# Optical character recognition using a memory matrix generated from singular value decomposition

**Osami Sasaki,** MEMBER SPIE
**Kenichi Sakata**
**Akihito Shibahara**
**Takamasa Suzuki,** MEMBER SPIE
Niigata University
Faculty of Engineering
Niigata-shi 950-21
Japan
E-mail: osami@eng.niigata-u.ac.jp

**Abstract.** A memory matrix provides output vectors that are specified by the corresponding input vectors. If the input vectors represent input characters, the memory matrix performs character recognition. When the input characters contain noise, it is difficult to recognize the characters. To overcome this difficulty, we generate a new memory matrix by using singular value decomposition and manipulating the singular values. The effectiveness of the memory matrix is made clear by recognizing 26 alphabets characters containing noise with an optical recognition system. 
[S0091-3286(98)00209-8]

Subject terms: associative memory; pattern recognition; singular value decomposition; optical signal processing.

Paper 01127 received Dec. 1, 1997; revised manuscript received Apr. 14, 1998; accepted for publication Apr. 15, 1998.

## 1 Introduction

A number of methods based on associative memory models have been studied for pattern recognition.[1–4] It is effective in classifying similar looking patterns to project input patterns to orthogonal vectors using the pseudoinverse algorithm[1] or the least-squares linear mapping technique.[2] In addition, in the case where the input patterns are distorted by random noise, some recursive operations are performed to acquire the correct output vector.[3]

In this paper, a pseudoinverse is calculated with singular value decomposition (SVD) and a memory matrix is generated. The effects of the noise contained in the input pattern on the output vector are clearly understood with singular values in the pseudoinverse.[5,6] Since small singular values amplify the noise, we modify the small singular values by replacing them with a specified value. We recognize alphabet characters with a better memory matrix, which is generated by manipulating the singular values. In the experiment, we construct an optical system for character recognition in which two liquid crystal displays are employed. Through the recognition of 26 alphabets characters containing noise, the usefulness of this system is made clear.

## 2 Memory Matrix Generated from SVD

A memory matrix $\mathbf{M}$ provides output vectors $\mathbf{y}_i$ that are specified by the corresponding input vector $\mathbf{x}_i$ as

$$\mathbf{y}_i=\mathbf{M}\mathbf{x}_i \quad i=1,\ldots,K, \tag{1}$$

where $K$ is the number of character patterns to be stored in the memory matrix. We define an input matrix $\mathbf{X}$ and an output matrix $\mathbf{Y}$ whose $i$'th columns are the input vectors $\mathbf{x}_i$ and the output vectors $\mathbf{y}_i$, respectively as

$$\mathbf{X}=[\mathbf{x}_1\cdots\mathbf{x}_K] \quad \mathbf{Y}=[\mathbf{y}_1\cdots\mathbf{y}_K]. \tag{2}$$

Then Eq. (1) becomes

$$\mathbf{Y}=\mathbf{M}\mathbf{X}. \tag{3}$$

If we have a pseudoinverse matrix $\mathbf{X}^+$ of $\mathbf{X}$, we obtain a memory matrix

$$\mathbf{M}=\mathbf{Y}\mathbf{X}^+. \tag{4}$$

The pseudoinverse $\mathbf{X}^+$ is calculated with SVD as

$$\mathbf{X}^+=\sum_{i=1}^{K} s_i^{-1}\mathbf{u}_i\mathbf{v}_i^T, \tag{5}$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are column vectors and orthogonal each other, and $T$ denotes the transpose of the vector. Singular value $s_i$ has the property

$$s_1\geqslant s_2\geqslant\cdots\geqslant s_K. \tag{6}$$

In our study, the input vector $\mathbf{x}_i$ is formed from an alphabet character whose pixel size is $7\times7$ and whose pixels are black or white. The number of the elements of the input vector is 49, their values are 0 or 1 according to the black or white pixels, and $K=26$. The $i$'th element of the output vector $\mathbf{y}_i$ is 1, and the others are 0, which means the corresponding input is the $i$'th one of 26 alphabet characters. Thus the output matrix $\mathbf{Y}$ becomes a $26\times26$ unit matrix, and $\mathbf{M}=\mathbf{X}^+$. By defining a coefficient and a matrix as

