모양에 기반한 영상 검색을 위한 2-D Invariant Descriptor

박종승° 장덕호
한국전자통신연구원 시각정보팀
park@etri.re.kr dhchang@etri.re.kr

2-D Invariant Descriptors for Shape-Based Image Retrieval

J. S. Park D. H. Chang
Visual Information Team, ETRI, 161 Kajong-Dong, Yusong-Gu, Taejon, 305-350, South Korea

요 약

모양 정보를 이용하는 내용기반 영상 검색 시스템에서 검색 정확도는 시스템에서 사용되는 모양 기술자에 매우 의존된다. 정확한 검색을 위해서 기술자는 이동, 회전, 스케일에 불변해야 한다. 본 논문에서는 모멘트 불변량과 푸리에 기술자를 복합적으로 사용하는 유사도 기법을 제시한다. 이 방법은 하나의 불변량 기술자를 사용하는것보다 더 우수한 결과를 나타내었다. 푸리에 기술자와 네개의 모멘트 불변량(Hu의 모멘트 불변량, Taubin의 모멘트 불변량, Flusser의 모멘트 불변량, Zernike 모멘트 불변량)을 구현하여 성능을 측정하였다. 영상분할된 이진 영상 데이터베이스로부터 각 기술자의 검색 정확도를 계산하였다. 실험 결과 경계선에 기초하는 푸리에 기술자와 영역에 기초하는 모멘트 불변량을 동시에 사용하는 방법이 영상 검색에 있어서 우수한 성능을 보였다.

Keywords: Shape-based image retrieval, Fourier descriptors, Moment invariants, Object recognition

1. Introduction

There have been numerous works on similarity measures for 2-D objects. Among them, moment invariants (Hu's moment invariants [1,2,3], Taubin's moment invariants [4], Flusser's moment invariants [5], and Zernike moments [2,3,6]) and Fourier descriptors [3] are considered as the most representative features in 2-D shape matching. All of the measures are invariant to translation, scale change, and rotation in 2-D space.

In general, the invariant property does not hold in a digital image due to the generic problem of the discrete function. We overcome the numeric problem by employing a two-stage similarity scheme. Note that moment invariant is the region-based measure and Fourier descriptor is the boundary-based measure. First, we compute moment invariants to extract relevant images. Then, the retrieval is verified using Fourier descriptors, which increase the retrieval effectness substantially.

In Section 2-3, we describe invariant properties of moment invariants and Fourier descriptors. that are used in our method as similarity measures for image retrieval. Then, in Section 4, our similarity scheme which uses both moment invariants and Fourier descriptors is described. Experimental results are presented in Section 5. We conclude this paper with discussion and future works in Section 6.

2. Fourier descriptors

Fourier descriptors are 2-D invariant features available from boundary points. Suppose that the boundary of a particular shape has N pixels numbered from 0 to $N-1$ and the contour is described as two parametric equations: $x(k) = x_k,\ \ y(k) = y_k,\ \ k = 0,\ldots,N-1$. By considering the equations in the complex plane, the direct parametric representation $z(t)$ is possible: $z(t) = x(t) + jy(t)$. The Fourier descriptors $Z(k)$ of the curve is the discrete Fourier transform coefficients of the complex valued curve $z(t)$: $Z(t) = \frac{1}{N}\sum_{k=0}^{N-1} z(k)\exp\left(\frac{-j2\pi kt}{N}\right)$. A simple normalization of $Z(t)$ makes the Fourier descriptors invariant to the starting point of sampling, rotation, scaling and translation. Each coefficient of a Fourier descriptor has two components, amplitude and phase. By using only the amplitude component, we achieve rotation invariance as well as the invariance to the starting point. By dividing all amplitudes by the amplitude of the first non-zero frequency coefficient, we achieve the scale invariance. Since only the DC coefficient is dependent on the position of shape, it is discarded to achieve the translation invariance. We compute the m-dimensional feature vector F_F from m Fourier descriptors $Z(-m/2),\ldots,Z(-1),Z(2),\ldots,Z(m/2+1)$ by dividing the magnitudes by $|Z(1)|$. In our system, we choose $m = 16(=2^4)$ so that the transformation can be conducted efficiently using FFT.

