International Journal of Information Technology, Vol. 12 No. 2 2006

The shape recognition based on structure moment invariants [1]

Li Zongmin[1, 2, 3], Kunpeng Hou[1],Liu Yujie[1, 2, 3], Diao Luhong[2, 3] , Li Hua[2]

1 School Of Computer Science And Communication Engineering,University Of Petroleum, Shandong China 257062
2 Institute of Computing Technology Chinese Academy of Science, Beijing 100080, P.R.China
3 Graduate School of Chinese Academy of Science, Beijing 100039, P.R.China
(zmli,lihua,yjliu,lhdiao@ict.ac.cn)

Abstract

Object shape comparability is a challenging problem in the field of pattern recognition and computer vision. The method based on the geometric moment invariants is the typical method in these applications. This paper introduces structure moment invariants based on the geometric moment invariants from transforming the density in geometric moments into a new density. The difference in the shapes is increased by using the structure moment invariants. Therefore, this method can be used in object shape analysis. To support our new theory , an algorithm for object shape analysis is designed and experiments based on square transform are conducted. Experiments give an encouraging high recognition rate by using the structure moment invariants.

1 Introduction

The shape of an object is a very important character in human's perception, recognition, and comprehension. Because geometric shape represents the essential characteristic of an object, and has invariance with respect to translation, scale, and orientation, the analysis and discernment like geometry is of important significance in computer vision.

The shape of an object has versatility, which results in many kinds of expression ways. Thus far there is no way that is totally in accord with human's recognition and comprehension of shape [1]. No matter in the search of 2-D data or in the 3-D model, researchers has proposed a lot of methods that describe the shape characteristics.

There already has a great deal of work on the research of matching in shape characteristic of the 2-D objects [2] [3] [4]. The expression ways of 2-D picture contour have Fourier descriptor [5], moment invariants [6] [7] [8] and boundary energy function [9], and etc. And there is some research which is based on shape matching of 3-D objects, such as expanded gauss method which is used in 3-D surface [10], the

[1] This paper is supported by National High-Tech Research and Development Plan(grant No: 2001AA231031), National Key Basic Research Plan (grant No: 2004CB318000) and National Special R&D Plan for Olympic Games (grant No: 2001BA904B08).

spherical harmonic function [11], 3-D Zernike moment invariants function [12] and topological structure method [13], and etc. Among them moment invariants have good invariance on translation, orientation, scale and anti-interference, thus they gained many important achievements in research. Unfortunately, these descriptors usually provide an insufficient discrimination between complicated objects. For this reason, this paper attempts to develop an invariant moment method from a new perspective, which is used to describe the shape characteristic of complicated objects.

This paper is organized as follows. Firstly, structure moment invariants are defined in Section 2. The difference in the shapes is increased by using the structure moment invariants.Then, some retrieved examples are presented to analyze the validity of our method in Section 3. We conclude in Section 4 by summarizing our results and discussing topics for future work.

2 Descriptions of Structure Moment Invariants

We can use the projection of object function $f \in L2$ on the area of Ω to define the moment μi which is used in analysis of object's shape. The function on the area of Ω was defined as $\Psi = \{\psi i\}$ where $i \in N$, and ψi is defined as follows:

$$\mu i = \langle f, \psi i \rangle = \int_{\Omega} f(x) \bullet \overline{\psi_i(x)}\, dx \qquad (2.1)$$

All research was concentrated on function Ψ in the past 20 years. With the difference of Ψ, different moments were accordingly produced. For example, when non-orthonormal basis xpyq was given, we could get Geometric Moments. In order to better study various characteristics of moment, Teague proposed in his paper that if we use orthonormal basis--Zernike and Legendre multinomial to replace non-orthornormal basis, we got the Zernike moments and the Legendre moments. We can also get wavelet moment if we use wavelet basis. All of these methods have played an important role in practical applications, such as pattern recognition, scenery matching, and the image sequence analysis, and etc. However, in practical applications, we find that using the methods said above, a lot of problems with complicated structure are difficult to solve, such as complicated shape being difficult to recognize.

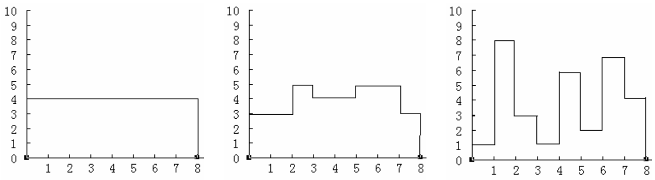
In information optics, we see that there is such a thought which provide the analytical method to the structure complexity of the 2-D objects [14].