$$c_i=s_i^{-1} \quad \mathbf{U}_i=\mathbf{u}_i\mathbf{v}_i^T, \tag{7}$$

from Eq. (5) we rewrite the memory matrix $\mathbf{M}$ as

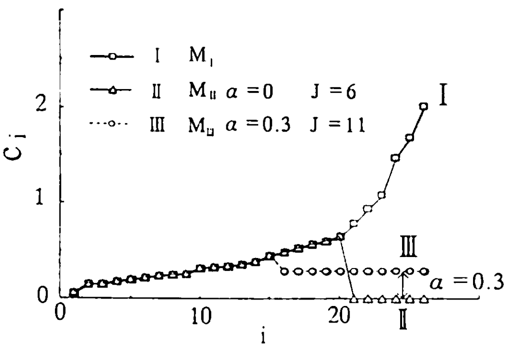


**Fig. 1** Distributions of $c_i$ for $\mathbf{M}_{\mathrm{I}}$, $\mathbf{M}_{\mathrm{II}}$, and $\mathbf{M}_{\mathrm{III}}$.

$$\mathbf{M}_{\mathrm{I}}=\sum_{i=1}^{K} c_i\mathbf{U}_i. \tag{8}$$

Figure 1 shows the distribution of the coefficient values using Roman numeral notation. When input characters contain noise, this noise component is amplified by the coefficients whose values are large. To eliminate the effect of noise, we do not use $J$ coefficients whose values are larger than the others, and we denote a memory matrix $\mathbf{M}_{\mathrm{II}}$ as

$$\mathbf{M}_{\mathrm{II}}=\sum_{i=1}^{K-J} c_i\mathbf{U}_i. \tag{9}$$

The memory matrix $\mathbf{M}_{\mathrm{II}}$ provides the result that an average recognition rate over all characters increases, but the recognition rate for some characters decreases. To avoid this low recognition rate, the $J$ coefficients are replaced with a nonzero value of $\alpha$ to obtain memory matrix $\mathbf{M}_{\mathrm{III}}$ as

$$\mathbf{M}_{\mathrm{III}}=\sum_{i=1}^{K-J} c_i\mathbf{U}_i+\sum_{i=K-J+1}^{K} \alpha\mathbf{U}_i. \tag{10}$$

Memory matrix $\mathbf{M}_{\mathrm{III}}$ provides the result that the low recognition rate for some characters is improved, and the average recognition rate for all characters increases. Figure 1 shows examples of distributions of $c_i$ and $\alpha$ for $\mathbf{M}_{\mathrm{II}}$ and $\mathbf{M}_{\mathrm{III}}$ by the notation of II and III, respectively.

We performed computer simulations to determine the values of $J$ and $\alpha$. The elements of alphabet characters are 0 or 1. We added Gaussian noise with zero mean and variance $\sigma^2$ to the alphabet characters. Then we assigned 0 and 1 to the elements of the alphabet characters if the value of the element was less than or more than 0.5, respectively. The signal-to-noise ratio (SNR) is defined to be $1/\sigma$. We tried to recognize 50 characters containing noise for one alphabet character. We assumed that recognition could be accomplished if the output $y_i$ corresponding to the input character had a maximum value and the other outputs were less than 90% of the value of the output $y_i$. A ratio of the number of correctly recognized characters to the total number of characters was defined as the recognition rate. An average recognition rate over 26 alphabet characters was obtained to estimate the ability of the memory matrix with different values of $\alpha$ and $J$. Figure 2 shows the average recognition rate versus the value of $\alpha$ for different values of $J$ at SNR = 1.5. When $\alpha=0$, values of $J=6$ or 4 gave us the best average recognition rate at different values of SNR. The value of $\alpha=0.3$ gave us the best average recognition rate at different values of SNR for $J=8$ to 12. These results indicates that an appropriate value of $\alpha$ is 0.3. Figure 3 shows the average recognition rate versus the value of $J$ for different values of SNR at $\alpha=0.3$. These results indicates that an appropriate value of $J$ is 11, since the average recognition rate becomes almost constant from $J=11$. Considering these results, we constructed the memory matrix $\mathbf{M}_{\mathrm{II}}$ with $\alpha=0$ and $J=6$ and the memory matrix $\mathbf{M}_{\mathrm{III}}$ with $\alpha=0.3$ and $J=11$, as shown in Fig. 1.

