연상 메모리를 사용한 3차원 물체(항공기) 인식

Associative Memories for 3-D Object (Aircraft) Identification

秦 成
(경북대학교 전자공학과)

■ 차 례 ■

Abstract

The (L,ψ) feature description on the binary boundary aircraft image is introduced of classifying 3-D object (aircraft) identification. Three types for associative matrix memories are employed and tested for their classification performance. The fast association involved in these memories can be implemented using a parallel optical matrix-vector operation. Two associative memories are based on pseudoinverse solutions and the third one is interoduced as a paralell version of a nearest-neighbor classifier. Detailed simulation results for each associative processor are provided.

Ⅰ. Introduction

The general problem we consider is the recongition, classification, and orientation estimation of an object with no stable position. To bound the problem and to allow quantitative data, we consider the recognition and identification of a moving (i.e. airborne) aircraft. We also assume that the object is isolated from the background and this allows us to use feature extraction techniques.

The aircraft image identification problem

has been pursued using various methods. Dudani et. al[1] extracted moment invariants from aircraft image and analyzed these image features using a Bayesian classifier and a K-nearest neighbor classifier with K=10. However, these moment invariants are very noise-sensitive and also have large dynamic range requirements. Fourier descriptors have been studied for shape recognition by several authors. Wallase and Wintz[2] have applied normalized Fourier descriptors to aricraft identification. They used 6 aircraft types and 143 3-D views of each aircraft type as an image data base. The computational load at the matching stage is intensive, compared to our vector-inner product (VIP) associative processor, in which the key feature vector is simply multiplied by a memory matrix to achieve the classification estimation. The syntactic pattern recognition method[3] has been applied to aircraft boundary data. The main drawback of the syntactic method for 3-D object recognition arises when random viewing angles must be handled. In this case, many hundreds of match possibilities must be checked, and the labor needed to derive appropriate grammars becomes prohibitive. Brooks[4] introduced ACRONYM, which is a domain-independent image understanding system. Casasent and Chien[5] introduced rule-based interpreter for aircraft image. Both approaches have merits for high flexibility but this approaches are quite time-consuming to develop and limited in its capabilities for 3-D problems we now consider.

Our associative memory approach is attractive in two respects. First, the parallel matrix vector processor can be employed to enhance the sppeed of matching and classification due to its massive parallelism. Second, the learning procedure is quite simple and the resultant classifier can be easily introduced and tested.

Ⅱ. Associative Processors

Ⅱ-1. Pseudoinverse Associative Memories

We denote the dimension of the input key vectors by N, the number of keyrecollection vector pairs by M and the dimension of the output recollection vector by K. The key vectors are denoted by x_k, the recollection vec tors by y_k and the memory matrix by M.

The association between input and output vectors is performed by a memory matrix described by

$$y_k = Mx_k (k=1,\cdots,M) \tag{1}$$

M is a K×N matrix. We can introduce two rectangular matrices X and Y with each x_k and y_k as their columns. The above equation is then rewritten as

$$Y = MX \tag{2}$$

where $X=[x_1,x_2,\cdots,x_M]$ and $Y=[y_1,y_2\cdots,y_M]$. The memory matrix is solved by the Moore-Penrose method and is given by $M=YX^+$.[6]

Here, X' is the pseudoinverse of X. The recollection of the output vector y' associated with an input key vector x' is the matrix-vector product y'=Mx'. If the key vectors are linealy independent, the recollection vector y' is identical to the original vector asociated with the corresponding key vector x. When Y≠X, the system is termed a heteroassociative memory (HAM). Generally, M is also known as an optimal linear associcative memory

matrix.

Ⅱ-2. Data Matrix Associative Processor

We employ and consider $M_D=X^T$. We refer to this as data metrix(it is known as direct storage memory). This memory is useful, since it gives the key vector with the closest Hamming distance for binary key vectors (if the maximum output elements of y is selected)[7] and this analysis has been extended to analog key vectors[8] under the condition $||x_k||=1$. The data matrix is thus a nearest-neighbor classifier.

