HUMAN-PERCEPTION-LIKE IMAGE RECOGNITION SYSTEM BASED ON THE ASSOCIATIVE PROCESSOR ARCHITECTURE

Masakazu YAGI, Tadashi SHIBATA*, and Kenji TAKADA**

Department of Electronics Engineering, School of Engineering, The University of Tokyo

*Department of Frontier Informatics, School of Frontier Science, The University of Tokyo

7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan,

PHONE & FAX +81-3-5841-8567, goat@if.t.u-tokyo.ac.jp, shibata@ee.t.u-tokyo.ac.jp

**Department of Orthodontics, Faculty of Dentistry, Osaka University

1-8 Yamadaoka, Suita, Osaka, 565-0871 Japan ktakada@dent.osaka-u.ac.jp

ABSTRACT

A human-perception-like image recognition system has been developed aiming at direct hardware implementation based on the Associative Processor architecture. The Principal Axis Projection (PAP) method [1] has been employed for representing images for processing using Associative Processor chips [2-4]. The PAP vectors very well preserve the human perception of similarity among images in the vector space and ideal for use in our system. In this study, the system has been applied to medical radiograph analysis as well as binary image recognition including corrupted handwritten patterns. By introducing two new techniques, namely, "winner score/pattern mapping" and "macro-scale matching", the recognition performance has been drastically improved as compared to our previous work [1]. By utilizing the macro-scale matching technique, the percent correct for three important Cephalometric landmarks (Nasion, Orbitale and Sella) has been improved to be 97.5%, 97.5%, and 87.5 % from 87.5%, 85%, and 62.5%[1], respectively. In order to expedite the PAP vector generation processing, dedicated VLSI chips have been developed and their proper operation has been experimentally demonstrated.

1. INTRODUCTION

In spite of the remarkable progress in the VLSI technology and intensive efforts in software sophistication, human-like image recognition systems having a real-time response capability have not yet been realized. Most of the systems have been designed for very specific problems and often operate too slow to achieve the real-time responses due to the very expensive computation. One of the reasons for this is that the hardware organization and the software algorithms are being developed and optimized independently. Such scheme is quite different from the processing in our brains.

Our approach is based on the Associative Processor architecture, a psychologically-inspired VLSI brain model [5]. In the architecture, image recognition is carried out by recalling the most similar image in a huge database based on the maximum likelihood search. The Associative Processor is the maximum likelihood search VLSI hardware with a huge on-chip database. Due to its parallelism, the template matching, a very time-consuming operation can be conducted almost cost free. The Associative Processor chips in analog [2,3] as well as in digital [4] technologies have been already developed aiming at real time processing. Here the essential issue is the image representation algorithm suitable for use in the Associative Processor chips. For this purpose, Principal Axis Projection (PAP)-method [1] has been developed. With the PAP-method, the edge feature of an image is extracted and compressed into a vector representation that very well preserves the human perception of similarity among similar images while achieving a substantial dimensionality reduction.

In our previous work, PAP vector representation has been successfully applied to both binary image recognition [2,6] and Cephalometric landmark identification [1] and its robust nature has been demonstrated. However, the issues of transformation, multi scale, and multi resolution have not yet been properly treated.

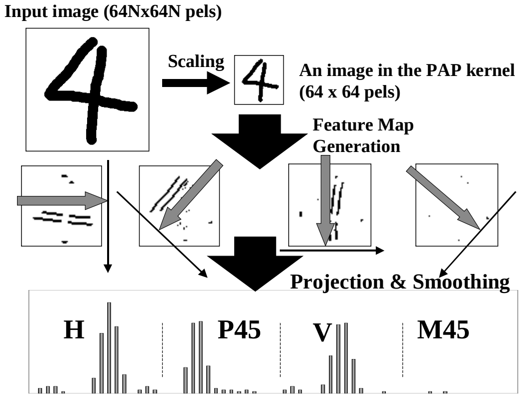

Fig. 1 Characteristic vector formation by Multi Resolution Principal Axis Projection (MRPAP).

The purpose of this paper is to introduce two new techniques to the PAP-based image representation algorithm [1], i. e., “winner score/pattern mapping” and “macro-scale matching”. The winner score/pattern mapping technique has allowed us to perform more robust binary image recognition, and the macro-scale matching technique has enhanced the recognition performance in Cephalometric landmark identification, one of the most important clinical practices in Orthodontics in dentistry.