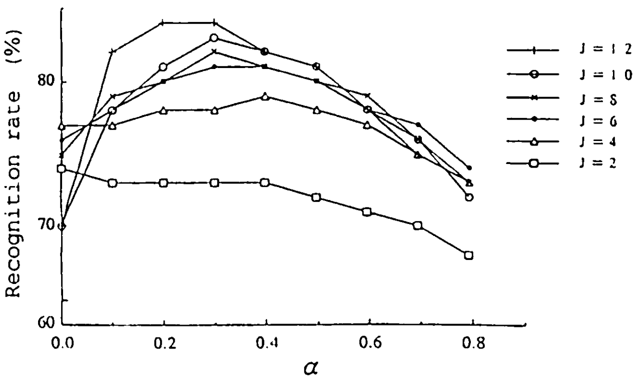


**Fig. 2** Average recognition rate versus the value of $\alpha$ for different values of $J$ at SNR = 1.5.

## 3 Optical Pattern Recognition System

Figure 4 shows an optical character recognition system in which the computation of Eq. (1) is executed. An input character $\mathbf{x}_i$ is displayed on liquid crystal display 1 (LCD1) with binary levels in the region denoted by $x_1$ to $x_n$ ($n=49$) in the vertical direction, and the same input character $\mathbf{x}_1$ is repeated $m$ times ($m=K=26$) in the horizontal direction, as shown in Fig. 4. The elements of matrix $\mathbf{M}$ are divided into nonnegative and negative elements. The matrix

