A HARDWARE-FRIENDLY SOFT-COMPUTING ALGORITHM FOR IMAGE RECOGNITION

Masakazu Yagi
Department of Information and Communication Engineering
The University of Tokyo
7-3-1 Hongo,Bunkyo-ku,Tokyo, 113-8656, Japan
goat@if.t.u-tokyo.ac.jp

Masayoshi Adachi
Department of Electronic Engineering
The University of Tokyo
7-3-1 Hongo,Bunkyo-ku,Tokyo, 113-8656, Japan
adachi@if.t.u-tokyo.ac.jp

Tadashi Shibata
Department of Frontier Infomatics
School of Frontier Science
The University of Tokyo
7-3-1 Hongo,Bunkyo-ku,Tokyo, 113-8656, Japan
shibata@ee.t.u-tokyo.ac.jp

ABSTRACT

A robust image recognition algorithm has been developed aiming at direct implementation in a bio-inspired hardware accelerator chip. The characteristic features in an original gray-scale image of 64x64-pels (i.e., a 4096-dimension vector) are extracted by a newly-developed Principal Axes Projection (PAP) method and compressed to form a 64-dimension characteristic vector. Despite the large dimensionality reduction, the essential features in the original image are well retained in the vector representation. As a result, very robust image recognition has become possible by using a simple template matching technique. Although the matching with a large number of templates is computationally very expensive, we rely upon the already-developed vector matching LSI's featuring about 1000 GOPS performance for template matching [1-3]. The present algorithm has been successfully applied to medical radiograph analysis and handwriting pattern recognition, and its robust nature in recognition tasks has been proven by intensive computer simulation.

1. INTRODUCTION

Despite the phenomenal progress in computer technology in the hardware performance as well as in the software sophistication, a human-like robust real-time recognition system has not yet been implemented. One of the reasons for this is that the architectures and algorithms of electronic systems are quite different from those of biological systems.

Our approach is based upon a bio-inspired recognition model, where *recognition* is carried out by recalling the most relevant event in the past experiences abundantly stored in our brains. In order to builda image recognition system based on the model, there are two major issues to resolve. One of the issues is that the template matching with a huge database is an extremely computationally-expensive processing. Regarding the problem, however, we have already developed right-brain-like LSI systems based on the digital as well as analog technologies, which are maximum-likelihood search engines aiming at real-time event recognition [4][5]. The most important issue unresolved so far is to find out a robust image representation scheme which is fully compatible to the algorithm employed in the hardware engine chips.

In this paper, we propose a new robust hardware-intensive feature extraction algorithm optimized for our vector matching engines. The algorithm can achieve an efficient dimensionality reduction in the original image without losing their characteristic features, thus providing a robust vector representation scheme for image recognition by template matching. The robustness of the algorithm has been proven by applying to automatic medical radiograph analysis [6,7] as well as to handwriting pattern recognition.

2. IMAGE REPRESENTATION BY PRINCIPAL AXES PROJECTION (PAP)

Since the data in an image are massive in quantity and redundant, an effective dimensionality reduction is a key to build a real-time recognition system. Firstly, the image data must be projected into a vector space having a much smaller dimension. Then it is an essential requirement that the two vectors representing two similar images in appearance are close in the vector space.

The newly-developed scheme of converting a 64x64-pels grayscale image in a 4096-dimension space to a characteristic vector in a 64-dimension space is illustrated in Fig.1.

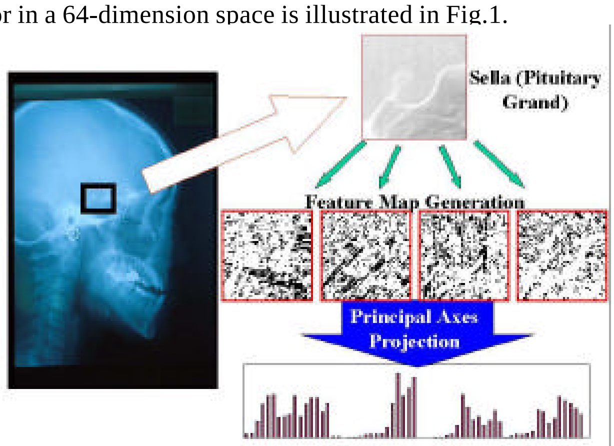


Fig. 1 Characteristic vector formation by Principal Axes Projection (PAP).

