IMAGE CODING USING A WAVELET BASED ZERNIKE MOMENTS COMPRESSION TECHNIQUE

G. A. Papakostas[1], D. A. Karras[2], and B. G. Mertzios[1]

[1] Democritus University of Thrace, Department of Electr. and Comp. Emg., 67100 Xanthi, Greece
e-mail: gpapakos@ee.duth.gr, mertzios@ee.duth.gr
[2] Hellenic Aerospace Industry and University of Hertfordshire, Rodu 2, Ano Iliupolis, Athens 16342, Greece,
e-mail: dakarras@hol.gr, dkarras@haicorp.com, dakarras@usa.net

Abstract: In this paper, we present a new method of image coding using two popular imaging tools, Zernike moments and Wavelets. The main idea is that we can produce appropriate image descriptors by involving an appropriate number of moments, compressed in a form suitable to represent an image with low reconstruction error for pattern recognition applications. At this point the concept of wavelet compression is involved, which has already been discussed in many technical papers. We use an existent wavelet based compression algorithm, to compress not the 2-D image, but the resulted moment based 1-D signal. So, using this formulation we can achieve a compressed representation of the image, suitable for pattern recognition purposes and image retrieval tasks. It is very important to notice here the ability of Zernike moments to provide a very high level of image reconstruction, using the inverse wavelet transform, establishing a useful method.

1. INTRODUCTION

Have been passed four decades, since Hu [1] introduced the concept of moment invariants and the use of image moments for 2-D pattern recognition. Since then, many moment-based techniques have found wide applications [2,3,4].

This paper presents a processing sequence, consisting of three general stages. At the first stage, translation, scale and rotation invariant Zernike moments are extracted from a 2-D image. At the second stage the 1-D wavelet decomposition is applied to the resulted signal that has been constructed from the moments. At the third stage a compression algorithm is applied to the derived wavelet coefficients, in order to keep the useful content of the moments based signal.

Wavelets have played important role in image processing, as a powerful tool for filtering, denoising, compression [5] etc. The role of wavelets in image coding is important, in the sense that they can give features containing large information from the image [6,7]. Thus, a new class of descriptors has been obtained, the wavelet descriptors [8,9,10], which are complementary to Fourier descriptors, since they have additional advantages.

The power of the features is completed by the ability to reconstruct the original image from their values. This procedure is not always simple, because the inverse procedure must exist for every process that is applied to the image. This demand is satisfied in our approach and its efficiency is derived using a measurement that describes the reconstruction error.

2. ZERNIKE MOMENTS

The desired properties that distinguished features must have in any pattern recognition system are translation, scale and rotation invariance. These demands have guided for the investigation of methods in order to derive invariant features.

In the proposed method we try to construct a one-dimensional "moment signal", using a modified version of Zernike moments, to achieve all the necessary invariances.

Let consider a MxM 2-D binary image with intensity function $f(x,y)$, resulted from a gray level image with a simple binarization method using thresholding. To keep the dynamic range of m_{pq} consistent for different size images, the MxM image plane is first mapped onto a square defined by $x \in [-1,+1]$, $y \in [-1,+1]$ [11]. This kind of normalization took place for one more reason. That is, the Zernike moments to be constructed in this section, are defined over the interior of the unit circle, $x^2 + y^2 = 1$ The regular geometrical moments of this image defined as

$$m_{pq} = \sum_{x=-1}^{+1} \sum_{y=-1}^{+1} x^p y^q f(x,y) \mathrm{dxdy} \tag{1}$$

where m_{pq} is the (p+q)th order moment of the image.

This type of moments does not have any of the desired invariances, so they are not suitable for pattern recognition tasks. To achieve translation invariance we describe each image point according to its centroid. Working by this way we derive the well known *central moments*

0-7803-7503-3/02/$17.00 ©2002 IEEE

$$\mu_{pq} = \sum_{x=-1}^{+1} \sum_{y=-1}^{+1} (x-\bar{x})^p (y-\bar{y})^q f(x,y) \quad (2)$$

where, $\bar{x} = m_{10}/m_{00}$, $\bar{y} = m_{01}/m_{00}$ are the centroid coordinates of the original image.

To ensure scale invariance we can define the *normalized central moments*

$$\eta_{pq} = \frac{\mu_{pq}}{\mu^{\gamma}{}_{00}} \quad , \ \gamma = \frac{p+q}{2} + 1 \quad (3)$$

At this point we have a modified type of geometrical moments, called normalized central moments that satisfy two of our three invariance demands. The last one is rotation invariance, an important property that any good distinguished feature must have.

