J. Res. Inst. Indus. Tech.
Cheju Nat. Univ., 5 : 59-66, 1994.

濟州大 産技硏報, 5 : 59-66, 1994.

Shape Invariant Korean Character Recognition Using Optical Associative Memory

Yang Hoi Doh*, Jeong Woo Kim**, and Soo Joong Kim**

光 聯想 記憶裝置를 利用한 模樣不變 한글 文字認識

都良會*, 金正雨**, 金秀重**

ABSTRAET

A holographic implementation of an optical associative memory system for shape invariant recognition of the printed Korean character is proposed. The structure of the recognition system is a single-layer neural network employing feedback. To form an optical memory loop, we use cascaded Vander Lugt correlators. The first holographic correlator is made of MACE filter having a high discrimination capability. And to reduce the size of the recognition system, we multiplex four MACE filters into one. Thus, 14 consonant MACE filters and 10 vowel MACE filters can be synthesized into four consonant MMACE filters and three vowel MMACE filters, respectively. The second holographic correlator, on the other hand, is made of a CMF having broadband characteristics so that the reconstructed image has high fidelity. Computer simulation results show that the proposed system has shape invariance and high discrimition capability.

Introduction

Korean characters are composed of simple 10 vowels and 14 consonants which can be combined into commonly used 1,500 characters. Thus it is preferred to recognize each consonant and each vowel (from now

* 공과대학 전자공학과
** 경북대학교 공과대학 전자공학과

on simply referred to as element of the character; 〈example〉 강 → ㄱ, ㅏ, ㅇ) rather than to recognize the whole character (Kim et. al, 1994; Kim et. al, 1993). The Korean consonants and vowels are similar and/or symmetric in their shapes, and diverse variations exist in their combinatorial form. Thus the recognition system is required not only to discriminate the similar and/or symmetric patterns but also to adapt to the diverse variations.

Here, a holographic implementation of an optical associative memory for shape invariant recognition of the printed Korean characters is proposed. The structure of the proposed system is a single-layer neural network employing feedback (Hsu et. al, 1990). Its three main operations are vector-matrix multiplication, thresholding and feedback. Two computer generated holograms (CGHs), to provide an interconnection matrix, are placed on intermediated plane of cascaded Vander Lugt correlators to form an optical memory loop. The holographic correlators store the reference images in a hologram and retrieve them in a coherently illuminated feedback loop. The first correlator has not only to discriminate the similar and/or symmetric patterns but also to adapt to the diverse variations in the combinatorial form. To satisfy these requirements, we synthesized a minimum average correlation energy (MACE) filter which produces sharp correlation peak and minimum sidelobe (Mahalanobis et. al, 1987). And to reduce the size of system, we synthesized multiplexed MACE (MMACE) filter using carrier frequency (Kim et. al, 1993). This MMACE filter contains four different subfilters (MACE filters), each of a different reference elements. All the subfilters are spatially compounded, but the output correlation between the input character and the MMACE filter is divided into four subplane corresponding to the location of the each subfilter. The second correlator is made of a conventional matched filter (CMF) so that the reconstructed image has high fidelity with respect to the input image.

Thus, an input image which has diverse variations of the element is applied to the system and be optically correlated with all of the stored reference images, simultaneously. The results of the second correlator is thresholded and fed back to the input, where the strongest correlation reinforces the input image. This reconstructed image is thresholded by an optimal value, and the thresholded value is feedback to the input until the system stabilized on the desired element. The computer simulation results show that the proposed printed Korean character recognition algorithm has high discrimination, and shape-invariance capability.

Implementation of holographic filters

In this study, target images are printed Korean characters. The holographic filters in an optical correlator system have the capability of an image discrimination and reconstruction. There are two CGHs in the proposed system, the first is used to generate correlation distribution and the second is used as an associative momory.