The four-direction edges (i.e., horizontal, vertical, +45degrees, -45 degrees) are detected by pixel-by-pixel spatial filtering operations.

Then, four feature maps are generated, each representing a two-dimensional distribution of binary edge flags. Since these feature maps are still two-dimensional and containing voluminous information, they are reduced to a 64-dimensinon vector by a new algorithm which we call Principal Axes Projection (PAP). In the PAP method, edge flags are accumulated in the direction parallel to the direction of each edge and projected onto the axis normal to the direction. Namely, the horizontal edge flags are projected onto the vertical axis, the vertical edge flags onto the horizontal axis, and +45 degrees / –45 degrees to –45 degrees / +45 degrees axes. The data in the four columns are merged to a single vector element, and a 64-dimension characteristic vector is formed by cascade–connecting the four-direction data sets.

3. EXPERIMENTAL

The PAP image representation was applied to the identification of specific anatomical points (Sella, Nasion, and Orbitale) in cephalometric radiographs. Sella is the center of a pituitary grand. Nasion is the most anterior point of the frontonasal suture. Orbitale is the lowest point on the average of the right and left borders of the bony orbit. This is one of the most important practices of dentists in orthodontics. 140 head films were randomly selected from retention files of the Department of Orthodontics at the University of Osaka. Then they were scanned at a resolution of 300dpi and then reduced to an appropriate resolution for experiment by pixel averaging. For the purpose of the algorithm development, eight samples were randomly selected from 140 head films and template vectors were generated from a 64x64-pels block containing each cephalometric landmark. In every search experiment, a head film from one patient was used as an input image and seven vectors taken from other patients were employed as template vectors. The position of a landmark in an input image was detected by pixel-by-pixel matching and the high score positions were identified. Then, the top 50 highest-ranking points were selected as candidates. The top 25 candidates were marked by black dots and second 25 by white dots.

The algorithm was also applied to the handwriting pattern recognition, where the input data were binary images.

4. RESULTS AND DISCUSSION

4.1 Cephalometric Landmarks Identification

Fig. 2 shows the patterns of Sella taken from three patients as examples.

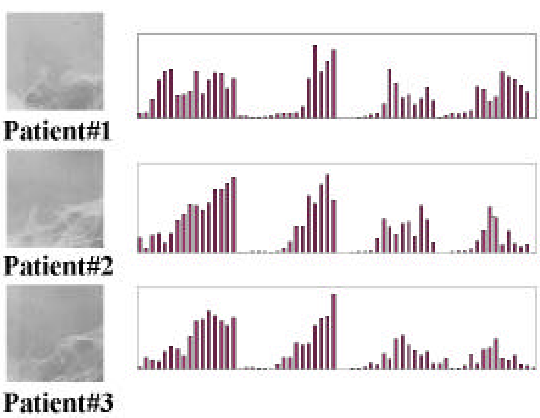


Fig. 2 Vector representations of Sella patterns by the PAP.

Although the images are different from a patient to patient, their vector representations are very similar. Therefore, it is suitable for use in landmark identification by vector matching.

The effect of histogram equalization (HE) [8] in the input image on the robustness of recognition is demonstrated in Fig. 3.

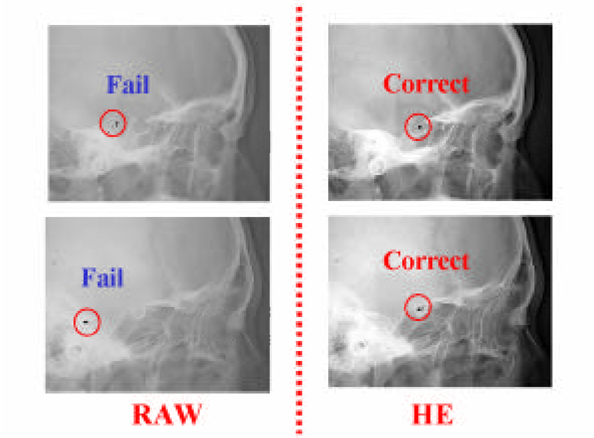


Fig.3 Results of Sella search with (left) or without (right) histogram equalization (HE).

While false points are detected in the Sella search using raw images, correct positions are identified when HE was employed. From the results, quality improvement in an input image by HE is important for robust recognition.

The effect of vector smoothing operation like the one in Eq.1 on the Sella recognition is demonstrated in Fig.4.