Ⅱ-3. Output Coding and Optical Achitecture

As detailed elsewhere[8], the chojce of the y_k affects the performance, memory size, and the storage capacity(M) of an HAM. We consider M vector pairs (differenct object classes and orientations) with a differenct unit output vector y_k denoting each intput, i.e., Y=I in Eq.(2) or $M^+=X^+$. We refer to this as the identity reall vector pseudoinverse matrix M_I. This M_I is quite large N×M (i.e.K=M) with the position of the single "1" in the output denoting the class as well as the orientation of the input.

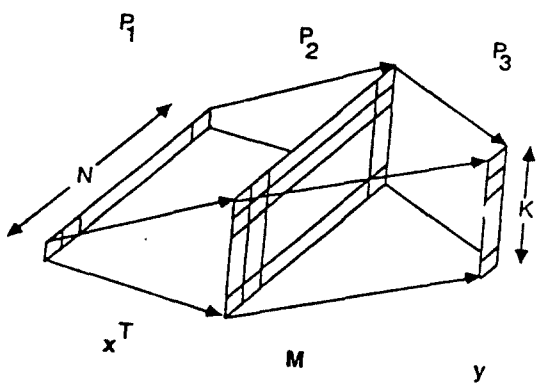

Fig. 1 Optical Matrix Vector Multiplier

The second associative processor matrix we consider uses only C elements (C=6) and the unit recollection vector output now denotes only the class (not the orientation) of the input data. This unit vector matrix $M_{I'}$ is of singnificantly reduced size (N×C) vs (N×M). The third associative processor we consider is the data matrix M_D with the size of N×M we described above. Recollection procedure for these three processors can be paralelly achieved by an optical vector-matrix multiplier schematically shown in Fig. 1

Ⅲ. Database and Key Vectors

Ⅲ-1. Database Generation

As our database we will use an available set of six aircraft(DC10, F104, MiG21, B747, Phantom and Mirage) stored in a model-based description of the vertices of a polygon of each object[9]. We employ binary 128×128 images.

Fig.2 shows top-down views of the six aircraft. We now consider 3-D orientation and first define 3-D orientation angle, θ_1, θ_2 and θ_3 in the coordinate system shown in Fig.3. The aircraft is initially located at the center of the coordinate system shown with its nose in the direction of the -y axis, its right wing in the direction of the +x axis, and the viewer along (-z) axis (i.e., top-down view). The distorted versions of the aircraft are specified by sequentially changing the rotation angles θ_1(pitch), θ_2(roll) and θ_3(yaw) angles.

For building training images used to train the associative memories, we assume that the training images are required at 10° increments $\Delta\theta_1$ and $\Delta\theta_2$. θ_3 variation produce the rotated versions in the direction of the field of view.

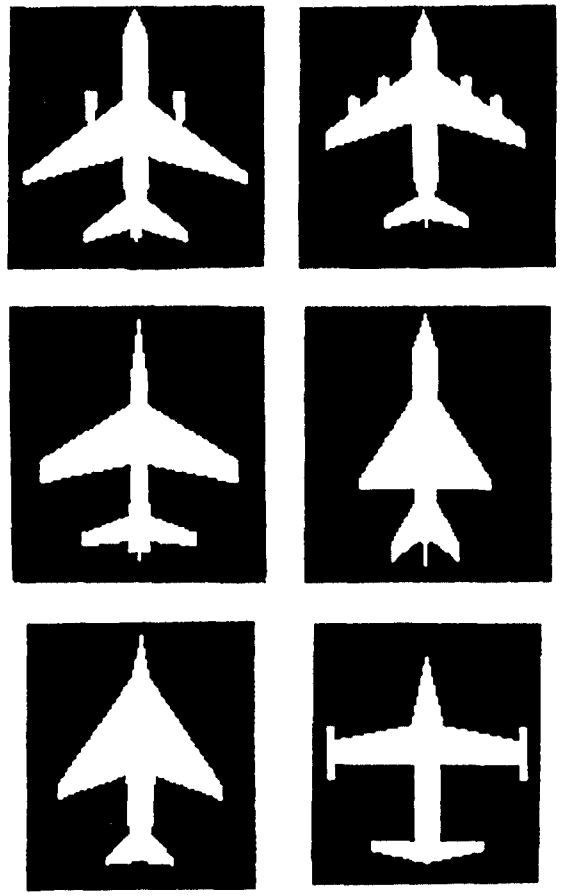

Fig. 2 Image Database (128×128)
(a) DC10 (b) B747 (c) Mirage (d) MiG21
(e) Phantom (f) F104

