2-D FUNCTIONAL AR MODEL FOR IMAGE IDENTIFICATION

Miki Haseyama and Isao Kondo

School of Engineering, Hokkaido University
N-13 W-8 Kita-ku Sapporo 060-8628, JAPAN
{mikich, isao}@media.eng.hokudai.ac.jp

ABSTRACT

This paper proposes a 2-D Functional AR Model for image identification. The definition of the proposed model includes functions that can exploit the self-similarity nature in images to throughly extract image features. By introducing the functional scheme into the model, only a few number of parameters, which are called 2-D Functional AR parameters, can describe the image features simply and accurately. These characteristics make the model suitable for image identification applications. Some experiments of image identification are performed, and the results verify that the proposed model accurately represents the image feature, and the image can be correctly identified. The calculation time is fast enough for practical use in image retrieval.

1. INTRODUCTION

The popular use of the Internet, higher bandwidth access to the network, widespread of digital recording and storage, together with available of powerful and easy to use image processing software, digital content can be easily created, replicated, transmitted, and distributed[1]. Further, it make users easily view materials of remote image databases and keep a large amount of images in local storage or in the network. In such an environment, the users need effective image retrieval systems for quickly finding their desired images from databases based on image content.

In order to realize an effective image retrieval system, we need an accurate image identification scheme, where a query image is compared to a database to execute a one-to-many matching identification. Several model identification methods have been already proposed[2], and some are applicable to 2-D image signals. However, since they are mainly proposed for image restoration[3], recognition[4], indexing [5], or coding[6] rather than for image identification, they cannot be adopted for image retrieval. For example, the identification accuracy generally becomes high if we increase the number of the feature parameters used in the model definition, but this also increases the calculation time greatly, so this practice cannot be employed in an image retrieval system. Therefore, development of an image identification model, which is a suitable for image retrieval, is needed.

In this paper, we present a novel image identification model, which is called a 2-D functional AR model. In the model definition, we introduce some functions into the model to represent image features. These functions are different from the other 2-D AR models[7]. With these functions, the total number of feature parameters is reduced without decreasing the image identification accuracy. This leads to fewer computation cost, and thus it can be utilized for image retrieval. Experimental results verify its accuracy of image identification.

This research is partly supported by Grant-in-Aid for Scientific Research (C) 14550343, Japan Society for the Promotion of Science (JSPS).

2. 2-D FUNCTIONAL AR MODEL

An image x and a white noise image u, whose intensities are $x(i,j)$ and $u(i,j)$ $(i = 0,\dots,L; \quad j = 0,\dots,M)$, respectively, are obtained. The image x is partitioned into size $N \times N$ pixels, $\lfloor L/N \rfloor \times \lfloor M/N \rfloor$ non-overlapping *reference blocks* $X^R_{i,j}$, where i and j denote the horizontal x and vertical y coordinates of the upper-left pixel of a block. In the same way, the image u is partitioned into $\lfloor L/N \rfloor \times \lfloor M/N \rfloor$ blocks $U_{i,j}$. Additionally, the image x is also partitioned into size $\alpha N \times \alpha N$ pixels ($\alpha \in \{1, 2, \dots, \min(\lfloor L/N \rfloor, \lfloor M/N \rfloor)\}$) , overlapping *domain blocks* $X^D_{i,j}$, and the total number of the domain blocks is $(L - \alpha N + 1) \times (M - \alpha N + 1)$. With these partitions, the 2-D functional AR model is defined as follows:

$$U_{i,j} = \sum_{s=-i}^{i-L+N} \sum_{t=-j}^{j-L+N} F_{s,t}(X^D_{i-s,j-t}) \tag{1}$$