2. IMAGE RECOGNITION ALGORITHMS

2.1. Multi-Resolution-Principal-Axis-Projection (MRPAP) Technique

Fig.1 illustrates the technique for generating vector representation from an image in several resolutions, named Multi-Resolution-Principal-Axis-Projection (MRPAP). It has been developed by introducing multi resolution concept to Principal Axis Projection (PAP)-method [1].

In the original PAP, a 64x64-pels image is converted to a 64-dimension feature vector. Therefore, if we handle an image having 64Nx64N (N: positive integer), the number of pixel data must be reduced to fit the PAP kernel size of 64x64 pels. This is done by averaging and normalization. Then the 64x64-pels image is subjected to the PAP kernel processing as described below. When a 64-dimension PAP vector is formed from a 64x64N-pels image, it is called a 1/N-resolution processing.

The PAP kernel processing is as in the following. An image (64x64 pels) is subjected to pixel-by-pixel 5x5 spatial filtering to extract four-direction edge information, i.e., horizontal, vertical, and ±45-degree directions at each location. After the spatial filtering, the edge threshold operation determines whether an edge really exists or not. After the edge threshold operation, the detected edge is represented by a digital flag and four feature maps are generated. The edge flags are accumulated in the direction normal to the edge direction gradient and edge histograms are generated. Then, the edge histograms, i. e., the projected data sets of edge flags, are series connected in the order of horizontal, +45,vertical, and –45 to form a one dimensional array of numerals, namely an image representation vector. After the element smoothing operation, a 64-dimensinal vector is finally formed.

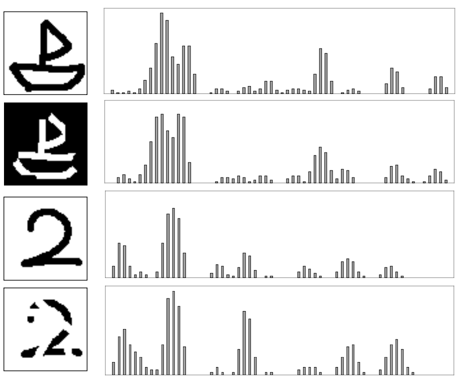

Fig. 2 Vector representations formed by the unity resolution PAP-method (N=1) for several patterns.

The vectors formed by the unity resolution PAP-method (N=1) for binary patterns (a boat and a handwritten digit “2”) are shown in Fig. 2. Their vector representations are very similar to each other while some images suffer from inversion and corruption. As shown by examples, here the PAP-generated vectors very well preserve the human perception of similarity among patterns.

2.2. Winner Score/Pattern Mapping Technique

Winner score/pattern mapping technique is illustrated in Fig. 3. The example shows a target image in which several handwritten digits are randomly drawn. A matching kernel (64x64-pels area, for instance) is taken from the target image, transformed to a PAP vector, and then matched with template vectors stored in the system. By scanning the matching kernel pixel-by-pixel in the target image, the winner template and the dissimilarity score is recorded at each location and the winner score map and the winner pattern map are generated. From these maps, the numerals are identified.