Whether the "structure" of a picture is abundant or not, means that is it sharply or gently of the picture's light intensity when varying with the position. The degree of abundant structure of the 2-D object is consistent with the integral as follows:

$$S_F = \iint_F I_0^2 (x, y)\, dxdy$$

That is to say, based on the premise that the total light energy is given definitely in area F, the bigger S_F is, the more abundant the structure of the 2-D object is.

$$\iint_F I_0 (x, y)\, dxdy = \text{const}$$



(a) (b) (c)

Fig 1 examples of the distribution of light intensity

Just like the distribution of light intensity in Fig 1, the total energy of each one is 32. However, the integral value will be different after square of three integrable functions respectively:

$$\text{Fig 1(a): } \int I_0^2 \, dx=128$$

$$\text{Fig 1(b): } \int I_0^2 \, dx=134$$

$$\text{Fig 1(c): } \int I_0^2 \, dx=180$$

Namely the more abundant the structure is, the greater the value is.

For this reason, in order to achieve the goal of recognition, we mapped the object function f(x) to another transformation space, then we got a new moment and we called it structure moment invariant:

$$\mu i =<F(f), \psi i>=\int_{\Omega} F(f(x)) \bullet \overline{\psi_i(x)} \, dx \qquad (2.2)$$

Note F(f) is function of f, and the others are defined in (2.1)

F can be linear transformation, and can be nonlinear transformation too. Especially, when F is a constant we will get (2.1).

If $g(x)= F(f(x))$, then $\nu i =< g(x), \psi i >=\int_{\Omega} g(x) \bullet \overline{\psi_i(x)} \, dx$. The complicated objects will be recognized through the existing pattern recognition methods.

According to (2.2), if f (x , y) is a limitary two-dimensional function, $F(f)=f^2$ and the basis function is a non-orthonormal basis, then moment of order (p + q) can be defined as

$$m_{pq}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^p y^q f^2(x,y)) \, dxdy \quad (p,q=0,1,2,\cdots)$$

The central moments of order (p + q) can be defined as

$$\mu_{pq}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x-x0)^p (y- y0)^q f^2 (x,y)dxdy \quad (p,q=0,1,2,\cdots)$$

Where (x0, y0) is the coordinates of the centre of mass

$$x0=\frac{m_{10}}{m_{00}}, \quad y0=\frac{m_{01}}{m_{00}}$$

To the digital picture

$$m_{pq}=\sum_x \sum_y x^p y^q f^2 (x,y)$$

$$\mu_{pq}=\sum_x\sum_y (x-x0)^p (y-y0)^q f^2(x,y)$$

Central moments μ_{pq} is invariant with respect to position. Under the following orthogonal transformation

$$\eta_{pq}=\frac{\mu_{pq}}{\mu_{00}^{r}} \quad (r=(p+q+2)/2，p+q\geq 2)$$

the moments will be invariant with respect to scale.
Similarly, we defined seven invariants from Hu's moment invariats.

3 Analysis and Results

In order to verify the result, we regard complicated manikin as the basic experimental data, and adopt one group of pictures which got from different angles for test. We obtain many pictures as testing database with the body that is raising the hand and bowing step on different angles at random, and give the serial number of every picture respectively.

There are two groups of test, and we compared the results of the test of structure moments (method 1) and Hu's moments (method 2).

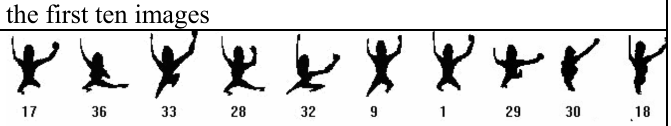
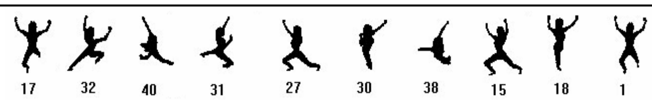
	c1	the first ten images
the first method		
the second method		

Figure 2. the first ten images indexed with the two methods about c1

Table 1. the order of images which are similar to c1 from the testing image database

the order（the second ethod）		the distance（the second method）	the distance（the first method）
1	17	0.000125348007680	0.000187478906854
2	32	0.000660715253218	0.002750807052167
3	40	0.000744803344816	0.078550511379515
4	31	0.000833070784763	0.004666607387328
5	27	0.001442010616861	0.013709215879515
6	30	0.001711294876661	0.003710746779861

7	38	0.001767158498375	0.018007635220625
8	15	0.001824070831385	0.009271637049499
9	18	0.002057676349881	0.003870371347967
10	1	0.002191317765542	0.002934754291871
the order （the first method）		the distance （the second method）	the distance （the first method）
1	17	0.000125348007680	0.000187478906854
2	36	0.005590079907031	0.000752931498535
3	33	0.006720642349892	0.001744045803518
4	28	0.002593539914291	0.001795322506055
5	32	0.000660715253218	0.002750807052167
6	9	0.009918330235619	0.002889837522023
7	1	0.002191317765542	0.002934754291871
8	29	0.004133192402461	0.003471168987230
9	30	0.001711294876661	0.003710746779861
10	18	0.002057676349881	0.003870371347967