where

$$F_{0,0}(X^D_{i,j}) = X^R_{i,j}. \tag{2}$$

When the pixel (k,l) in the domain block $X^D_{i,j}$ has intensity $f_{X^D_{i,j}}(k,l)$ $(1 \le k < \alpha N, 1 \le l < \alpha N)$, and the pixel (k,l) in $F_{s,t}(X^D_{i-s,j-t})$ has $f_{X^{D'}_{i-s,j-t}}(k,l)$ $(1 \le k < N, 1 \le l < N)$, the function $F_{s,t}$ transforms $X^D_{i,j}$ into $X^{D'}_{i,j}$ as follows:

$$f_{X^{D'}_{i-s,j-t}}(x,y) \quad = \quad s_{s,t} f'_{X^D_{i-s,j-t}}(x',y') + o_{s,t}, \tag{3}$$

where

$$\begin{aligned} & f'_{X^D_{i-s,j-t}}(x',y') \\ & \quad = \frac{1}{\alpha^2} \sum_{k=0}^{\alpha-1} \sum_{l=0}^{\alpha-1} f_{X^D_{i-s,j-t}}(\alpha x' - 1 + k, 2y' - 1 + l) \\ & \qquad (x' = 1,\dots,N; \quad y' = 1,\dots,N) \end{aligned} \tag{4}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta_{s,t} & -\lambda_{s,t}\sin\theta_{s,t} \\ \sin\theta_{s,t} & \lambda_{s,t}\cos\theta_{s,t} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \tag{5}$$

where $\theta_{s,t}$ ($\in \{\, 0,\, \pi/2,\, \pi,\, 3\pi/2\ [rad]\}$) is an angle of rotation; $\lambda_{s,t} \in \{1, -1\}$; e and f are coordinate offsets, which are automatically set by the other parameters; and $s_{s,t}$ and $o_{s,t}$ are the parameters of the affine transformation. The parameters $\theta_{s,t}$, $\lambda_{s,t}$, $s_{s,t}$ and $o_{s,t}$ are called *2-D functional AR parameters*.

The function $F_{s,t}$ defined in Eqs. (4) and (5) is similar to the transformation in the global IFS method[8] for image coding[9]. Though any functions can be chosen for the proposed functional AR model, we adopt $F_{s,t}$ since it well represents the self-organization property for the images, and is thus suitable for the autoregressive scheme in our model.

3. COMPUTATION FOR PARAMETERS OF 2-D FUNCTIONAL AR MODEL

For a given image x, the parameters of the 2-D functional AR model are computed by minimization of the following criterion:

$$E(x(i,j), \hat{x}(i,j)) \triangleq \sum_{i=1}^{L}\sum_{j=1}^{M} (x(i,j) - \hat{x}(i,j))^2 \qquad (6)$$

where

$$\hat{x} = \bigcup_{\substack{i=0,N,2N,\ldots,N\lfloor L/N \rfloor \\ j=0,N,2N,\ldots,N\lfloor M/N \rfloor}} \hat{X}^R_{i,j} \qquad (7)$$

and

$$\hat{X}^R_{i,j} = \sum_{s=-i}^{i-L+N}\sum_{t=-j}^{j-L+N} \hat{F}_{s,t}(X^D_{i-s,j-t}). \qquad (8)$$

where the parameters of $\hat{F}_{s,t}$ are denoted by $\hat{\theta}_{s,t}$, $\hat{\lambda}_{s,t}$, $\hat{s}_{s,t}$, and $\hat{o}_{s,t}$; and

$$\hat{F}_{0,0}(X^D_{i,j}) = 0_{N\times N}. \qquad (9)$$

After the minimization, we can obtain the optimal 2-D functional AR parameters. However, it requires an extensive calculation because the criterion includes too many parameters, and moreover, the optima of $\theta_{s,t}$ and $\lambda_{s,t}$ need to be determined by a search algorithm. Therefore, in order to reduce the calculation costs, we simplify the model definition in Eq.(1) to the following:

$$\hat{X}^R_{i,j} = \hat{F}_{s_{i,j},t_{i,j}}(X^D_{i-s_{i,j},j-t_{i,j}}) \qquad (10)$$

where $s_{i,j} \in \{1,\ldots,L_M\}$; $t_{i,j} \in \{1,\ldots,L_M\}$; and $\hat{F}_{s_{i,j},t_{i,j}}(X^D_{i-s_{i,j},j-t_{i,j}})$ satisfies

$$\hat{F}_{s_{i,j},t_{i,j}}(D_{i-s_{i,j},j-t_{i,j}}) = \min_{\substack{s=1,\ldots,L_M \\ t=1,\ldots,L_M}} E(R_{i,j}, \hat{F}_{s,t}(X^D_{i-s,j-t})). \qquad (11)$$