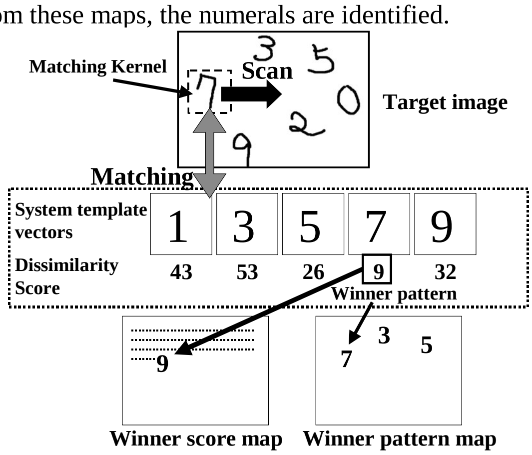

Fig. 3 Winner score/pattern mapping technique

2.3 Macro-Scale Matching Technique

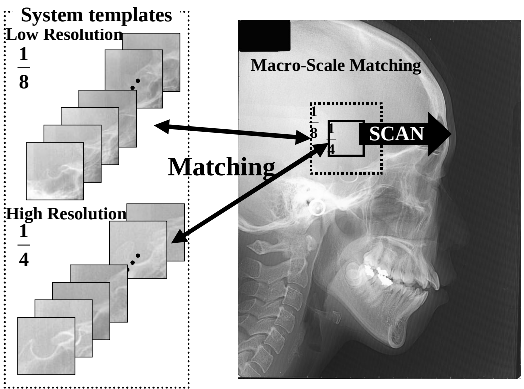

Fig. 4 Macro-scale matching technique for Cephalometric landmark identification

The macro-scale matching is a technique to repeat the winner score mapping using various resolution PAP vectors. The procedure is explained in Fig.4 taking the Cephalometric landmark identification for a Sella pattern search as an example. In our previous work on the Cephalometric landmark identification, 1/4-resolution PAP vector were utilized. In terms of the macro-scale matching technique, the Sella pattern is carried out in a lower scale (1/8-resolution, for instance) in the first stage. In the 1/8-resolution search, the area of matching kernel becomes four times as large as that of the 1/4-resoution, the search is carried out taking more surrounding information into account. Therefore it identifies the area where Sella is locating more robustly. After identifying the area of Sella location by low-resolution search, the correct location in the area is detected by the high-resolution (1/4-resolution) search. Since the search area is restricted to a more probable region, the chance of detecting false location is diminished.

3. EXPERIMENTALS

The image recognition system with the winner score/pattern mapping technique was applied to binary image recognition including corrupted handwritten digit patterns. The unity resolution PAP vectors (N=1) were formed from ten numerical patterns in Times New Romans font and utilized as a group of system templates for matching. In addition, the vacant vector having all zero components was also included as one of the system templates to identify the area of no patterns.

The macro-scale matching technique was applied to Cephalometric landmark identification using 100 head X-ray films for template generation and 40 head X-ray films as target image samples. Three important anatomical points (Sella, Nasion, and Orbitale) in a medical X-ray film of head for the medical X-ray analysis were selected as the target patterns. It should be noted that Orbitale (the lowest point on the average of the right and left borders of the bony orbit) is one of the most difficult points even for expert dentists. 15 template vectors were generated from the 100 samples using Generalized Lloyd learning algorithm [7,8].

Dedicated vector generation chips were designed and fabricated in CMOS digital technology. The Sella pattern search was also carried out using the analog Associative Processor chip developed in a separate project [2].

4. RESULTS AND DISCUSSION

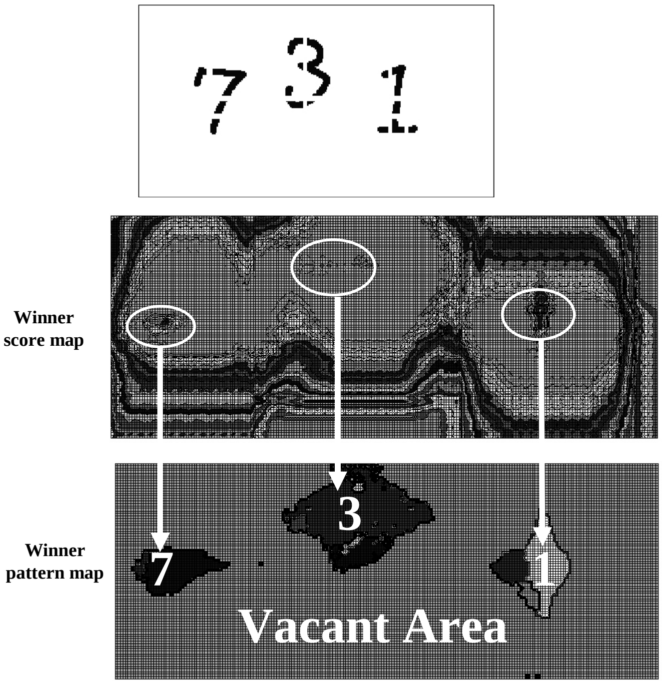

Fig. 5 Target image, the winner score map, and the winner pattern map for the target image

The target image including corrupted handwritten digit patterns, and the winner score map and the winner pattern map are shown in Fig. 5. It is seen that the vacant vector has contributed to enhance the recognition performance by identifying the area of no existing patterns (vacant area) in the winner pattern map. The locations of digit patterns are accurately recognized as the minima in respective areas in the winner score map, and the digit pattern is correctly identified by the winners in the corresponding area in the winner map. These results show the robust nature of winner score/pattern mapping technique.

The results of Sella search in the Cephalometric landmark identification experiment are demonstrated in Fig.6, where the case of a false detection in our previous work [1] is taken as an example (left). By introducing the macro-scale matching technique correct identification result is obtained (middle). Shown on the right is the Sella location identified by an expert dentist having more than 10 years of clinical experience in the university hospital. The search strategy of the macro-scale matching is somehow analogous to the processing by expert dentists in that they firstly restrict the area of Sella location using their expertise knowledge and then proceed to more detailed search.