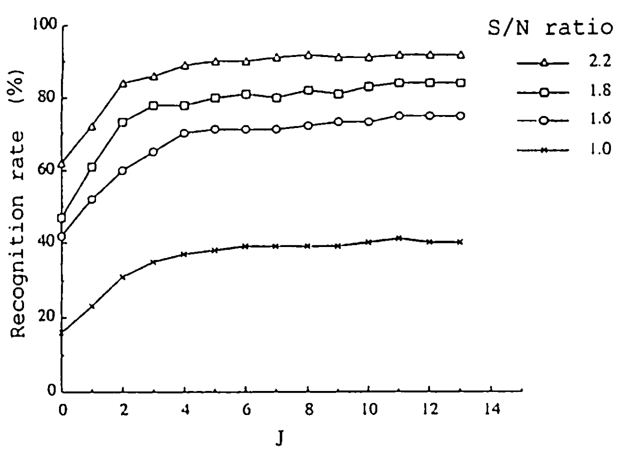


**Fig. 3** Average recognition rate versus the value of $J$ for different values of SNR at $\alpha=0.3$.

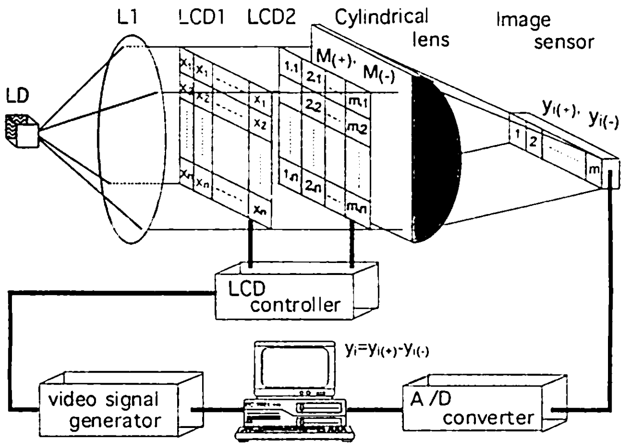


**Fig. 4** Optical character recognition system.

is decomposed as $\mathbf{M}=\mathbf{M}_{(+)}-\mathbf{M}_{(-)}$. In $\mathbf{M}_{(+)}$ negative elements of $\mathbf{M}$ are replaced by zero values and nonnegative elements of $\mathbf{M}$ remain. In $\mathbf{M}_{(-)}$ nonnegative elements of $\mathbf{M}$ are replaced by zero values and negative elements of $\mathbf{M}$ remain, changing the minus signs to plus signs. The memory matrices $\mathbf{M}_{(+)}$ and $\mathbf{M}_{(-)}$ of $m\times n$ matrix size are displayed on LCD2 with 256 gray levels. An element of the memory matrix at the $i$'th row and the $j$'th column is on the position denoted by $(i,j)$, as shown in Fig. 4. Light from the laser diode (LD) are collimated with lens 1 (L1). The collimated light passes through the two LCDs and the transmitted light is collected with a cylindrical lens onto the image sensor. Outputs of the image sensor are represented by $y_{i(+)}$ and $y_{i(-)}$ $(i=1,...,m)$ when $\mathbf{M}_{(+)}$ and $\mathbf{M}_{(-)}$ are displayed on LCD2, respectively. The $i$'th element of the outputs is obtained from the light intensity on the region of the image sensor denoted by a number $i$. The outputs of the image sensor are fed to a personal computer through an 8-bit analog-to-digital (A/D) converter. After the optical matrix-vector multiplication is performed twice, a final output $y_i=y_{i(+)}-y_{i(-)}$ is obtained in the personal computer. If the $i$'th element $y_i$ of the final output is the largest among the other elements, the input character is recognized as $\mathbf{x}_i$.

The input character and the memory matrix are displayed on the LCDs with the personal computer, a video signal generator, and an LCD controller.

## 4 Experimental Results

The optical recognition system shown in Fig. 4 was constructed to perform recognition of 26 alphabets characters. First we used $\mathbf{M}_{\mathrm{I}}$, given by Eq. (8), as the memory matrix $\mathbf{M}$. When the input character did not contain any noise, we obtained the results for four input characters of A, M, N, and Z, as shown in Fig. 5. The input characters are shown on the left side. Next to the input character, the corresponding output is shown. The vertical axis is the value of the final output $y_i (i=1,...,26)$, and the horizontal axis is the number of $i$. Along the horizontal axis, instead of the number of $i$, the characters A to Z are indicated. For the input of the character A, the value of $y_i$ is the largest at A. For the other input characters, the situations are the same. When the input characters contained noise whose SNR was 1.5,

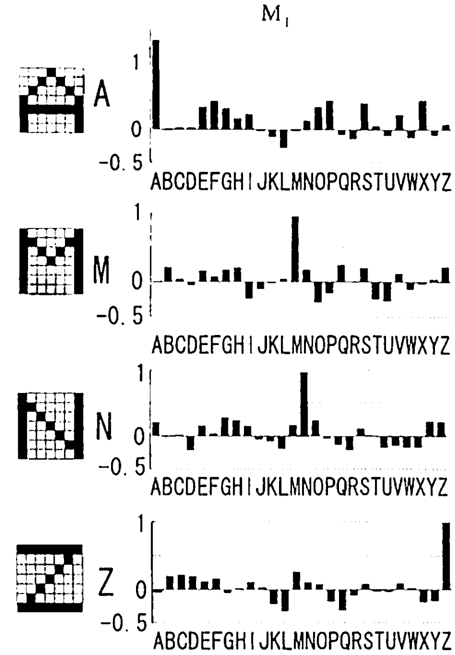


**Fig. 5** Output vectors obtained with memory matrix $\mathbf{M}_{\mathrm{I}}$ for input characters containing no noise.

we obtained the results shown in Fig. 6. Input characters corrupted with noise are shown on the left side. Because of the noise we cannot recognize the characters with the memory matrix $\mathbf{M}_{\mathrm{I}}$. We made 50 different patterns corrupted with noise for one alphabet character by changing the noise, and tried to recognize these patterns. The recognition rate was calculated using the definition described in Sec. 2. The recognition rates of each of 26 alphabet char-

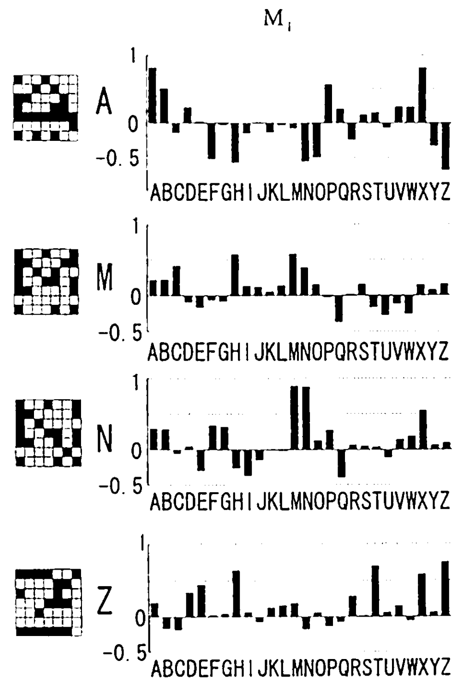


**Fig. 6** Output vectors obtained with memory matrix $\mathbf{M}_{\mathrm{I}}$ for input characters corrupted with noise.