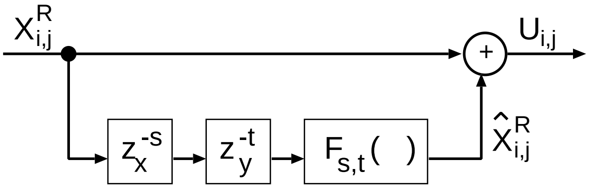

Fig. 1. Simplified 2-D Functional AR Model: z_x^{-s} and z_y^{-t} denote shift operators for x and y axes, which work as $z_x^{-s}z_y^{-t}D_{i,j} = D_{i-s,j-t}$.

The simplified model is shown in Fig. 1. This simplification does not decrease the identification quality, while can reduce the calculation costs. It will be verified in Section 5 with experiments.

4. IMAGE IDENTIFICATION BASED ON 2-D FUNCTIONAL AR MODEL

A query image $x(i,j)$ is given, and its 2-D functional AR parameters are computed by the method described in Section 3. By using the 2-D functional AR model, we can identify whether an image in a database and the query image x are the same or similar.

(i) Image Identification Measurement E_1

Suppose an image in the database is g, and its intensities are $g(i,j)$ $(i = 1,\ldots,L;\quad j = 1,\ldots,M)$. Now $\hat{g}(x,y)$ is computed by applying the 2-D functional AR parameters of the query image x as

$$\hat{g} = \bigcup_{\substack{i=0,N,2N,\ldots,N\lfloor L/N \rfloor \\ j=0,N,2N,\ldots,N\lfloor M/N \rfloor}} \hat{G}^R_{i,j} \qquad (12)$$

$$\hat{G}^R_{i,j} = \sum_{s=-i}^{i-L+N}\sum_{t=-j}^{j-L+N} \hat{F}_{s,t}(G^D_{i-s,j-t}). \qquad (13)$$

The function $\hat{F}_{s,t}$ is determined by the 2-D functional AR model computed from the query image x. An identification measurement can be defined as follows:

$$E_1 = E(x(i,j), \hat{g}(i,j)) \qquad (14)$$

For this identification measurement, (1) g is identified with the same query image x if $E_1 = 0$; or (2) g is considered similar to the query image x if $E_1 > C_{Th}$, where C_{Th} is a predefined threshold. Since we have to compute $\hat{G}^R_{i,j}$ for E_1 in the image retrieval application, the computational costs may become a problem if the database is huge. In this case, the following measurement E_2 becomes a solution.

(ii) Image Identification Measurement E_2

Suppose the 2-D functional AR model parameters of each image are pre-calculated and also stored in the database. The calculation costs for image identification can be reduced with the following measurement:

$$E_2 = \sum_{i=0}^{L-N}\sum_{j=0}^{M-N} I\left((s_{i,j},t_{i,j}),(\hat{s}_{i,j},\hat{t}_{i,j})\right) \qquad (15)$$

where $\hat{s}_{i,j}$ and $\hat{t}_{i,j}$ are the parameters computed from the image g in the database, and the function $I(x)$ is defined as

$$I\left((s_{i,j},t_{i,j}),(\hat{s}_{i,j},\hat{t}_{i,j})\right) = \begin{cases} 1 & \text{if } (s_{i,j},t_{i,j}) = (\hat{s}_{i,j},\hat{t}_{i,j}) \\ 0 & \text{otherwise.} \end{cases} \qquad (16)$$

For this identification measurement, (1) g is identified with the same query image x if $E_2 = (L-N)\times(M-N)$; or (2) g is considered similar to the query image x if $E_2 > C_{Th}$.