We search for images which are similar to c1 from the testing database of images and order them according the similarity. Table 1 lists the first ten images indexed with the two methods said above. If we use the second method, image 32 lines in the 2nd place, image 40 the 3rd, image 36 the 17th, image 33 the 19th; But if we use the first method, image 32 lists in the 5th place, image 40 the 37th, image 36 the 2nd, image 33 the 3rd. Actually, image 36 and image 33 are more similar to c1 than image 32 and image 40. The 2nd method distributes images of the database in interval of [0.000125348007680，0.083966473378103] while the 1st method enlarge the interval to [0.000187478906854，0.33841268978052].

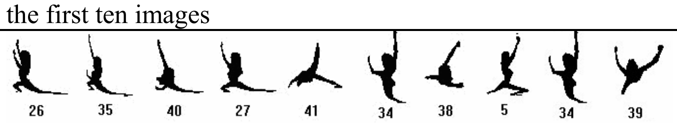
	c2	the first ten images
the first method		
the second method		

Figure 3. the first ten images indexed with the two methods about c2

Table 2. the order of images which are similar to c2 from the testing image database

the order （the second ethod）	the distance （the second method）	the distance （the first method）

1	2	0.0002502642309862	0.019653609500517
2	18	0.0003557256302715	0.020322267083941
3	20	0.0010677099722689	0.026540867092134
4	40	0.0014080399153509	0.007943660597961
5	1	0.0017195433637831	0.027716073040467
6	27	0.0017560566150470	0.008988837384352
7	33	0.0017614170171022	0.031277224419745
8	9	0.0018143206345173	0.039675097874939
9	15	0.0027032293616430	0.017966020832443
10	35	0.0028824764196008	0.012329131789070
the order （the first method）		the distance （the second method）	the distance （the first method）
1	26	0.0092484681237414	0.003983001617729
2	35	0.0028824764196008	0.012329131789070
3	40	0.0014080399153509	0.007943660597961
4	27	0.0017560566150470	0.008988837384352
5	41	0.0054291964270121	0.009501716602267
6	6	0.0155735038090230	0.012164447007566
7	38	0.0103354515234100	0.005426964517334
8	5	0.020020090301977	0.012777078763895
9	34	0.041937012576623	0.013831520185084
10	39	0.0078330039169631	0.014360517523508

And we search for images which are similar to c2 from the testing database of images and order them according the similarity. Table 2 lists the first ten images indexed with the two methods said above. If we use the second method, image 2 lines in the 1st place, image 18 the 2nd, image 26 the 28th, image 35 the 10th; But if we use the first method, image 2 lists in the 14th place, image 18 the 16th, image 26 the 1st, image 35 the 2nd. Actually, image 26 and image 35 are more similar to c2 than image 2 and image 18. The 2nd method distributes images of the database in interval of [0.0002502642309862，0.061255879295977] while the 1st method enlarge the interval to [0.003983001617729，0. 41747553628997].

	c3	the first ten images
the first method		

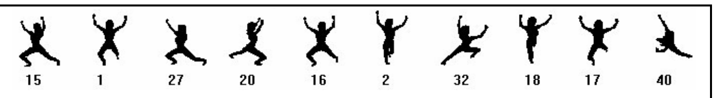
the second method		

Figure 4. the first ten images indexed with the two methods about c3

Table 3. the order of images which are similar to c3 from the testing image database

the order（the second ethod）		the distance（the second method）	the distance（the first method）
1	**15**	0.0002859202548459	0.004602570051266
2	**1**	0.0004141667254640	0.000777874103655
3	**27**	0.0004804623746547	0.010230014115013
4	**20**	0.0008599789888517	0.004495787153661
5	**16**	0.0011943735757784	0.001357933352222
6	**2**	0.0014766260741646	0.003163291060699
7	**32**	0.0017629151694908	0.000916833388346
8	**18**	0.0017998653088092	0.003030208816386
9	**17**	0.0019945557335414	0.001330589884956
10	**40**	0.0021560221947273	0.078676375926922
the order（the first method）		the distance（the second method）	the distance（the first method）
1	**8**	0.0096238306568775	0.000449517577671
2	**33**	0.0041012356529578	0.000727789290503
3	**1**	0.0004141667254640	0.000777874103655
4	**32**	0.0017629151694908	0.000916833388346
5	**22**	0.0053477297144838	0.001198510382485
6	**17**	0.0019945557335414	0.001330589884956
7	**16**	0.0011943735757784	0.001357933352222
8	**9**	0.0032875230664497	0.001649140445942
9	**7**	0.0042553411850559	0.001682104332752
10	**21**	0.0034088331458849	0.002108374212194

And we search for images which are similar to c3 from the testing database of images and order them according the similarity. Table 3 lists the first ten images indexed with the two methods said above. If we use the second method, image 15 lines in the 1st place, image 1 the 2nd, image 8 the 28th, image 33 the 13th; But if we use the first method, image 15 lists in the 20th place, image 1 the 3rd, image 8 the 1st, image 33 the 2nd. Actually, image 8 and image 33 are more similar to c3 than image

15 and image 1. The 2nd method distributes images of the database in interval of [0.0002859202548459，0.0713609637900520] while the 1st method enlarge the interval to [0.000449517577671，0.331014962711620].