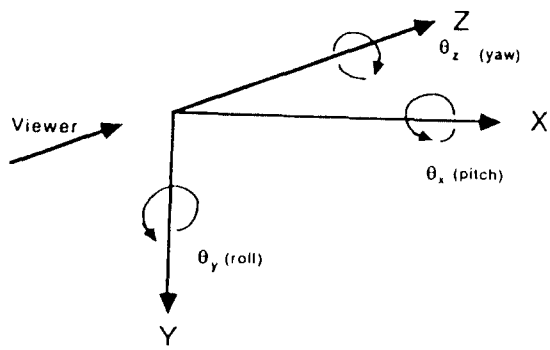

Fig. 3 Coordinate System

They are invariant when we extract invariant features from the objects and thus θ_v is fixed arbitrarily. To provide a meaningful database to test an associative processor, we consider the reduced problem consisting of 6 aircraft types, $-60° \leqq \theta_i < 60°$, and $0° \leqq \theta_v < 60°$ with $\Delta\theta_i = \Delta\theta_v = 10°$. This results in 432 images as the associative processors training set. Each aircraft has 72 different views.

Once the associative memory matrix has been formed form these training set images, we test its performance on the test set data (not present in the training set). The test set images used are at θ_i and θ_v values midway between the training set images (i.e., varies from -55° to 55°in 10° increments and θ_v varies from 5° to 55° in 10° increments). The total number of test images are again 432 for six aircraft.

Ⅲ-2. Key Vector Representation

A feature vector description is adopted for describing one aircraft image. To form an N-dimensional feature space, we first determine the centroid fo the object We then meausre the perimeter of the object and divide it into N equal length intervals. We then draw chords from the centroid to the end points of each arc and thus produce N chords or an L (chord length) objcect description. The starting point is the vertex located furthest form the centroid and thus these features are rotation-invariant. The second feature space considered lists the angle ψ_n for each of the L chords. Finally, the high-dimensional feature space is a combined (L, Ψ) listing. Fig. 4 shows the L_n *and* Ψ_n feature space elements for one arc. Fig.5 shows the preprocessed vertex image of a DC10 with $\theta_i = 50°$, $\theta_v = 20°$, and $\theta = 60°$. If we fix N=144, Fig. 6 shows

the combined (L,Ψ) feature space for Fig.5 (The first 72 elements represent L and the other 72 element Ψ). To achieve scale invariance, the maximum values of L and Ψ are normalized to be one.

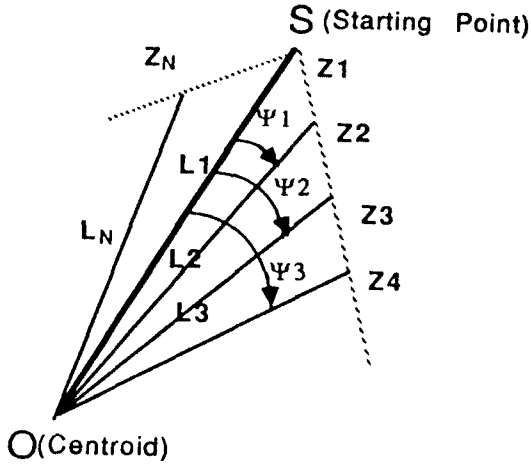

Fig. 4 Schematic showing high-dimensional feature spaces

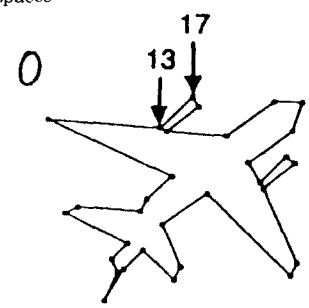

Fig. 5 Vertex representation of a DC10

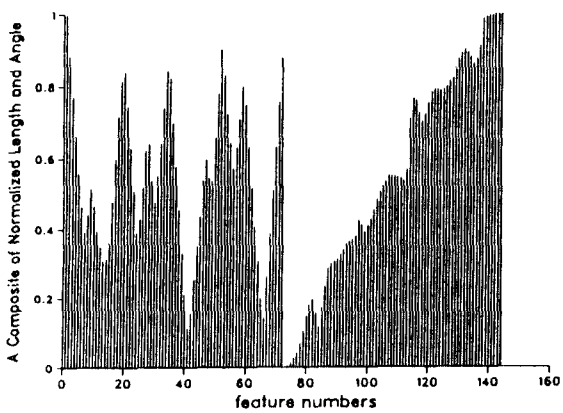

Fig. 6. The combined(L,Ψ) feature space for the object in Fig. 5 with N=144 total samples