The first CGH is made of MACE filter having a high discrimination capability for recognition of the similar and/or symmetric images, and diverse variations. And to reduce the size of the recognition system, we multiplex MACE filters. The MACE filter is synthesized in the frequency domain, subject to the constraint that the average correlation plane energy be minimized while producing a correlation peak of a prespecified value corresponding to the location of the training image (Mahalanobis et. al, 1987). Thus, we obtain the Fourier transform (FT) of the consonants or vowels to be recognized. And we generate an each consonant (or vowel) MACE filter for the every consonants (or vowels) having different shapes. $H_{C\text{-MACE-}k}(\xi, \eta)$, the FT of the consonant MACE filter function, is described by the linear combination of the FT of the diverse variations of the consonant reference image, $c_{k\text{-}i}(x, y)$

$$H_{C\text{-MACE-}k}(\xi, \eta) = D_k^{-1} C_k [C_k^+ D_k^{-1} C_k]^{-1} u_{c\text{-}k} \tag{1}$$

where $k=1, 2, \cdots, 14$ and $i=1, 2, \cdots, 6$, (or 7 or 8). And the vector C_k is the FT of $c_{k\text{-}i}(x, y)$, the diagonal elements of D_k describe the power spectrum of $c_{k\text{-}i}(x, y)$, '+' denotes the conjugate transpose of a complex vector, and the element of the constraint vector, $u_{c\text{-}k}$, is the user specified value of the correlation function at the origin. In similar method, $H_{V\text{-MACE-}l}(\xi, \eta)$, the FT of the vowel MACE filter function, is described by the linear combination of the FT of the diverse variations of the vowel reference image, $v_{l\text{-}j}(x, y)$

$$H_{V\text{-MACE-}l}(\xi, \eta) = G_l^{-1} V_l [V_l^+ G_l^{-1} V_l]^{-1} u_{v\text{-}l} \tag{2}$$

where $l=1, 2, \cdots, 10$ and $j=1$ or $1, 2$. And G_l, V_l, '+' and $u_{v\text{-}l}$ have the same meaning with D_k, C_k, +, and $u_{c\text{-}k}$ in the consonant MACE filter respectively. In this study, the size of the input character is 64×64, the size of the MACE filter and MMACE filter are 128×128 (the size of the correlation plane) respectively. And four 128×128 MACE filters can be synthesized into the only one MMACE fulter plane because the correlation distribution between MACE filter and input character is restricted to 64×64. Therefore, the 14 consonant MACE filters and 10 vowel MACE filters are synthesized into four consonant MMACE filters and three vowel MMACE filters using the multiplexing technique. The correlation distribution between each MACE filter in the MMACE filter plane and the input character must be sufficient distance for no aliasing. Therefore, we move the center of the 64

$\times$64 correlation plane to the center of the 64$\times$64 subplane of the 128$\times$128 correlation plane. Then the consonant MMACE filters are given by

$$H_{\text{C-MMACE-}k}(\xi,\eta)=\sum_{i=1}^{N}K_iH_{\text{C-MACE-}i}(\xi,\eta)\times\exp[j2\pi(a_i\xi+b_i\eta)] \quad (3)$$

where N=3 or 4, k=1, 2, 3, 4. And K_i denotes the amplotude of the ith reference wave, a_i is a horizontal distance from the reference point source to the ith cloloumn and b_i is a vertical distance from the reference point to the ith row. We choose the a_i, and b_i as $a_1=b_1=b_2=a_3=32$ and $a_2=b_3=a_4=b_4=-32$, respectively. Similarly, vowel MMACE filters are described by

$$H_{\text{V-MMACE-}l}(\xi,\eta)=\sum_{i=1}^{M}K_iH_{\text{V-MACE-}i}(\xi,\eta)\times\exp[j2\pi(a_i\xi+b_i\eta)] \quad (4)$$

where M=3 or 4, l=1, 2, 3. And K_i, a_i, and b_i are the same meaning as each paramenter in the consonant MMACE filters.