There have been presented several methods to obtain rotational independence, the most popular of them using orthogonal polynomials as Zernike, Legendre etc. The necessity of using orthogonal polynomials, except that we can secure the rotation invariance property, is the fact that the regular geometrical moments are the projection of $f(x,y)$ onto the monomial $x^p y^q$. The basis set $x^p y^q$ is not orthogonal and this results to information redundancy. This drawback of regular moments can be fixed by using an orthogonal basis of functions that also provides a very useful and efficient reconstruction property.

In our experiments we use the Zernike moments for obtaining invariant moment features, to construct the "moment signal" that will be processed in order to extract useful information from our image.

Zernike introduced a set of complex polynomials, which form a complete orthogonal set over the interior of the unit circle $x^2 + y^2 = 1$. These polynomials [11,12] have the form

$$V_{nm}(x,y) = V_{nm}(\rho,\vartheta) = R_{nm}(\rho)\exp(jm\vartheta) \quad (4)$$

where n is non-negative integer, m is a non zero integer subject to constraints n-$|m|$ even and $|m| \leq n$, ρ is the length of vector from origin $(\bar{x},\bar{y})$ to (x,y) pixel and θ the angle between vector ρ and x axis in counter-clockwise direction, $R_{nm}(\rho)$ are the Zernike radial polynomials in (ρ,θ) polar coordinates defined as

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|/2} (-1)^s \cdot \frac{(n-s)\,!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!} \rho^{n-2s} \quad (5)$$

Note that $R_{n,-m}(\rho) = R_{nm}(\rho)$

These polynomials are orthogonal and satisfy the orthogonality principle

$$\iint_{x^2+y^2\leq 1} [V_{nm}(x,y)]^* V_{pq}(x,y)\,dxdy = \frac{\pi}{n+1}\delta_{np}\delta_{mq} \quad (6)$$

where $\delta_{\alpha\beta}$ =1 for α=β and $\delta_{\alpha\beta}$ =0 otherwise, is the Kronecker symbol.

The Zernike moment of order n with repetition m for a digital image with intensity function $f(x,y)$ that vanishes outside the unit disk is

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) V^*_{nm}(\rho,\theta), \quad x^2 + y^2 \leq 1 \quad (7)$$

The rotation invariance property of these Zernike moments has been already analyzed [12]. These investigations led to the conclusion that the magnitudes of Zernike moments are invariant to any rotation of the image. Thus, we can use for our experiments the magnitudes of the resulted Zernike moments beyond a high order.

Because the Zernike moments are only rotationally invariant, we must give to these moments the additional properties of translation and scale invariance, in some way. As discussed in the previous section, we can ensure these invariances by converting the absolute pixel coordinates (2), (3).

Guided from these equations we achieve translation invariance by transforming the image into one whose origin is the image centroid. In other words, the origin is moved to the centroid before moment calculation. The new image has intensity function $f(x+\bar{x}, y+\bar{y})$ [12].

We obtain scale invariance by enlarging or reducing each object such that its zeroth-order moment m_{00}, is set to a predetermined value, β [12]. Thus the original image function is transformed into a new function $f(x/a,y/a)$ with $\alpha=(\beta/m_{00})^{1/2}$.

Finally, an image function $f(x,y)$ can be normalized with respect to scale and translation by transforming it into

$$g(x,y) = f\left(\bar{x} + \frac{x}{\alpha}, \bar{y} + \frac{y}{a}\right) \quad (7)$$

In our experiments we use as invariant moment features the following Zernike moments magnitudes

$$|A_{nm}| = \left| \frac{n+1}{\pi} \sum_x \sum_y g(x,y)\, V^*_{nm}(\rho,\vartheta) \right| \quad (8)$$

These magnitudes are translation, scale and rotation invariant and therefore, they are useful to any pattern recognition application.

From these moments the initial image can be obtained, using the reconstruction formula

$$\hat{f}(x,y) = \sum_{n=0}^{n_{max}} \sum_m A_{nm} V_{nm}(\rho,\theta) \quad (9)$$

An alternative formula for reconstruction can be found in [12]. Note that as n_{max} approaches infinity $\hat{f}(x,y)$ will approach $f(x,y)$.

3. WAVELET COMPRESSION

Wavelet theory constitutes a very useful tool in image processing [13]. In some sense, we can say that it comes to complement the so important Fourier theory.

Applying the 1-D discrete wavelet transform (1-D DWT) to the "moment signal", obtained in the previous processing step, using the next definition

$$W_\psi\left(f\left(k2^{-s},2^{-s}\right)\right) = 2^{s/2}\int_{-\infty}^{\infty} f(t)\cdot\psi\left(2^s t - k\right)dt \quad (10)$$

where ψ is the *mother wavelet*, the decomposition of our signal can be achieved as follows

$$f(t) = \sum_k u_{j_0,k}\cdot\phi_{j_0,k}(t) + \sum_{j=j0}^{\infty}\sum_k w_{j,k}\cdot\psi_{j,k}(t) \quad (11)$$

with

$$u_{j,k} = W_\varphi\left(f(j,k)\right),\; w_{j,k} = W_\psi\left(f(j,k)\right) \quad (12)$$

the scaling and wavelet coefficients respectively.

