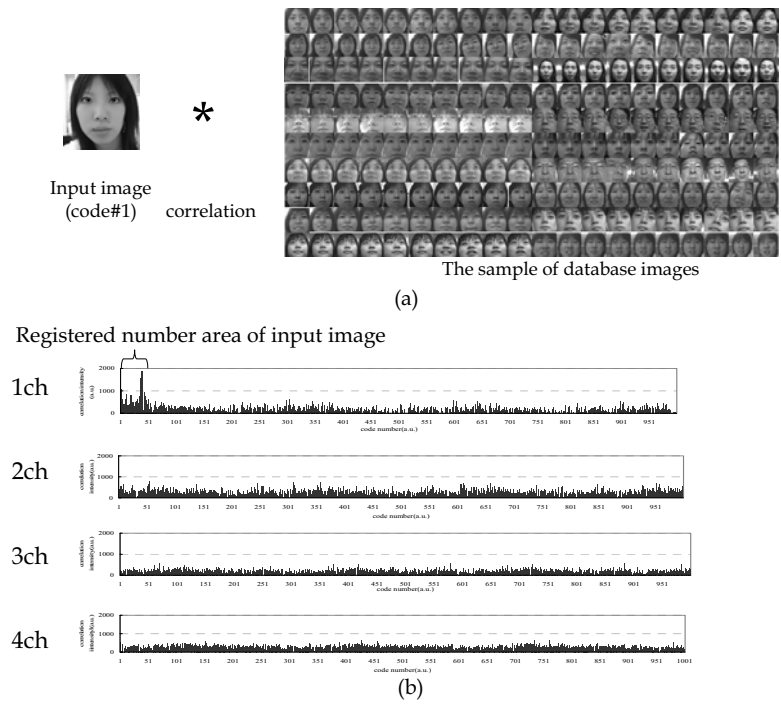


Figure 10. Evaluation accuracy of FARCO according to the biometrics guideline with database images of 300 persons. The lowest error rates of the two values were recorded for 300 data, with FAR and FRR error rates of less than 1.3% and 2.6% respectively

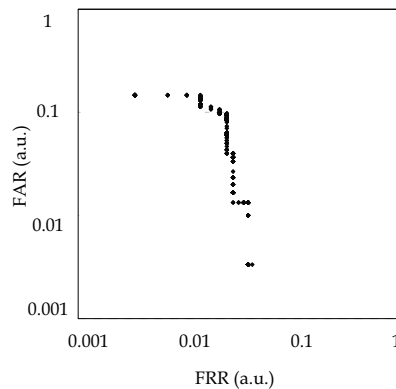


Figure 11. Evaluation accuracy of FARCO according to biometrics guideline using 300 people database images

#### 4.4.2 10 channel Experiment Results

We examined the correlation using multi-object reference images. When the correlation filter is applied to multi-object images, there are problems in that optical intensity is reflected by transmission intensity and diffraction angles. By using all 1280x768 pixels of SLM pixels, we can implement parallel operations of more than 10 channels. In this experiment, however, we designed and performed experiments limited to 10 parallels, for the sake of checking the accuracy of each channel. Fig.12 (c) illustrates 10 ch multi-object image auto-correlation intensity using the designed optical correlation filter. The sharp correlation value was acquired by 10ch as well as by 1ch. The design arrangements of multi-object images are based on reference ((b)Watanabe & Kodate , 2005). This result shows that we can implement 10,000 faces/s, if a 10ch photo detector array is installed in the FARCO system.

#### 4.5 Experimental Recognition Results for Various Facial Images

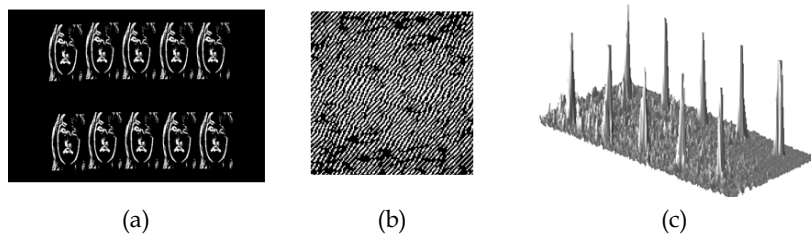


Figure 12. (a) Input images,(b) correlation filter, (d) correlation intensity 10ch

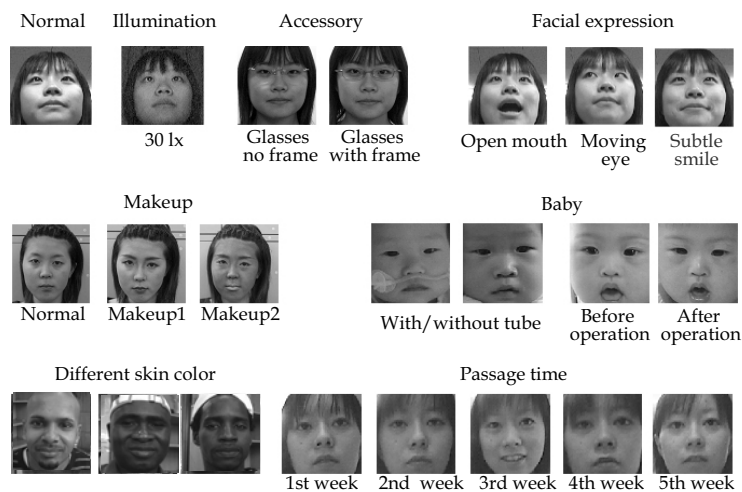


Figure 13. Input images example of successful recognition

Various facial images for input are shown in Fig.13. Facial images taken after a period of time, with glasses and various changed features (e.g. suntan and heavy make-up) can all be identified. Conventionally, the facial images of racially mixed persons, twins and infants have been regarded as the most difficult objects of all. However, the system has enabled these cases to be recognized as separate entities, enhancing its systemic robustness. The following two reasons may be considered responsible. Firstly, an edge-enhancement

binarization as a pre-process breaks down the shape of facial images, regardless of make-up and different skin color, clearly delineating unchanged elements of the images. Secondly, the matched filter lies within the range of the spatial frequency codifying facial features. These results ascertained a considerably high recognition rate of the proposed algorithm, alongside high robustness. Making the most of the high-speed data processing capability of this device, even higher robustness can be achieved for various recognition conditions when registering many category data (e.g. multiplex data extracted from a three-dimensional model) for a single person.

### 5. Highly Precise Algorithm Using Multiple Database Images for FARCO

Section 6 applies this system to a temporal sequence of moving images. The multiplexed database is extracted from video data, and contains various images taken from different angles. Our experiments confirmed that temporal sequential images functioned effectively as part of the system. From the results, we conclude that this is a promising system for a variety of purposes such as security and medical safety, where a large number of images have to be handled at high speed (Watanabe & Kodate et al., 2005).

#### 5.1 Highly Precise Algorithm Using Multiple Database Images