Ⅳ. Simulation Results

We now address various major issues in associative process design. These include : N / M ratio effects and analysis of our results that demonstrate and quantify good performance with M>N. In fact, N=144 is fixed since it gives best results.

To determine the effects of N / M on performance (*P*=percent correctly recognized), we first consider the results obtained with tests on the training set images (Fig.7) and

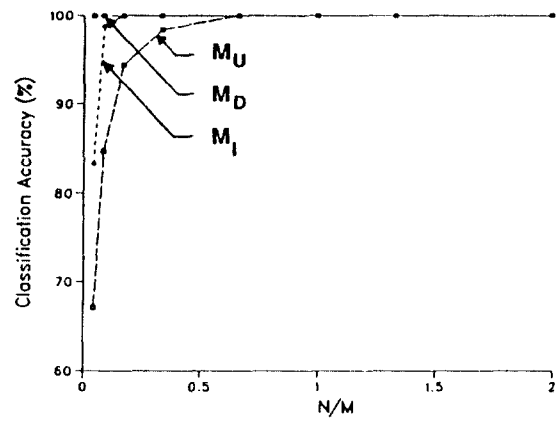

Fig. 7. The Comparative Classification Results of the Three Assocative Processor Memories on Traning Images

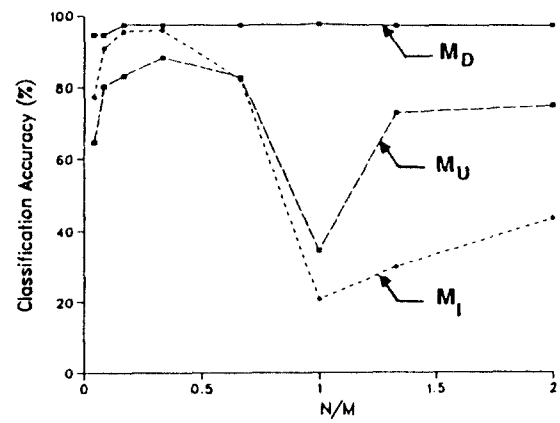

Fig. 8. The Comparative Classification Results of the Three Associative Processor Memories on Test Images

then the results obtained with test set mages (Fig.8). In Fig.7, we note that all three associative processors memories give perfect performance when $N/M \geq 0.667$. The data memory matrix yields perfect performance for any N/M values, M_I yields perfect performance when $N/M \geq 1/6 = 0.167$ and M_U perfect performance when $N/M \geq 2/3 = 0.667$. All three associative processor memories perform well for $M>N$ (i.e., perfect performance on the traning set images).

We now consider P_c vs N/M in Fig.8 for test images (different by 5° from the training set images). We consider the $M>N$ range of interest ($N/M<1$). We note that M_D performs best(as expected since it is a nearest-neighbor classifier).

The X^T memories attempt to produce different vector outputs for the difference in classes. The suprising dip in the performance curve about $M \simeq N$ is noted and analyzed elsewhere[10].

We also find that the M_I memory performs worst and the M_U memory performs betterthan the M_D memory.This is expected since the M_I memory output per class can be viewed as the weighted sum of 72 outputs, and hence will have more output noise than the M_U memory. Recall that if any of the 72 outputs per class in the M_U memory are the largest, we assume that the class is correct and ignore angle distortion estimates from M_U. The P_c performance of thress memories (and the associated memory size) for training and test images are summarized in Table 1. For the training set data, the minimum N at which perfect performance is available is given. For the M_D memory, P_c varies little with N (P_c=95.6% -97.7% as N varies 36 to 288). The M_U memory offers the best combination of P_c and memory size(if performance of 88.4% can be tolerated). Form these data, we find that associative processors with $M>N$ perform well