In the above equations φ is the relative to the mother function *scaling function*, *j,k* are indices of the translation and dilation parameters and j_0 represents the coarsest scale.

This yields a number of wavelet coefficients from which we can reconstruct the original signal. The importance of the wavelet transform is that captures local characteristics of the signal and in this way we have a localized view of the signal's behavior.

Bearing in mind that our aim is to keep the least coefficients possible without losing useful information, we apply to this set of coefficients a compression algorithm that uses the following simple soft thresholding (shrinkage) procedure

$$Y = \begin{cases} \text{sign}(x)\left(\left(|x| - \text{thr}\right)\right)_+ , & |x| > \text{thr} \\ 0 , & |x| \le \text{thr} \end{cases} \quad (13)$$

where x is the input signal, Y is the compressed version of x and *thr* is the compression threshold that can be set manually or by using a specified algorithm.

Doing this, we obtain a truncated set of coefficients from which we can reconstruct the initial image.

4. PROPOSED METHOD

This section describes the new set of invariant features that resulted after several processing stages. The flow diagram is depicted in figure (1).

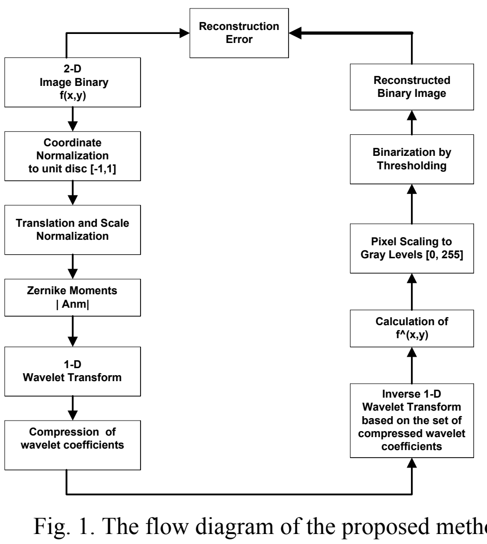

Fig. 1. The flow diagram of the proposed method

The desired steps that perform the procedure shown in the above figure, can be summarized as follows:

Step 1: Taking a 2-D binary image, we transform the image coordinates in such way, that the image is mapped to a unit disk [-1,1].

Step 2: Transform the image density function $f(x,y)$ to a translation and scale invariant version as shown in equation (7). At this step we get the $g(x,y)$ intensity function.

Step 3: Computation of Zernike moment magnitudes with the help of equation (8). Using these measurements, we construct a one-dimensional signal, which consists of those magnitudes. The way to place the magnitudes is the number of generation of each magnitude, which also constitutes the order that each moment participates in the reconstruction procedure. So, first we position the $|A_{00}|$ moment, second the $|A_{11}|$ and so on.

Step 4: Decomposition of the "moment signal" using equation (11) and the wavelet transform (10), results to a set of wavelet coefficients (12) able to reconstruct the original "moment signal".

Step 5: At the last step we perform a compression procedure onto the set of wavelet coefficients, applying thresholding (13).

The above processing yields a compressed set of Zernike moments. To reconstruct the original image from these compressed moments, we perform exactly the inverse steps.

The inverse process is described in Fig. (1), in which we can see that the inverse path is legal since all the intermediate steps have their inverse procedures.

After the completion of this "loopback" operation we take a measurement of the effectiveness of the proposed method that is the reconstruction error. We define as reconstruction error the Hamming distance between the original and the reconstructed images

$$\text{error} = \sum_x\sum_y\left|\hat{f}(x,y) - f(x,y)\right| \quad (14)$$

with $\hat{f}(x,y), f(x,y) = 0$ or 1.

5. SIMULATION RESULTS

To study the capabilities of the proposed method, a number of experiments have been conducted. The simulation results have been obtained using the MATLAB software package.

The experiments were implemented according to the following settings: β=800, binarization threshold equal to 128, images of size 256x256, one level wavelet decomposition using *haar* wavelet, n=0,1,.30, and m according to the same constraints with equation (4).

With these settings a set of 256 translation, scale and rotation invariant Zernike moments were derived, that were used to construct the "moment signal". In this signal the wavelet based compression algorithm was applied that resulted in a signal that can be reconstructed from 25% fewer wavelet coefficients.

These moments based wavelet coefficients are the features we need.

In Fig.2, a sample of the original (solid line) and the compressed (dashed line) “moment signal” consisting of the last 50 values are depicted.