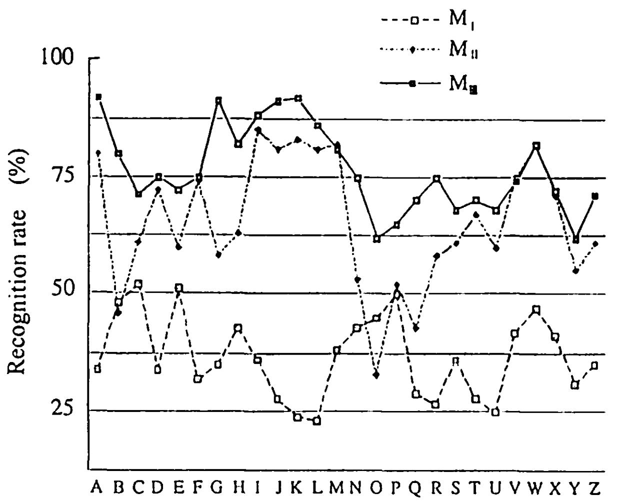


**Fig. 7** Recognition rate of 26 alphabet characters obtained with $\mathbf{M}_{\mathrm{I}}$, $\mathbf{M}_{\mathrm{II}}$, and $\mathbf{M}_{\mathrm{III}}$.

acters obtained with the memory matrix $\mathbf{M}_{\mathrm{I}}$ are shown in Fig. 7. The horizontal axis is the input character. The recognition rate is very low when we use the memory matrix $\mathbf{M}_{\mathrm{I}}$. In the following experiments we used input characters corrupted with noise at SNR$=1.5$.

Next we used $\mathbf{M}_{\mathrm{II}}$, given by Eq. (9), as the memory matrix $\mathbf{M}$ and obtained the results shown in Fig. 8. The input characters A, M, and Z are recognized. The recognition rate obtained with the memory matrix $\mathbf{M}_{\mathrm{II}}$ is shown in Fig. 7. The recognition rate is improved for most of the

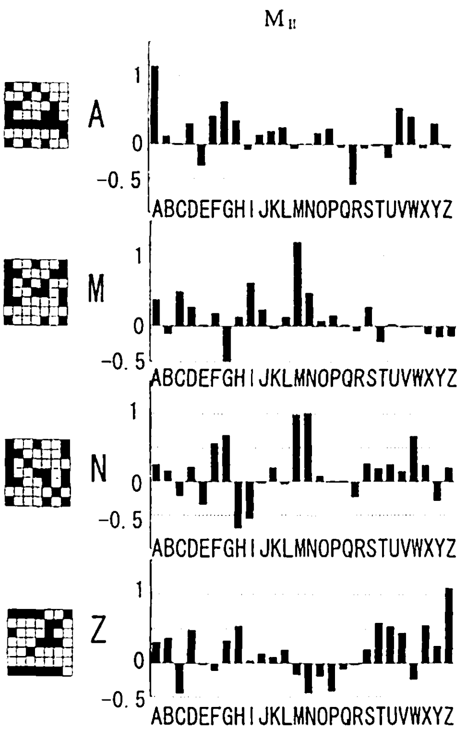


**Fig. 8** Output vectors obtained with memory matrix $\mathbf{M}_{\mathrm{II}}$ for input characters corrupted with noise.