3. Moment Invariants

Moment invariants are useful in 2-D object recognition. Moment invariants are functions of moments that are invariant under certain transformations. Although, moments are defined on a continuous image intensity function, a simple approximation is possible for a discrete binary image using summation operation. Let f be a binary digital image matrix with dimension $M \times N$, and let $S = \{(x,y)|f(x,y)=1\}$ represent a 2-D shape. The *moment* of order (p,q) of shape S is given by $m_{pq}(S) = \sum_{(x,y)\in S} x^p y^q$. The *central moment* of order (p,q) of shape S is given by $\mu_{pq}(S) = \sum_{(x,y)\in S}(x-\bar{x})^p(y-\bar{y})^q$ where $(\bar{x},\bar{y})$ is the center of gravity: $\bar{x} = m_{10}(S)/m_{00}(S)$, $\bar{y} = m_{01}(S)/m_{00}(S)$. From the central moments, the *normalized central moments*, denoted by η_{pq}, are defined as $\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}$ where $\gamma = \frac{p+q}{2}+1$ for $p+q = 2,3,\ldots$.

3.1 Hu's moment invariants

From the second- and third-order normalized central moments, a set of seven invariant moments, which is invariant to translation, scale change and rotation, has been derived by Hu[7]:

$$M_{H1} = \eta_{20} + \eta_{02}$$

$$
\begin{aligned}
M_{H2} &= (\eta_{20}-\eta_{02})^2+4\eta_{11}^2\\
M_{H3} &= (\eta_{30}-3\eta_{12})^2+(3\eta_{21}-\eta_{03})^2\\
M_{H4} &= (\eta_{30}+\eta_{12})^2+(\eta_{21}+\eta_{03})^2\\
M_{H5} &= (\eta_{30}-3\eta_{12})(\eta_{30}+\eta_{12})\left[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2\right]\\
&\quad +(3\eta_{21}-\eta_{03})(\eta_{21}+\eta_{03})\left[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2\right]\\
M_{H6} &= (\eta_{20}-\eta_{02})\left[(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2\right]\\
&\quad +4\eta_{11}(\eta_{30}+\eta_{12})(\eta_{21}+\eta_{03})\\
M_{H7} &= (3\eta_{21}-\eta_{03})(\eta_{30}+\eta_{12})\left[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2\right]\\
&\quad -(\eta_{30}-3\eta_{12})(\eta_{21}+\eta_{03})\left[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2\right]
\end{aligned}
$$

Hu[7] has proved the invariance properties of the seven moments for the case of continuous functions. For each object, we compute the feature vector $M_H = (M_{H1},\ldots,M_{H7})$ to obtain the similarity measure.

3.2 Taubin's moment invariants

Taubin and Cooper[4] defined affine moment invariants by introducing the concept of covariant matrix. From the central moments of order 2 of shape S, a 2×2 matrix is computed : $M_{11} = \begin{bmatrix} \mu'_{20} & \mu'_{11} \\ \mu'_{11} & \mu'_{02} \end{bmatrix}$ where $\mu'_{pq} = \mu_{pq}/m_{00}(S)$. From M_{11}, a lower triangular matrix $L = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix}$ such that $LM_{11}L^T = I$ is found using cholesky decomposition and matrix inversion. From the center of gravity $(\bar{x},\bar{y})$ and the matrix L, the new moments η'_{pq}, for $p+q>2$, are found using the equation :

$$\eta'_{pq} = \frac{1}{m_{00}(S)} \sum_{(x,y)\in S} \left[l_{11}(x-\bar{x})\right]^p \left[l_{21}(x-\bar{x})+l_{22}(y-\bar{y})\right]^q .$$