SV[k]=(V[k-1]*2+v[k]+V[k+1]*2)/4
SV[0]=V[0],SV[63]=V[63] (Eq.1)

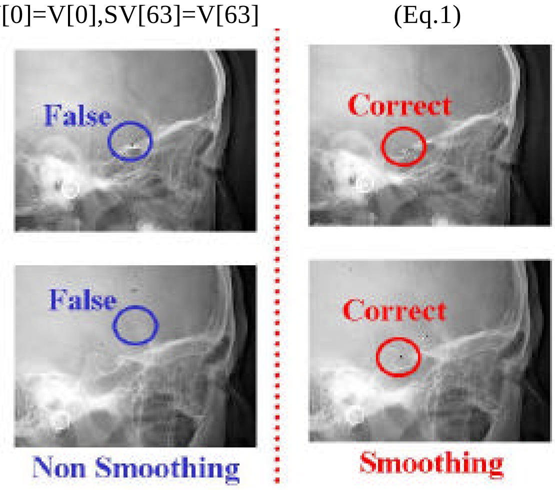


Fig.4 Results of Sella search with (left) or without (right) vector smoothing.

False points are detected in the left images without smoothing, while correct recognition is carried out in the right images with smoothing. It is known that the robustness of recognition is improved by vector smoothing operation.

Results of Sella search experiments are given in Fig.5. Almost all of Sella patterns are recognized correctly. In addition to the correct detection in sample #3, false points are also detected. The false detection was pointed out by an expert dentist as a kind of errors frequently made by humans, thus showing the similarity of our algorithms to human recognition.

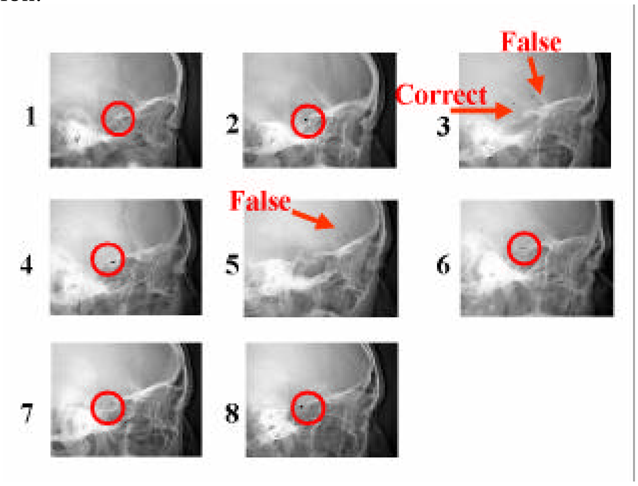


Fig.5 Results of Sella identification.

In sample #5, only false points are detected. However, as shown in Fig.6, the correct location is identified as the local minimum of distance. Therefore this is also easily identified. Then the false detection points in sample #3 and #5, can be removed without difficulty based on the expertise knowledge of dentists.

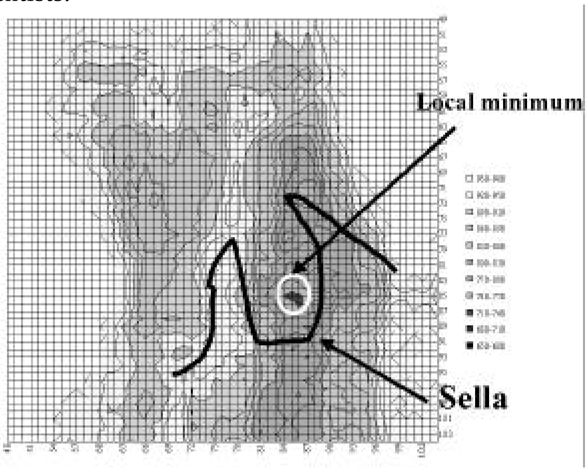


Fig. 6 Dissimilarity (distance) contour map of sample #5.

The experiments were also carried out for the search of Nasion and Orbitale. Orbitale is one of the most difficult points to identify even for expert dentists. However, correct identification was successfully carried out for both Nasion and Orbitale. This is due to the very robust nature of vector representation by PAP as demonstrated in Fig. 7. Although it is not very easy to find similarity to our eyes among the original Orbitale images, the corresponding characteristic vectors all look alike. Thus they are easily detected by vector matching.