The second CGH is made of a CMF having broadband characteristics so that the reconstructed image has high fidelity. $H_{CMF}(\xi,\eta)$, the CMF function, is described by

$$H_{CMF}(\xi,\eta)=[F(\xi,\eta)]^* \quad (5)$$

where $F(\xi,\eta)$ is the FT of the refernce image function, $f(x,y)$, and '*' denotes the complex conjugate.

System architecture for experiments

Recognition algorithm

A flowchart of the printed Korean character recognition algorithm using the holographic associative memory is shown in Fig 1. It consists of a single feedback loop for a consonant or a vowel filter. In order to recognize the full printed Korean character, the system must have four consonant recognition loops and three vowel ones, and it is carried out in parallel.

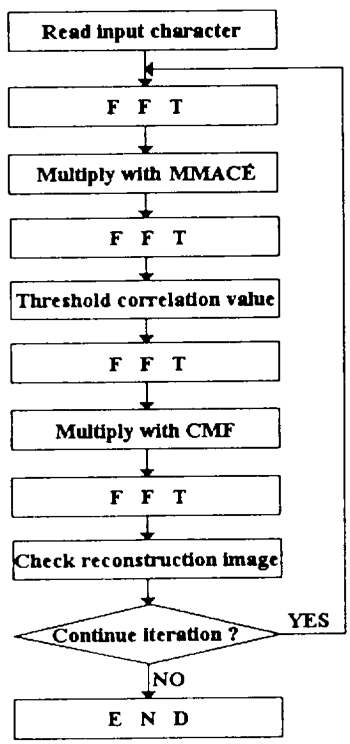


Fig. 1. Flowchart of Korean character recognition.

This algorithm is performed by four step. First, the 2-D correlation between the input character $s(x, y)$ and each of the reference elements stored in the first correlation filter, $H_{MMACE}(\xi, \eta)$, is calculated. Second, the correlation output is thresholded by an optimal value. If the strongest correlation value is larger than the thresholding level, the stored reference element exists at the level, the stored reference element exists at the position of the correlation value. And this correlation output preserves location information of the input characters. Third, 2-D correlation between the thresholded correlation output and the second correlation filter, $H_{CMF}(\xi, \eta)$, is calculated. At the center of the output plane of the second correlator we obtain the superposition of the four stored images. The stored image that is most similar to the input element gives the strongest correlation output, hence the brightest reconstructed image is appeared. Fourth, this reconstructed image is thresholded by an optimal value, and the thresholded value is fed back to the input until the system stabilizes on the desired element. In this way the stable patten established in the loop is typically the stored image that is most similar to the original input.

Optical setup

The schematic diagram of the optical architecture is shown in Fig 2. It is constructed by two 4f correlator (Vander Lugt corelator) systems. The first correlator consists of the liquid crystal television (LCTV) controlled by a computer, the first hologram H1, the FT lenses L1 and L2, the mirror M1, and thresholding device at plane P1. The LCTV is used as an input device. The second correlator consists of P1, M2, the second hologram H2, L3, L4, and CCD camera.

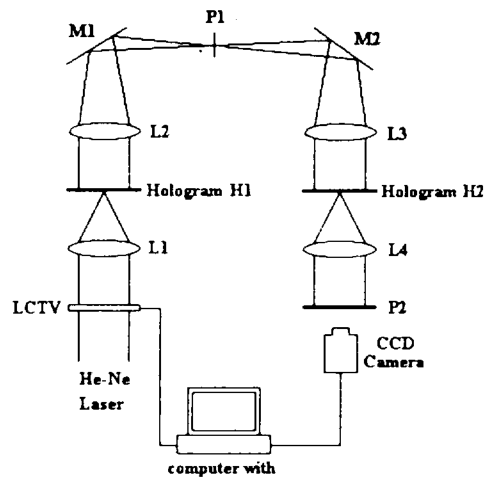


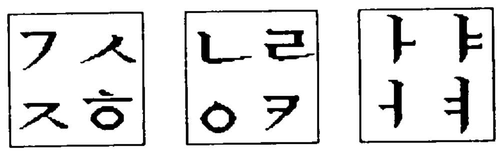
Fig. 2. Schematic diagram of optical experime...