Instead of the description of the test image with the 256 Zernike moments, we can represent this image using the 192 resulted wavelet coefficients, from which we can construct the compressed “moment signal” and thus the original image.

Therefore we can represent the initial image with 25% fewer invariant features, and of course we can reconstruct it with minimum reconstruction error.

In Fig. 3 the original and the reconstructed image are illustrated, using these wavelet based Zernike moments. The reconstruction error with this method was 547 pixels, meaning that the original image has been reconstructed loosing 0.87% of its pixels. This error is about 0.85% using the usual Zernike moments. Errors are almost the same, but using the proposed method, the feature vector corresponding to the image has been reduced by about 25%.

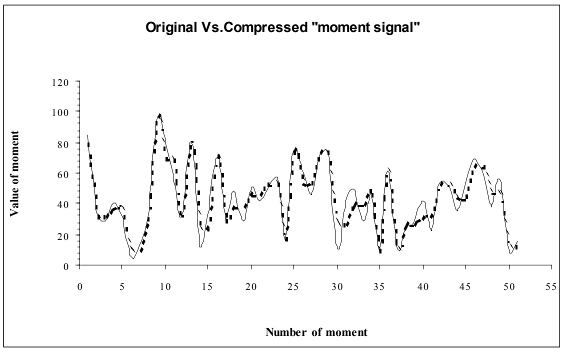

Fig. 2. Original vs. Compressed “moment signal”

(a) (b)

Fig. 3. (a) original image, (b) reconstructed image from the compressed Zernike moments

6. CONCLUSIONS

A novel method for image representation was presented in this paper. We used a combination of the moment’s theory and the wavelet transform, to derive invariant features, proper for pattern recognition applications.

In this paper we investigate the ability of these features to carry enough information of the original image, so we can reconstruct it with optimal reconstruction error. The simulation results are very promising, for further investigation of this approach. Future work will emphasize in finding an optimal method of constructing the “signal moment” and in testing these features in real world pattern recognition problems.

REFERENCES

[1] M.K. Hu, “Visual pattern recognition by moment invariants”, *IRE Trans. Inform. Theory,* vol. IT-8, pp. 179-187, Feb. 1962.

[2] S. Dudani, K. Breeding, R. McGhee, “Aircraft identification by moment invariants”, *IEEE Trans. Comput*, C-26, pp. 39-45, 1977.

[3] K. Tsirikolias, B.G. Mertzios, “Statistical pattern recognition using efficient two-dimensional moments with applications to character recognition”, *Pattern Recognition*, vol. 6, pp. 877-882, 1993.

[4] M. Zhenjiang, “ Zernike moment-based image shape analysis and its application”, *Pattern Recognition* Letters, vol. 21, pp. 168-177, 2000.

[5] P.N. Topiwala, “Wavelet Image and Video Compression” , Kluwer Academic Publishers, 1998.

[6] D.A. Karras and S.A. Karkanis and B.G. Mertzios, “Texture Discrimination for Quality Control Using Wavelet and Neural Network Techniques”, *Proceedings IWISSP ’96,* pp. 191-194, 4-7 November 1996, Manchester U.K.

[7] S.A. Karkanis, D.A. Karras and B.G. Mertzios, “A Novel Signal Modeling Method Using the Wavelet Transform”, *Neural Parallel & Scientific Computations*, vol. 6, pp. 417-428, 1998.

[8] G.C.-H. Chang and C.-C. Jay Kuo, “Wavelet Descriptor of Planar Curves: Theory and Applications”, *IEEE Trans. on Image Processing.,* vol. IP-5, No.1, pp.56-70, 1996.

[9] P. Wunsch and A. F. Laine, “Wavelet Descriptors for Multiresolution Recognition of Handprinted Characters”, *Pattern Recognition,* vol. 28, No. 8, pp. 1237-1249, 1995.

[10] D. Shen, Horace H.S. Ip, “Discriminative wavelet shape descriptors for recognition of 2-D patterns”, *Pattern Recognition*, vol. 32, pp. 151-165, 1999.

[11] A. Khotanzad, J.-H. Lu, “Classification of Invariant Image Representations Using a Neural Network”, *IEEE Trans. on Acoustics, Speech and Sign. Processing,* vol.38, No.6, pp. 1028-1038, 1990.

[12] A. Khotanzad and Y.H. Hong, “Invariant Image Recognition by Zernike Moments”, *IEEE Trans. on. Pattern Anal. Machine Intell.,* vol. PAMI-12, No.5., pp. 489-497, 1990.

[13] J.C. Goswami, A.K. Chan, “Fundamentals of Wavelets, Theory, Algorithms and Applications”, *Wiley Publications* 1999.