Since this measurement is not defined with any multiplications, the calculation costs can be reduced largely.

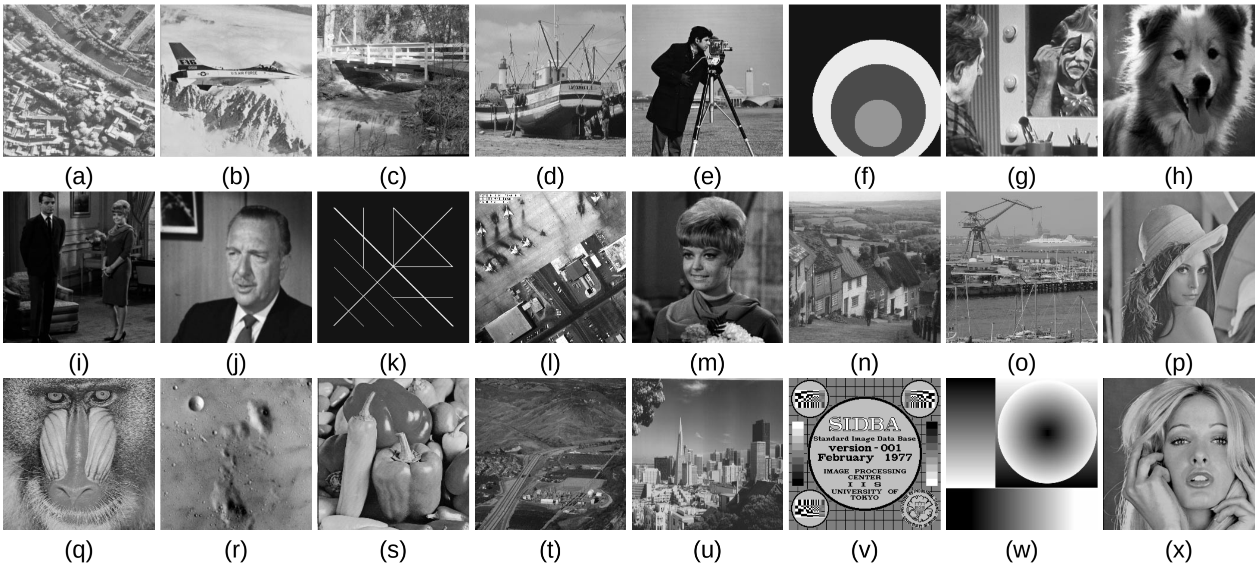

Fig. 2. Original Images:
24 images, (a) Aerial, (b) Airplane, (c) Bridge, (d) Boat, (e) Cameraman, (f) Circles (g) Clown, (h) Collie, (i) Couple, (j) Cronkite, (k) Crosses, (l) Eltoro, (m) Girl, (n) Goldhill, (o) Harbour, (p) Lena, (q) Mandrill, (r) Moon, (s) Peppers, (t) Plant, (u) San Francisco, (v) SIDBA Title, (w) Splope, and (x) Woman are used as query images and for constructing the database.

5. EXPERIMENTAL RESULTS

The proposed 2-D functional AR model is applied to image identification problems. In the experiments, we use 24 kinds of original images which are all with size of 256 × 256 pixels and 255 gray levels (Fig. 2). Each image is processed with 18 different ways as follows to construct a database containing 432 image objects.

(i) – (iii) Smoothing (1)
The median filters of sizes 2×2, 3×3, and 4×4 pixels are applied to each original image.

(iv) Smoothing (2)
A smoothing filter $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ is applied to each original image.

(v) – (xvi) JPEG compression
Twelve kinds of JPEG encoded images with parameters 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, and 90% quality without smoothing are prepared for each original image.

(xvii) Sharpening
A sharpening filter $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$ is applied to each original image.

(xviii) Random geometric distortion
With the software *stirmark* [1], each original image is contaminated by added distortion.