The result shows that structure moment method (method 1) is more effective than Hu's moments (method 2) in identifying complicated objects.

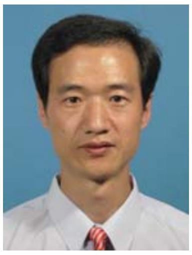
4 Conclusions and Future Work

This paper developed the moment invariants method from a new perspective. The structure moments through transforming the original density functions to the new ones are invariant with respect to translation, scale and orientation, and can describe the form of the complicated structure 2-D objects. In order to verify the method, take the example of manikin, we compared the results of the test of structure moments and Hu's moments, from which we can see that structure moments are distinctively better than Hu's moments. The focal point of work in the future is how to choose a transformation so as to improve the precision and robustness of pattern recognition.

References

1 1. R. Osada, T. Funkhouser, B. Chazells, D. Dobkin. Matching 3D Models with Shape Distributions，In: Proceedings of international conference on shape modeling and applications, genova, italy,2001,154-167
2 C.Jacobs, A.Finkelstein, and D.Salesin. Fast multiresolution image querying. SIGGRAPH, 1995,277-286.
3 YS.Kim, WY. Kim. Content-Based Trademark Retrieval System Using a Visually Salient Feature, IVC(16), No. 12-13, 24 August 1998, 931-939
4 MK Mandal, T. Aboulnasr and S. Panchanathan. Image Indexing using Moments and Wavelets, IEEE Transactions on Consumer Electronics, vol. 42, no. 3, pp. 557 - 565, August 1996
5 K. Arbter, W. E.Snyder, H.Burkhardt, and G. Hirzinger. Application of affine-invariant fourier descriptors to recognition of 3-d objects. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 7 (July), 1990,640-647.
6 M K Hu. Visual pattern recognition by moment invariants, IEEE trans. Information Theory, Vol.8,1962, 179-187
7 M. Elad, P. Milanfar, and Gene Golub, Shape from Moments - An Estimation Theory Perspective, the IEEE Trans. on Signal Processing on August 2002
8 M Teague. Image analysis via the general theory of moments. Journal Optical Society of America,1980, 70(8):920--930.
9 I. Young, J. Walker, and J. Bowie. An analysis technique for biological shape. Computer Graphics and Image Processing 25, 1974.pp 357-370.
10 B. Horn. Extended gaussian images. Proc. of the IEEE 72, 12 (December), 1984，1671-1686.
11 D. Saupe and D. V. Vranic. 3D Model Retrieval With Spherical Harmonics and Moments. In Proceedings of the DAGM 2001 (editors B. Radig and S. Florczyk), Munich, Germany,

2001，392-397
12 M. Novotni.3D Zernike Descriptors for Content Based Shape Retrieval, proceedings of The 8th ACM Symposium on Solid Modeling and Applications, June 2003
13 Robert Osada, Thomas Funkhouser, Bernard Chazells, David Dobkin,, Shape Distribution. ACM Transactions on Graphics, 21(4), pp. 807-832, October 2002
14 L.Cheng. optics : theory and development , Science Press，Beijing，China，1990



Zongmin Li, was born in 1965. He received his Ph.D. degree in Institute of Computing Technology, Chinese Academy of Sciences. His current research interests include CAGD, image processing, pattern recognition, and computer Graphics.

Kunpeng Hou, was born in 1981. He received his bachelor degree in Shandong University of Technology. He is a master candidate in University of Petroleum now. His interest area includes image processing and pattern recognition.

Liu Yujie, was born in 1971. He received his bachelor and master degree in University of Petroleum. He is a Ph.D. candidate in Institute of Computing Technology, Chinese Academy of Sciences now. His interest area includes image processing, and computer Graphics.

Luhong Diao, was born in 1978. He received his bachelor and master degree in Shandong University. He is a Ph.D. candidate in Institute of Computing Technology, Chinese Academy of Sciences now. His interest area includes CAGD, image processing, computer Graphics.

Li Hua, was born in 1957. He received his Ph.D. degree in Institute of Computing Technology, Chinese Academy of Sciences in 1989. His current research interests include CAGD, image processing, pattern recognition, and computer Graphics. He is a IEEE member