A collimated He-Ne laser beam illuminates the input image on the LCTV. The laser beam passed through LCTV is Fourier transformed by lens L1 and illuminates hologram H1. The corelation between the input image and each of the stored images is produced at plane P1. The correlation output at the P1 is detected by a CCD camera and thresholded by a computer. The remainder, the system from P1 back to the CCS camera, must become a replica of the first half, with the hologram H2 storing the same set of images as H1. The thresholding step in the P1 and P2 plane is performed by the CCD camera and a frame grabber mounted on the computer. In the first processing process, the thresholding value is determined by 70% of minimum auto-correlation peak value among every reference consonants or vowels to include the large variations of the same element to be recognized.

Computer simulation

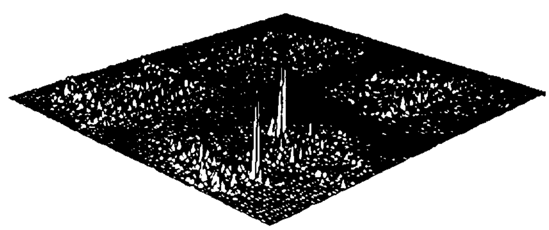
Input image and reference images of the filter for computer simulation are shown in Fig. 3 and Fig. 4 respectively. Fig. 5 shows the correlation outputs between the input image of Fig. 3 and the MMACE filters of the Fig. 4. Fig. 6(a), 6(b) and 6(c) show the intensity distribution which are thresholded by 70% of minmum auto-correlation peak value of every reference image. Fig. 6(d), 6(e) and 6(f) show the intensity distribution which are correlated between the former thresholded distribution and CMF of the Fig. 4, and thresholded by an optimal intensity level. These intensity distribution of Fig. 6(d), 6(e) and 6(f) represent the original image. These results are achieved after one iteration. The reconstructed image of the system is shown in Fig. 7.

Fig. 3. Input image '갉' for computer simulation.

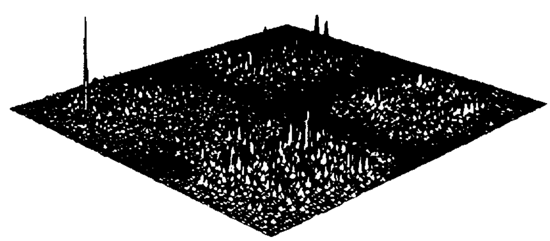


(a) (b) (c)

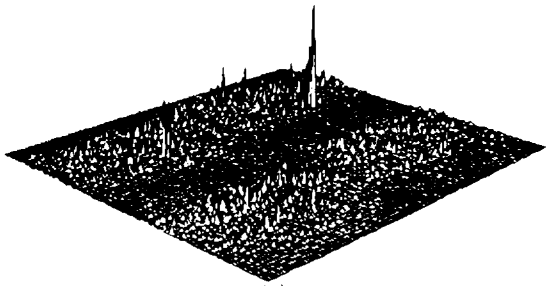
Fig. 4. Filter images of consonant and vowel for computer simulation.
(a) filter image of 'ㄱ,ㅅ,ㅈ,ㅎ',
(b) filter image of 'ㄴ,ㄹ,ㅇ,ㅋ', and
(c) filter image of 'ㅏ,ㅑ,ㅓ,ㅕ'.



(a)



(b)



(c)

Fig. 5. Correlation distributions between input image of Fig. 3 and the first CGHs of Fig. 4.