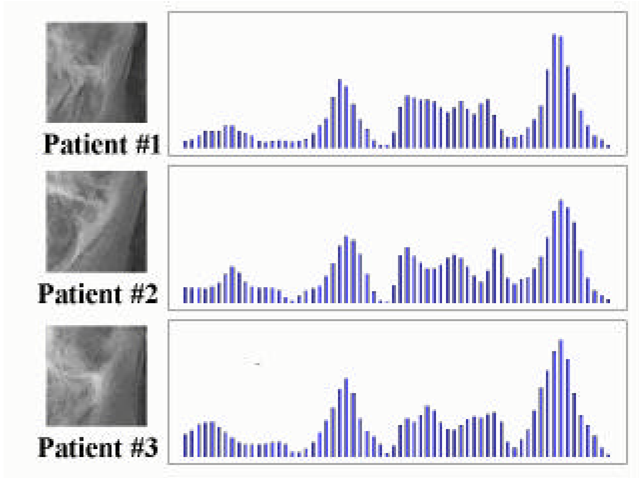


Fig. 7 Vector representation of Orbitale patterns by PAP.

The algorithm thus developed using the eight sample images was applied to a large number of samples (140 head films) to examine its practicality. 40 samples were selected randomly and used as input images. Template vectors were generated from the other 100 samples.

Firstly, the recognition performance was tested as a function of the number of templates, which was varied from five to 50. The percent correct for search of Sella, Nasion, or Orbitale was 60%, 70%, 80% at best, respectively. It was observed that the increase in the number of templates does not necessarily improve the performance. This is because there is an increased chance of including inappropriate template vectors.

In order to solve the problem, we have introduced a learning algorithm to form template vectors. Fig. 8 shows the percent correct as a function of the number of template vectors formed by the Generalized Lloyd Algorithm [9,10]. The 100 samples were used for learning.

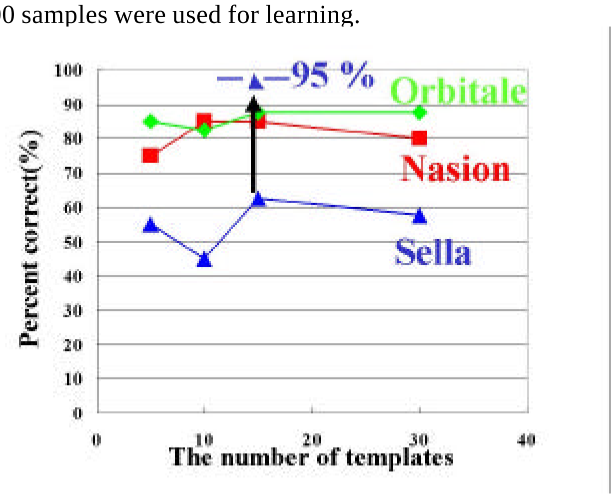


Fig.8 Percent correct of recognition as a function of the number of templates vectors.

The percent correct for Nasion, or Orbitale was improved to 85% or 87.5% at best, while no improvement was obtained for Sella. By taking the positions of local minima, however, the percent correct as high as 95% was achieved for Sella.

4.2 Handwriting Patterns Recognition

The present algorithm has been also successfully applied to the identification of simple handwriting patterns and characters. As an example to show the robustness of our system, the separation of two overlapping patterns is demonstrated in Fig.9. Template patterns are given at the top of the figure.

When the unknown pattern is given as an input image, a circle and a square is recalled as the first and second candidates. When the template of the square recalled is subtracted from the input in the vector space, the circle is recalled with the smallest distance, and vice versa. From these observations, the presented pattern can be identified as being composed of a circle and a square. This is one of the most difficult problems in artificial intelligence, which can be easily done in our system.