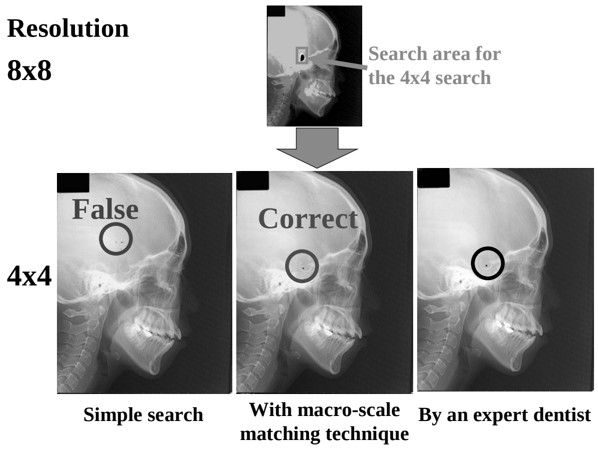

Fig. 6 Sella Recognition results

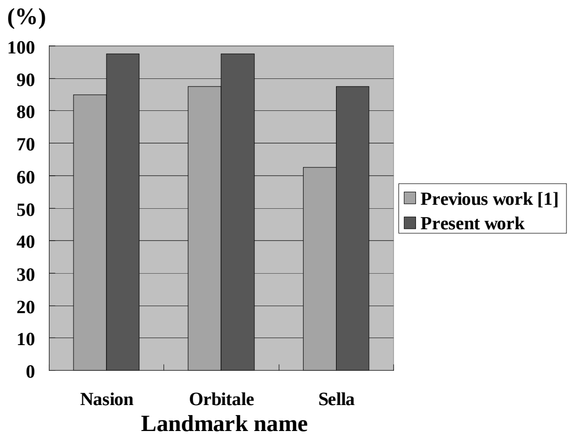

Fig. 7 Percent correct of the recognition results by the simple search system in the previous work [1] and by the system with macro-scale matching technique

The comparison of recognition performances for three

Cephalometric landmarks between those in the previous work [1] and in the present work is shown in Fig. 7. Percent correct for Nasion, Orbitale, and Sella identification as high as 97.5%, 97.5%, and 87.5%, respectively have been achieved by introducing macro-scale matching technique. On the other hand, the percent correct was 85%, 87.5%, and 62.5%, in the previous work for Nasion, Orbitale, and Sella, respectively.

The dedicated digital test chips have been developed to expedite the MRPAP vector generation. The photomicrographs of the chips and the measured data are demonstrated in Fig. 8. Chip #1 was implemented in a 0.6-μm double-poly three-metal-layer CMOS technology and is utilized for the edge threshold determination. Chip #2 was in a 0.35-μm single-poly five-metal-layer CMOS technology, which is for generating feature maps and feature vectors. With these chips, a PAP formed vector is generated in 320 μ seconds while it takes several minutes by the computer simulation.

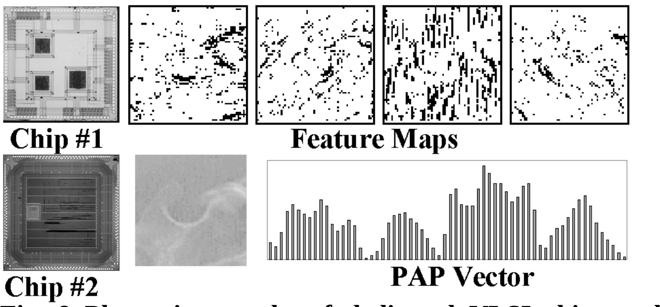

Fig. 8 Photomicrographs of dedicated VLSI chips and feature maps and a vector for a Sella (a pituitary gland) pattern generated by chip #2 at 25 MHz.

The successful Sella pattern search result using analog low power Associative Processor [2] is demonstrated in Fig. 9. It suggests that MRPAP vector representation is very suitable for the Associative Processor architecture. Describing the details of the dedicated VLSI chips and the hardware configuration of the analog Associative Processor [2] is beyond the scope of this paper, and will be presented elsewhere.