[1]version3.1, http://www.cl.cam.ac.uk/~fapp2/ watermarking/ stirmark/

The following identification measurement, which is slightly modified of E_2 is used.

$$E_2' = \frac{1}{10}\sum_{k=1}^{5}\sum_{(i,j)\in R} I\left((s_{i,j}^k, t_{i,j}^k), (\hat{s}_{i,j}, \hat{t}_{i,j})\right) \qquad (17)$$

where R consists of 10 *reference blocks* with from 1-st to 10-th highest variance selected from all the blocks in the query image; $(s_{i,j}^1, t_{i,j}^1)$ corresponds to $(s_{i,j}, t_{i,j})$ in Eq. (11), and $(s_{i,j}^k, t_{i,j}^k)$ is of the k-th smallest $E(R_{i,j}, \hat{F}_{s,t}(X_{i-s,j-t}^D))$. The 2-D functional AR model is realized with $N = 16$ and $\alpha = 2$; each domain block can be overlapped with less than 16 pixels; and the parameters λ and θ are fixed at 1 and 0, respectively, because of reduction of the computation.

The original images are used as query images. Each query image is compared to all 432 images in the database and 432 values of E_2' are computed and plotted in Fig. 3. We can see that the values always become bigger, or meaning well matched, only when the query image is compared with those generated from it. For example, Fig. 3(a) for query image Aerial shows the first 18 values are equal to or close to 100% and the remains are at 20% or less. The results clearly demonstrate that the 2-D functional AR model based measurement is suitable for the image identification. Furthermore, the calculation time for each query image in 1-to-432 comparisons is about 20 ms with a 450 MHz Pentium II processor. The speed can be further improved by some other hierarchical search algorithms. The calculation time is proportional to the size of the database. Therefore, the proposed model can be utilized for image retrieval even in a huge database.