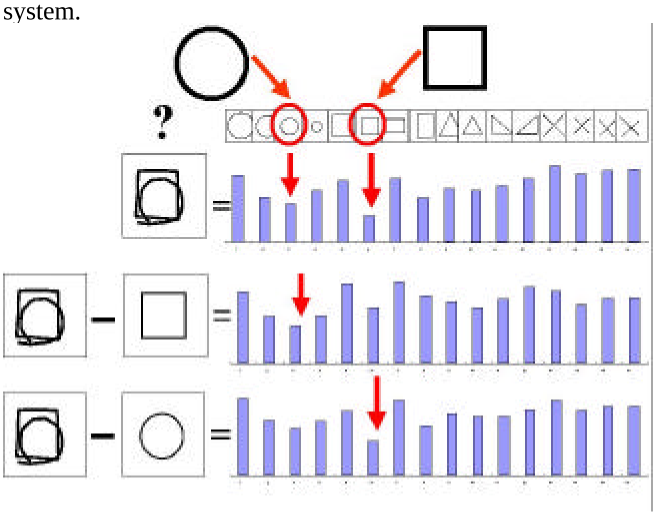
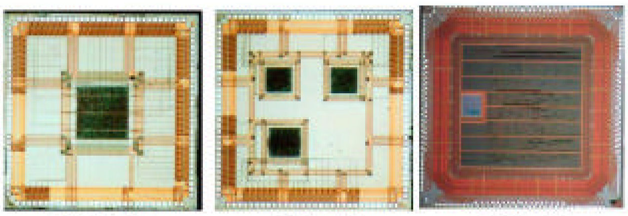


Fig. 9 Separation of overlapping handwriting patterns.

5. Hardware development

Currently it takes several minutes on a workstation (SUN UE450) to carry out a characteristic vector extraction and 30-40 minutes to accomplish cephalometric landmark identification by template matching. It is too slow to establish a real-time-response image recognition system. We have already finished the design of test chips dedicated for the characteristic vector extraction. (See Fig. 10) Judging from the result of logic simulation, the characteristic vector extraction is finished within a micro second. The vector extraction chip in combination with our already developed VQ chips is expected to complete the recognition task in less than 3 msec.



Chip#1 Chip#2 Chip#3

Fig. 10 Photomicrographs of test chips designed through VDEC. Chip#1 and Chip#2 are in a 0.6-um double-poly three-metal-layer technology. Chip#3 is in a 0.35-um single-poly five-metal-layer gate array technology. Chip#1 is for characteristic vector extraction. Chip#2 contains test circuitries for determining the edge detection threshold. Chip#3 was designed as a system for characteristic vector formation.

6. CONCLUSION

We have presented a robust and efficient image recognition algorithm aiming at hardware implementation in the right-brain computing architecture. It has been successfully applied to medical X-ray analysis and handwriting pattern recognition.

ACKNOWLEDGEMENT

The authors would like to thank Prof. Takada, Department of Orthodontics, Faculty of Dentistry, Osaka University for providing us with cephalometric radiograph samples as well as his expertise knowledge of landmark identification. The work is partially supported by the Ministry of Education, Science, Sports and Culture under the Grant-in-Aid for Scientific Research (No. 11305024) and by Japan Science and Technology Corporation in the program of Core Research for Evolutional Science and Technology (CREST).

References

[1] T. Shibata, A. Nakada, M. Konda, T. Morimoto, T. Ohmi, H. Akutsu, A. Kawamura, and K. Marumoto, "A fully-parallel vector quantization processor for real-time motion picture compression," in ISSCC Dig. Tech. Papers, Feb. 1997, pp. 236-237.

[2] 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.

[3] A. Nakada, M. Konda, T. Morimoto, T. Yonezawa, T. Shibata and T. Ohmi, "Fully-Parallel VLSI Implementation of Vector Quantization Processor uisng Neuron-MOS Technology," IEICE trans. Electronics. Vol. E82-C, No. 9, pp. 1730-1737 (1999).

[4] T. Shibata, T. Nakai, N. M. Yu, Y. Yamashita, M. Konda, and T. Ohmi, "Advances in neuron-MOS applications," in ISSCC Dig. Technical Papers, Feb. 1996, SA 18.4, pp. 304-305.

[5] T. Shibata and T. Ohmi, "Neural Microelectronics," Technical Digest, International Electron Devices Meeting (IEDM) 1997, Washington D. C., pp. 337-342.

[6] S. Parthasarathy, S. T. Nugent, P. G. Gregson, and D. F. Fay, "Automatic Landmarking of Cephalograms," COMPUTERS AND BIOMEDICAL RESEARCH 22 (1989), pp. 248-269.

[7] D. J. Rudolph, J. M. Coggins, H. Moon, "Investigation of filter sets for supervised pixel classification of cephalometric landmarks by spatial spectroscopy," International Journal of Medical Informatics 47 (1997), pp. 182-191.

[8] Gerhard X. Ritter, Joseph N. Wilson, "HANDBOOK OF Computer Vision Algorithms in Image Algebra," CRC Press, pp.51-123.

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

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