From the 3rd-4th order moments, two matrices are found :

$$M'_{12} = \begin{bmatrix} c_1\eta'_{30} & \eta'_{21} & c_1\eta'_{12} \\ c_1\eta'_{21} & \eta'_{12} & c_1\eta'_{03} \end{bmatrix},\ M'_{22} = \begin{bmatrix} c_2\eta'_{40} & c_2\eta'_{31} & c_2\eta'_{22} \\ c_1\eta'_{31} & \eta'_{22} & c_1\eta'_{13} \\ c_2\eta'_{22} & c_1\eta'_{13} & c_2\eta'_{04} \end{bmatrix}$$

where $c_1 = \frac{1}{\sqrt{2}}$ and $c_2 = \frac{1}{2}$. The eight moment invariants $M_T = (M_{T1},\ldots,M_{T8})$ are two eigenvalues of the 2×2 symmetric matrix $M_{12}M_{12}^T$, three eigenvalue of the 3×3 symmetric matrix M_{22}, two eigenvalue of the 2×2 symmetric matrix $M_{12}M_{22}M_{12}^T$, and the scalar $\frac{1}{4}(\eta'_{40}+\eta'_{04})$.

3.3 Flusser's moment invariants

Flusser and Suk[5] derived affine moment invariants which are invariant under general 2-D affine transformations.

$$
\begin{aligned}
M_{F1} &= \mu_{00}^{-4}(\mu_{20}\mu_{02}-\mu_{11}^2)\\
M_{F2} &= \mu_{00}^{-10}(\mu_{30}^2\mu_{03}^2-6\mu_{30}\mu_{21}\mu_{12}\mu_{03}+4\mu_{30}\mu_{12}^3+\\
&\quad 4\mu_{03}\mu_{21}^3-3\mu_{21}^2\mu_{12}^2)\\
M_{F3} &= \mu_{00}^{-7}((\mu_{20}(\mu_{21}\mu_{03}-\mu_{12}^2))-\\
&\quad \mu_{11}(\mu_{30}\mu_{03}-\mu_{21}\mu_{12})+\mu_{02}(\mu_{30}\mu_{12}-\mu_{21}^2))\\
M_{F4} &= \mu_{00}^{-11}(\mu_{20}^3\mu_{03}^3-6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03}-\\
&\quad 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03}+9\mu_{20}^2\mu_{02}\mu_{12}^2+12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03}+\\
&\quad 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03}-18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12}-8\mu_{11}^3\mu_{30}\mu_{03}-\\
&\quad 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12}+9\mu_{20}\mu_{02}^2\mu_{21}^2+12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12}-\\
&\quad 6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21}+\mu_{02}^3\mu_{30}^2)\\
M_{F5} &= \mu_{00}^{-6}(\mu_{40}\mu_{04}-4\mu_{31}\mu_{12}+3\mu_{22}^2)\\
M_{F6} &= \mu_{00}^{-9}(\mu_{40}\mu_{04}\mu_{22}+2\mu_{31}\mu_{22}\mu_{12}-\mu_{40}\mu_{12}^2-\mu_{04}\mu_{31}^2-\mu_{22}^3)
\end{aligned}
$$

3.4 Zernike moment invariants

The Zernike moment of order n with repetition m that vanishes outside the unit circle is

$$A_{nm} = \frac{n+1}{\pi} \sum_{(x,y)\in S,\ x^2+y^2\le 1} V^*_{nm}(\rho,\theta)$$

where $R_{nm}(\rho)$ and V^*_{nm} are defined in [6]. The magnitudes of the Zernike moments, $|A_{nm}|$, are invariant to rotation. To achieve scale and translation invariant property, we translate the data points so that the origin is moved to the centroid and scale the points so that the maximum distance from the origin is equal to 1. The normalization affects the first two features, $|A_{00}|$ and $|A_{11}|$. From second to fifth order moments, we compute the vector M_Z from the 10 features, $|A_{20}|, |A_{22}|, |A_{31}|, |A_{33}|, |A_{40}|, |A_{42}|, |A_{44}|, |A_{51}|, |A_{53}|, |A_{55}|$.