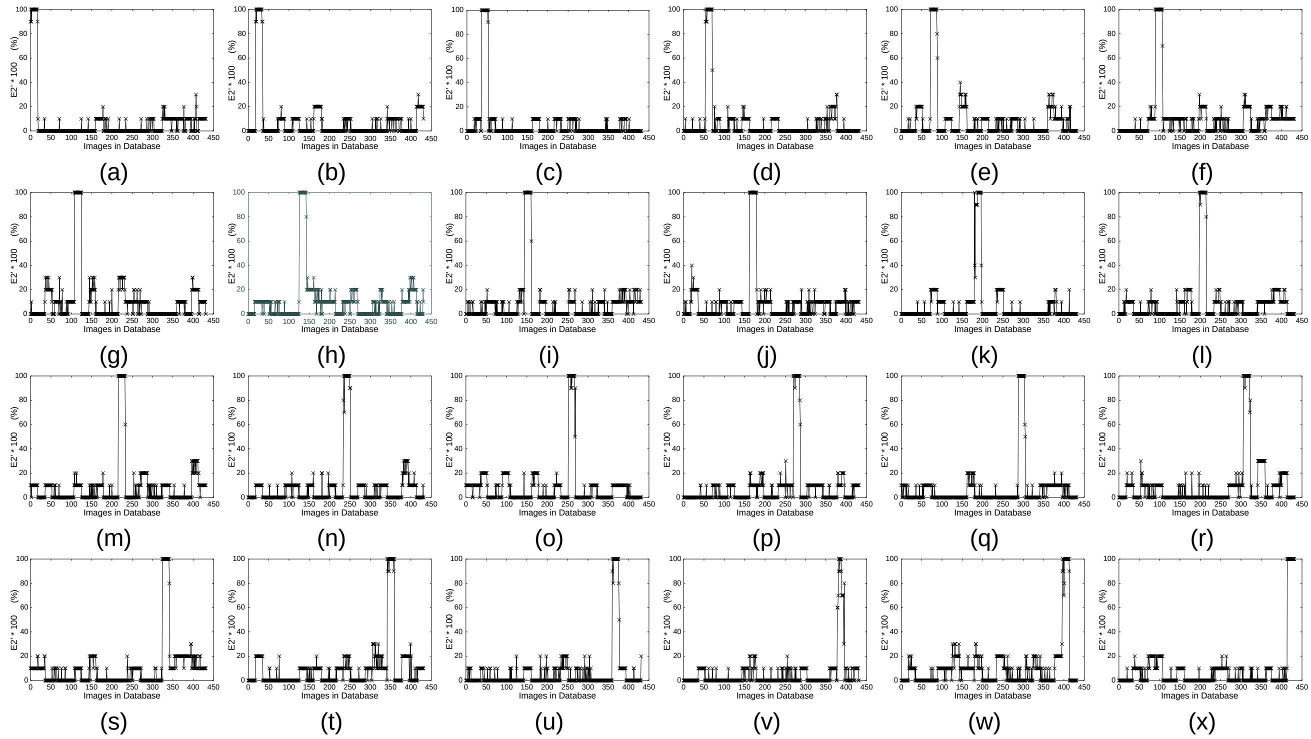

Fig. 3. Identification results:
The values $100E_2'$ (%) are computed and plotted for each query image of (a) Aerial, (b) Airplane, (c) Bridge, (d) Boat, (e) Cameraman, (f) Circles (g) Clown, (h) Collie, (i) Couple, (j) Cronkite, (k) Crosses, (l) Eltoro, (m) Girl, (n) Goldhill, (o) Harbour, (p) Lena, (q) Mandrill, (r) Moon, (s) Peppers, (t) Plant, (u) San Francisco, (v) SIDBA Title, (w) Splope, and (x) Woman. The x-axis represents 432 images in the database which are ordered as 24 blocks corresponding to 24 original images (a)-(j). Each block consists of 18 image objects generated from its corresponding original image.

6. CONCLUSIONS

We have proposed the 2-D functional AR model for image identification. Though its formulation is similar to the AR model, it is based on the blockwised scheme, and its definition includes the functions that can represent the self-similarity of images. With the proposed model, the images can be accurately modeled with only a few parameters. Consequently, the model is suitable for image retrieval. The experimental results verify the accurate performance of the model.

7. REFERENCES

[1] A. Tefas and Ioannis Pitas, "Image authentication using chaotic mixing system," in IEEE Int. Symposium on Circuit Syst., vol. I, 2000, pp. 216–219.

[2] S. Haykin, *Adaptive filter theory.* Englewood Cliffs, NJ: Prentice Hall, 1991.

[3] Hussen Youlal, Malika Janati-I, and Mohamed Najim, "Two-dimensional joint process lattice for adaptive restoration of images," IEEE Trans. Image Processing, vol.1, pp. 366–378, July 1992.

[4] V. F. Vasiliev, "Recognition of 2-D object on geometrical quasisimilarity of polygonal representations," in IEEE Int. Conf. Systems, Man, and Cybernetics, vol. 5, pp. 4396–4400, 1998.

[5] R. Stoica, J. Zerubia, and J. M. Francos, "The two-dimensional Wold decomposition for segmentation and indexing in image libraries," in Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, pp. 2977 – 2980, 1998.

[6] P. A. Maragos, R. W. Schafer, and R. M. Mersereau, "Two-dimensional linear prediction and its application to adaptive predictive coding of images," IEEE Trans. Acoust., Speech & Signal Process, vol. ASSP-32, no.6, pp. 1213–1229, Dec. 1984.

[7] J. M. Francos and A.Z. Meiri, "A 2-D autoregressive, finite support, causal model for texture analysis and synthesis," in IEEE Int. Conf. Acoustics, Speech, and Signal Processing, vol. 3, 1989, pp. 1552–1555.

[8] A. Jacquin, "Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations," IEEE Trans. Image Processing, vol.1, pp. 18-30, Jan. 1992.

[9] M. F. Barnsley, and L. Hurd, "Fractal Image Compression," AK Press, Wellesley, 1993.