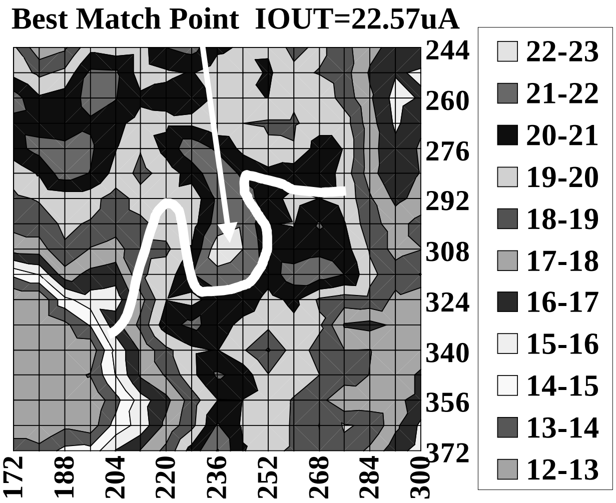

Fig. 9 Winner score map for a Sella pattern recognition. The correct position of a Sella pattern is indicated by a white arrow.

5. CONCLUSIONS

A very robust image recognition system has been developed by introducing two new technologies, the winner score/pattern mapping and macro-scale matching to the Associative-Processor-based system. As a result, Cephalometric landmark identification as well as the binary image recognition containing corrupted digit patterns has been successfully carried out by the present system. In order to accelerate the recognition processing, dedicated VLSI chips have also been developed and the proper operation has been confirmed. The Associative-Processor-based pattern matching system utilizing the multi-resolution principal-axis-projection vector generation algorithm would be a promising candidate to realize a human-perception–like image recognition system.

ACKNOWLEDGEMENT

The VLSI chip for threshold detection and PAP vector generation in this study has been fabricated in the chip fabrication program of VDEC, the University of Tokyo with the collaboration by Rohm Corporation and Toppan Printing Corporation and by Hitachi Ltd. and Dai Nippon Printing Corporation, respectively. The work is partially supported by the Ministry of Education, Science, Sports, and Culture under Grant-in-Aid for Scientific Research (No. 11305024) and by JST in the program of CREST.

REFERENCES

[1] M. Yagi, M. Adachi, and T. Shibata, "A Hardware-Friendly Soft-Computing Algorithm for Image Recognition," Proceedings of 10th European Signal Processing Conference (EUSIPCO 2000), pp. 729-732, Tampere, Finland, Sept. 4-8, 2000.

[2] T. Yamasaki, K. Yamamoto, and T. Shibata, “Analog Pattern Classifier with Flexible Matching Circuitry Based on Principal-Axis-Projection Vector Representation,” *Proceedings of the 27th European Solid-State Circuits Conference (ESSCIRC 2001)*, Ed. by F. Dielacher and H. Grunbacher, pp. 212-215 (Frontier Group), Villach, Austria, September 18-20, 2001.

[3] M. Ogawa and T. Shibata, “NMOS-based Gaussian-Element-Matching Analog Associative Memory,” *Proceedings of the 27th European Solid-State Circuits Conference (ESSCIRC 2001)*, Ed. by F. Dielacher and H. Grunbacher, pp. 272-275 (Frontier Group), Villach, Austria, September 18-20, 2001.

[4] A. Nakada, T. Shibata, M. Konda, T. Morimoto, and T. Ohmi, “A fully-parallel vector quantization processor for real-time motion picture compression,” IEEE Journal of Solid-State Circuits, Vol. 34, No. 6, pp. 822-830, June 1999.

[5] T. Shibata, “Intelligent VLSI Systems Based on a Psychological Brain Model,” Proceedings of 2000 IEEE International Symposium on Intelligent Signal Processing and Systems (ISPACS 2000), pp. 323-332, Honolulu, Hawaii, U.S.A., November 5-8, 2000.

[6] M. Adachi and T. Shibata, "Image Representation Algorithm Featuring Human Perception of Similarity for Hardware Recognition Systems," Proceedings of the International Conference on Artificial Intelligence (IC-AI'2001), Ed. by H. R. Arabnia, Vol. I, 229-234 (CSREA Press, ISDBN: 1-892512-78-5), Las Vegas, Nevada, USA, June 25-28, 2001

[7] S.P. Lloyd, “Least squares quantization in PCM,” IEEE Trans., IT-28, Mar. 1982, pp.127-135.

[8] K. Zeger and A. Gersho, A stochastic relaxation algorithm for improved vector quantiser design, Electronics Letters, No.25, Jul. 1989, pp.896-898.