Table 1. Selected performance of the 3 associative processor with different dimensionality N

	M_I	M_U	M_D
Training Set	Pc=100% NxM≥72x432	Pc=100% NxNc≥288x6	Pc=100% NxM≥18x432
Test Set (Best Pc)	Pc=96.1% N/M=144/432	Pc=88.4% N/M=144/432	Pc=97.7% N/M=288/432
(Other Pc)	Pc=91.0% N=36	Pc=80.3% N=36	Pc=95.6% N=18
	Pc=95.6% N=72	Pc=88.0% N=72	Pc=97.5% N=72

Table 2. Confusion matrix for M_I data (M=432, N=144)

	DC10	F104	MiG21	B747	Phantom	Mirage
DC10	68	0	0	4	0	0
F104	0	72	0	0	0	0
MiG21	1	0	71	0	0	0
B747	0	3	0	68	0	1
Phantom	3	0	0	2	66	1
Mirage	1	1	0	0	0	70

and yield good excellent P_c performance at modest memory size.

Table 2 shows the confusion matrix for $\mathbf{M}_I$ with $P_c=415/432\times100=96.1\%$ at N=144. The first data row indicates that(out of a total of 72 DC10 images) 68 images were correctly classified and 4 images were incorrectly calssified as B747.

V. Summary and Conclusion

From binary aircraft images, we introduced the(L,Ψ) feature space description that is invariant to shift, in-plane rotation, and scale changes. Based on this feature space, we employed and tested three basic associative memories with detailed performance analyses. Much interest is forcused on the memory size and performance. The $\mathbf{M}_D$ memory achieving a nearest-neighbor classifier shows the best performance. The $\mathbf{M}_C$ and $\mathbf{M}_I$ memories are found to perform well even with $M>N$ cases, and this implies that pseudoinverse memories can be used as high capacity memory classifiers.

References

1. S.D. Dudani, K. Breeding, and R. McGhee, "Aircraft Identification by Moment Invariants", IEEE *Trans, on Computers*, Vol. C-26, No.1, January 1977. pp.39-45.
2. T. Wallace and P. Wintz. "Algorithm Using Normalized Fourier Descriptor", *Computer Vision, Graphics, and Image Processing*, Vol.13, 1977, p. 99-126.
3. K.C. You, and K.S. Fu, "Syntactic Shape Recognition Using Attributed Grammars", Purdue University Technical Report TE-EE 78-88, August 1978.
4. Brooks, R.A., "Symbolic Resoning among 3D Models and 2D images", Artificial Intelligence, Vol.17, August 1981, pp.285-348.
5. D. Casasent and S.I. Chien, "Rule-Based Processing for String Code Identification : an Aircraft Case Study", *SPIE Proc.*, Vol. 974, August 1988.
6. T. Kohonen, "Self-Organization and Associative memory", Springer-Verlag, Berlin, 1984.
7. B. Montgomery and B.V.K. Vijaya Kumar, "An Evaluation of The use of the Hopfield Nevral Network Model as a Nearest-Neighbor Algorithm", *Applied Optics*, Vol.25, No. 15, November 1986, pp. 3759-3766.
8. B. Telfer, "Optical Associative Memories for Distortion-Invariant Pattern Recognition", Master's thesis, Carnegie Mellon University, 1987.
9. D. Casasent and S. Liebowitz, "Model-based Knowledge-based Optical Processors", *Applied Optics*, Vol. 26, No.10, May 1987, pp.1935-1942.
10. K. Murakami and T. Aibara, "An Improvement on the Moore-Penrose Generalized Inverse Associative Memory", *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. SMC-17, No. 4, July August 1987, pp. 699-766.

秦 成 一

저자약력

- 1977년 2 월 : 서울대학교 전자공학과 졸업
- 1981년 2 월 : 한국과학기술원 전기및 전자공학과 졸업(석사)
- 1988년 8 월 : 미국 Carnegie Mellon University 졸업(박사)
- 1977년~1979년 : 대영전자(주) (연구원)
- 1981년~현재 : 경북대학교 전자공학과 조교수

※주관심분야 : Computer Vision, Neural Network