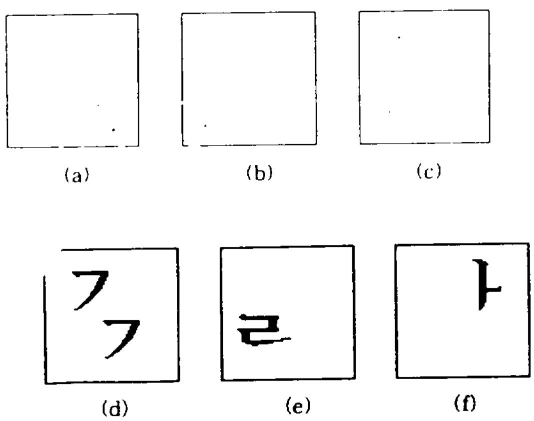


Fig. 6. Simulation results of input image of Fig. 3 and filter images of Fig. 4.
(a) thresholded correlation output at Pl for filter image of Fig. 4(a),
(b) thresholded correlation output at Pl for filter image of Fig. 4(b),
(c) thresholded correlation output at Pl for filter image of Fig. 4(c),
(d) sampled and thresholded result for filter image of Fig. 4(a),
(e) sampled and thresholded result for filter image of Fig. 4(b), and
(f) sampled and thresholded result for filter image of Fig. 4(c).

Fig. 7. Final result of the recognized Korean character.

In this simulation, we use '갉' as an input image. And the results show that we need only one iteration to recover a perfect original image. If consonant or vowel of input image is more similar to the reference image, several iterations are needed for perfect recovery.

Conclusion

Holographic implementation of a fully connected neural network is proposed for shape invariant recognition of the printed Korean characters. The proposed printed Korean character recognition system employs cascaded two holographic Vander Lugt correlators and thresholding operations to form the optical memory loop. The first CGH is made of a MMACE filter not only to improve the discrimination capability and recognize the shape variations of the same consonant or vowel but also to reduce the size of the recognition system. And the second CGH is made of CMF so that the reconstructed images have reference form with respect to the original input images.

Thresholding operations are performed by a CCD camera and a frame grabber mounted on a digital computer with an optimal thresholding value. The thresholding value is determined by 70% of minimum auto-correlation peak value among every reference consonants or vowels to adapt to the large variation of the same element to be recognized.

The performance of the proposed printed Korean character recognition system is evaluated by a computer simulation. The simulation results show that the proposed system has shape invariance and high discrimination capability.

References

Hsu, Ken-Yuh, Shin-Yu Li, and Demetri Psaltis, 1990. Holographic implementation of a fully connected neural network, *Proc. of the IEEE*, 78 : 1637-1645.

Kim, J. W., J. K. Bae and Y. H. Doh, 1994. Korean character recognitiion using optical associative memory, *Journal of KITE*, 31-A : 61-69.

Kim, J. W., C. S. Kim, W. H. Cho, Y. H. Doh, and S. J. Kim, 1993. Printed Korean character recognition using multiplexed MACE filter, *Proc. of the 1993 Korean Signal Processing Conference*, 6 : 337-341.

Mahalanobis, A., B. V. K. Vijaya Kumar, and D. Casasent, 1987. Minimum average correlation energy filters, *Appl. Opt.*, 26 : 3633-3640.

〈국문초록〉

글자모양의 변형에 무관하게 인쇄체 한글문자를 인식하기 위하여 홀로그램을 이용한 광 연상 기억 시스템을 제안하였다. 제안된 인식시스템의 구조는 궤환을 가지는 단층 신경회로망이며, 광기억 루프를 구성하기 위하여 종속접속된 Vander Lugt 상관기를 사용하였다. 첫번째 홀로그램 상관기는 변별력을 높이기 위해 MACE 필터를 이용하여 만들었으며, 인식시스템의 규모를 줄이기 위하여 4개의 MACE 필터를 하나로 다중화 하였다. 이리하여 4개의 MACE 필터와 10개의 모음 MACE 필터를 각각 4개 및 3개의 다중 MACE 필터로 합성할 수 있었다. 한편 두번째 홀로그램 상관기는 원래 영상을 그대로 복원할 수 있도록 광대역 특성을 가지는 CMF를 이용하여 만들었다. 컴퓨터 시뮬레이션을 통하여 제안된 시스템이 글자모양의 변형에 무관하고, 변별력이 우수한 특성을 가짐을 확인하였다.