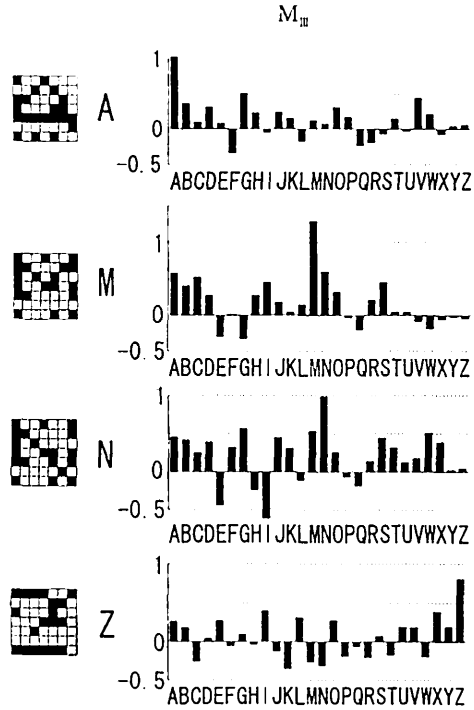


**Fig. 9** Output vectors obtained with memory matrix $\mathbf{M}_{\mathrm{III}}$ for input characters corrupted with noise.

characters, but the recognition rate of a few characters is still around 50% or less.

Finally we used $\mathbf{M}_{\mathrm{III}}$, given by Eq. (10), as the memory matrix $\mathbf{M}$ and obtained the results shown in Fig. 9. The input characters A, M, N, and Z are recognized. The recognition rate obtained with memory matrix $\mathbf{M}_{\mathrm{III}}$ is shown in Fig. 7. The recognition rate is greatly improved for all of the 26 characters.

## 5 Conclusions

We calculated a pseudoinverse with SVD to obtain the memory matrix that performed character recognition against the noise contained in the input characters. A memory matrix insensitive to noise was generated by manipulating the small singular values. We constructed an optical system to perform character recognition with the memory matrix by using two LCDs. Experimental results of the recognition of 26 alphabets characters containing the noise showed clearly that this optical system eliminates the effects of the noise. However, the optical system is sensitive to shift, rotation, and deformation of the characters. Some preprocessing operations such as Fourier transformation are required in the case where the optical system is applied to recognition of handwritten characters.

## *References*


1. T. Kohonen and M. Ruohonen, "Representation of associated pairs by matrix operators," *IEEE Trans. Comput.* C-22, 701–702 (1973).
2. Z. H. Gu, J. R. Leger, and S. H. Lee, "Optical implementation of the least-squares linear mapping technique for image classification," *J. Opt. Soc. Am.* 72, 787–793 (1982).
3. J. Honnma, Y. Kousugi, and T. Sato, "Evaluation and improvement of associative memory for pattern recognition," *Appl. Opt.* 29, 1675–1681 (1990).
4. B. S. Kiselyov, N. Y. Kulakov, A. L. Mikaelian, and V. A. Shkitin,

"Optical associative memory for high-order correlation patterns," *Opt. Eng.* **31**, 764–767 (1992).
5. C. K. Rushforth, A. E. Crawford, and Y. Zhou, "Least-squares reconstruction of objects with missing high-frequency components," *J. Opt. Soc. Am.* **72**, 204–211 (1982).
6. O. Sasaki and T. Yamagami, "Image restoration in singular vector space by the method of convex projections," *Appl. Opt.* **26**, 1216–1221 (1987).

**Osami Sasaki** received his BE and ME degrees in electric engineering from Niigata University in 1972 and 1974, respectively, and his DrEng degree in electric engineering from Tokyo Institute of Technology in 1981. He is a professor at Niigata University. His research interests include optical metrology and optical information processing.

**Kenichi Sakata** received his BE and ME degrees in electric engineering from Niigata University in Japan in 1994 and 1996, respectively. His research involved optical pattern recognition. Since 1996 he has worked on the design and quality control of LSI, Niigata Sanyo Corporation.

**Akihito Shibahara** received his BE and ME degrees in electric engineering from Niigata University in Japan in 1993 and 1995, respectively. His research involved optical pattern recognition. Since 1995 he has worked on the development of RISC system software and tools at Semiconductor System Engineering Center, Toshiba Corporation.

**Takamasa Suzuki** received his BE degree in electric engineering from Niigata University in 1982, his ME degree in electric engineering from Tohoku University in 1984, and his DrEng degree in electric engineering from Tokyo Institute of Technology in 1994. He is an associate professor at Niigata University. His research interests include optical metrology, optical information processing, and phase conjugate optics.