4. Similarity Measure

All of the descriptors described above have invariant properties to rotation, translation, and scale changes. The invariance properties of the features hold for the case of continuous functions. In discrete case, the set of moment invariants is still invariant under image translation although the moments are computed discretely. But the invariants are expected not to be strictly invariant under rotation and scale changes due to sampling, digitizing, and quantizing of the continuous image for digital computation[1]. Moment invariants are not sufficient for distinguishing all shapes, they can be very sensitive to noise, and their values drastically change with occlusion.

Figure 1 shows the invariant properties. Each two plots from the first row to the fifth row corresponds to the invariant property plots of Fourier descriptors, Hu's moment invariants, Taubin's moment invariants, Flusser's invariants, and Zernike moment invariants. Source object is an arrow shape image of size 337 × 145. The transform parameter is a normalized scalar (from 0 to 1) for the translation factor, the rotation factor, and the scaling factor. The translation factor ranges from -400 to 400 in pixel units with step size 8, the rotation factor ranges from 0 to 360 in degree with step size 3.6, and the scaling factor ranges from -2 to 2 with step size 0.04. All of the descriptors maintained good invariant properties during the transformation although Taubin's moment invariants oscillated within a narrow range.

Note that moment invariants are the region-based measure and Fourier descriptors are the boundary-based measure. Mehtre et. al[8] showed that the measure using both Fourier descriptors and moment invariants gives the best average retrieval efficiency. They thought this could be because the human perceptual mechanism uses both these aspects of shape in order to compute similarity. We followed the idea in computing the similarity of two objects.

Contrary to their method of similarity computation, we employ two-stage similarity scheme. First, we compute moment invariants to extract relevant images. Then, verification using Fourier descriptors is followed which increases the retrieval effectness substantially.

The similarity distance between two feature vectors is computed as the Euclidean distance.

First we compute the distance using a feature vector of moment invariants among the four kinds of moment invariants. The first n images of closest distance are selected as candidates for output. Among the n images, we reorder the sequence using the distance of Fourier descriptors and we select $m < n$ images in decreasing order of distance of Fourier descriptors.

5. Experimental Results

We have evaluated the retrieval effectiveness of the invariant descriptors described in the previous section. In our application, Hu's moment invariants plus Fourier descriptors and Zernike moment invariants plus Fourier descriptors showed good performance relative to other descriptors.

An example of evaluation is shown in Table 1. The database contains 85 animal images. Among the images, there are 10 rabbit images. For a query rabbit image, the number of retrieved rabbit images out of the first five(ten) retrieved images are shown in the table.

The method using both moment invariants and Fourier descriptors showed better results rather than method using moment invariants only or method using Fourier descriptors only,

The computation time on an Indigo2 IMPACT with a MIPS R10000 processor is shown in Table 2. The test image of size 337 with 718 boundary points and 17761 region points.

Unexpected results may appear when there is a 3-D perspective effect in the shape since the invariance holds only when the deformation is a kind of 2-D affine transformation. Bad results may also appear when the two boundaries are from the same object but a boundary was too much smoothed by a region extraction module.

6. Conclusion

In this paper we introduced invariant features that may be used for shape-based image retrieval. We considered Fourier descriptors, Hu's moment invariants, Taubin's moment invariants, Flusser's moment invariants, and Zernike moments. Among the invariants, combination of Fourier descriptors and moment invariants (Hu's moment invariants or Zernike moments) showed effective image retrieval.

As future works of our research, we are developing an image segmentation algorithm which extracts only objects of interest regardless of the complexity of the environment where the object is located in. Future research should also include the shape matching of partially recovered objects or objects having occlusions.

References

[1] C.-H. Teh and R. T. Chin, "On digital approximation of moment invariants," *CVGIP*, vol. 33, 1986.

[2] C.-H. Teh and R. T. Chin, "On image analysis by the methods of moments," *IEEE T-PAMI*, vol. 10(4), 1988.

[3] N. Ezer, E. Anarim, and B. Sankur, "A comparative study of moment invariants and fourier descriptors in planar shape recognition," in *Proceedings of 7th Mediterranean Electrotechnical Conference*, vol. 1, pp. 242–245, 1994.

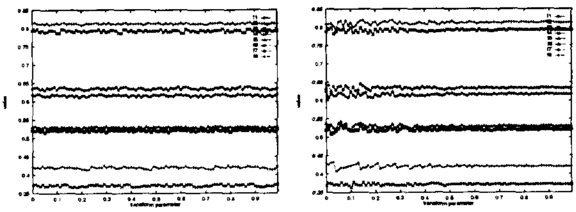
[4] G. Taubin and D. B. Cooper, "Object recognition based on oment (or algebraic) invariants," in *Geometric invariance in computer vision* (J. L. Mundy and A. Zisserman, eds.), pp. 375–397, MIT Press, 1992.

[5] J. F. T. Suk, "Pattern recognition by affine moment invariants," *Pattern Recognition*, vol. 26, pp. 167–174, 1993.

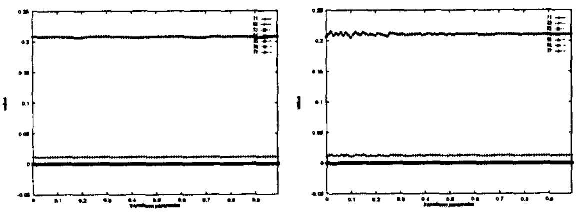
[6] A. Khotanzad and Y. H. Hong, "Invariant image reconstruction by zernike moments," *IEEE T-PAMI*, vol. 12(5), pp. 489–497, 1990.

[7] M. K. Hu, "Pattern recognition by moment invariants," in *Proc. IRE*, vol. 49, 1961.

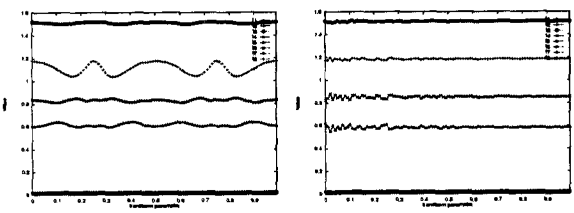
[8] B. M. Mehtre, M. S. Kankanhalli, and W. F. Lee, "Shape measures for content based image retrieval: a comparison," *Information Processing and Management*, vol. 33, no. 3, 1997.



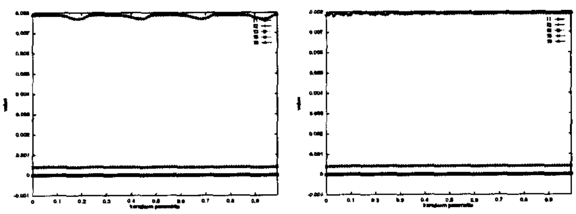
Fourier descriptors (FD)



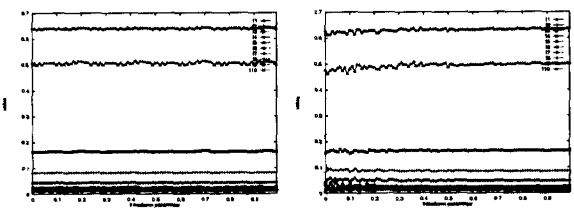
Hu's moment invariants (HM)



Taubin's moment invariants (TM)



Flusser's moment invariants (FM)



Zernike moments (ZM)

Figure 1: Invariant property of descriptors.
Left: rotation, Right: scaling

query\method	FD	HM	TM	FM	ZM	FH	FZ
query 1	5/8	4/7	4/7	2/3	5/9	5/8	5/9
query 2	5/8	5/8	4/5	2/5	5/9	5/8	5/9
query 3	5/8	2/6	1/1	1/3	4/8	5/8	5/8
query 4	5/8	5/7	4/6	1/3	5/9	5/8	5/9

Table 1: Comparison to other similarity schemes. (Top 5/Top 10) FH=FD+HM, FZ=FD+ZM

method	FD	HM	TM	FM	ZM	FH	FZ
time	0.01	0.02	0.02	0.02	0.93	0.03	0.94

Table 2: Comparison of computation time. (unit=sec)