Algorithm for the Extraction of the Features of Squamous Carcinoma Cells Based on MEFMs

Hu Qi1,2
1Department of Electronics and Information Engineering, Changchun University of Science and Technology, Changchun 130022, Jilin, China
2Information Engineering Department Jilin Business and Technology College, Changchun130507, Jilin, China

Duan Jin*
Department of Electronics and Information Engineering, Changchun University of Science and Technology, Changchun 130022, Jilin, China
*Correspondence author

Wang Li Ning
College of Biological and Agricultural Engineering, Jilin University, Changchun, Jilin, China

Zhai Di
Department of Public Security of Jilin Province, Changchun, Jilin, China

Abstract
Using the image moment as the characteristics to recognize biological cells has gradually become the research hotspot of academia. The extraction of image features is to find a set of data which is as small as possible and which can represent the image. However, all the current researches cannot well balance its accuracy and robustness, especially cannot guarantee the multi-distortional invariants such as scaling, translation and rotation when choosing the features of the images. Therefore, this paper proposes an algorithm that possesses the multi-distortional invariants features for the extraction of squamous carcinoma cells, which is based on Mapping Exponential-Fourier Moments (MEFMs). At the same time, extending it to the general cell feature extraction mode and finally using the image moment to reconstruct the image. From the result of the experiment, we can see that the accuracy of the feature extraction of the target image is significantly improved, which thus ensures a high recognition efficiency.

Key words: Image Moment, Biological Cell Recognition, MEFMs, Squamous Carcinoma Cells.

1. INTRODUCTION

One of the important problems in image recognition is how to obtain an effective description of the image with a very small data set to represent the image. At the same time, the description of image should not be sensitive to various distortion of image, namely, it has the multi-distortional invariants. The moment of the image is a highly condensed image characteristic, which has the translation, scaling, rotation, and gray scale invariance, therefore, the image moment is very suitable for image recognition, target recognition and scene analysis. For the limitation of the cancer diagnosis and treatment of current biomedical research that are mostly for the specific parameters identification, the method of image moment can ingeniously do the cell recognition according to different cell characteristics. Therefore, this paper used squamous cancer cells as an example, and put forward a method that possesses squamous cancer cells with the multi-distortional invariants features which based on Mapping Exponential-Fourier Moments (MEFMs). (Chow and Kaneko,1972;Gonzalez and Wintz,1987)

Chinese-American Ming-Kui Hu first put forward the concept of moment invariants in 1961, using geometric moment for the description of the image. (Hu,1962) However, the low order geometric moment contains less image detail information, while the high order geometric moment is susceptible to noise, and it is difficult to use geometric moment to restore image. Teh and Chin evaluated geometric, Legendre moment, Zernike moment, Pseudo-Zernike moment and complex moment from noise sensitivity, information redundancy and the ability of image description, and they found that the nature of the Zernike moment was the better than any other image of moment(Teh and Chin,1988). Sheng and some other people put forward Orthogonal Fourier-Mellin Moments (OFFMs) in 1994, and proved that Orthogonal Fourier-Mellin Moments is better than Zernike...
moment in the aspect of describing image. (Sheng and Shen, 1994) Ping put forward the Chebyshev-Fourier moment, which has similar nature of Orthogonal Fourier moment. Then, he put forward a new image moment: Jacobi-Fourier Moments, and he pointed out that the change of the polynomial $p$ and $q$ can form various orthogonal polynomial, which finally can for various orthogonal moments that the above mentioned. (Ping and Sheng, 2002; Ping, 2008)

From the perspective of the normalization method of realizing scale distortion, radial function can be any function, as long as the calculation of the moments of integral proportional changes have taken place in the image, the image moment can keep scaling invariance. In 2003, Ren H.P. put forward Trigonometric-Fourier Moments (TFMs). Due to the number and location of the zero radial function represents the image sampling frequency and sampling position of the image moments, we can see the nature of the trigonometric functions Fourier moments from the zero position. The multiple zeros were distributed equally, and when the first zero point reached at or very close to the origin (image center), the contribution to the moments was the same from center to everywhere of edge, which was very suitable for describing small image. Therefore, Trigonometric-Fourier Moments has strong ability of image description with small reconstruction error and high quality. However, from the aspects of the TFMs trigonometric function computing complexity and computing speed, there are still pressures. (Ren, 2003) From the construction process of the above-mentioned image moments, we can see that there is great pressure in both the computational complexity and operational speed. Therefore, based on the above mentioned, this paper constructed MEFMs, which improved the efficiency of image feature extraction.

2. IMAGE MOMENTS WITH MULTI-DISTORTED INVARIANT FEATURES

Moments were used to describe the images because of the features of rotation, translation and scaling invariance. But Bhatia and Wolf in their paper in 1954 proved that all the polynomial about the rotating origin representing the invariant should comply with the following format:

$$V(r \cos \theta, r \sin \theta) = R_n(r) \exp(j m \theta)$$

Therefore, moments with multi-distortional invariants features were defined under the polar coordinates, while the most images in our computer were stored with rectangular coordinates as reference, which would give troubles to calculate the image moments with rotation invariance feature. When calculating the moments in the existing literature, it usually adopted the method of interpolation, which put each pixel of the image into the discrete polar coordinates, and then calculated the image moments, which directly lead to the existence of error. For this purpose, this paper represents a method that calculating the image moments with multi-distortional invariants features directly under the rectangular coordinates. (Bhatia and Wolf, 1954; Shi and Zhang, 2007)

2.1. Construction of MEFMs

In the polar coordinate system $(r, \theta)$, the function system is defined as $P_{nm} (r, \theta)$, including two parts, radial function $J_n (r)$ and angular function $\exp(j m \theta)$:

$$P_{nm} (r, \theta) = J_n (r) \exp(j m \theta)$$

To guarantee the function system $P_{nm} (r, \theta)$ is orthogonal in the interval between $0 \leq r \leq 1$, $0 \leq \theta \leq 2 \pi$, that is

$$\int_0^1 \int_0^{2 \pi} P_{nm} (r, \theta) P_{pq}^* (r, \theta) r \, dr \, d\theta = \delta_{nm} \delta_{pq}$$

Among which, $\delta_{nm}$ is the Kronecker symbol, $r = 1$ is the maximum size of the object encountered under the specific circumstances. The image function $f(r, \theta)$ in the polar coordinates system can be decomposed according to function system $P_{nm} (r, \theta)$ as follows:

$$f(r, \theta) = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} \phi_{nm} J_n(r) \exp(j m \theta)$$

Among which,

$$\phi_{nm} = \int_0^1 \int_0^{2 \pi} f(r, \theta) J_n(r) \exp(-j m \theta) r \, dr \, d\theta$$

The above formula is the general form of the orthogonal image moments with multiple distortion invariance. For this reason, the image should be normalized to the unit circle when calculating the exponential moment of it according to the definition of the exponential moment in the polar coordinate system. (Li and Xu, 2009; Wang and Lu, 2011)
Figure 1. The Integral Region of the Exponential Moment

Among which, the unit circle S1 is the integral region and the rectangle S2 is the normalized image. For this purpose,

\[
\phi_{nm} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 f(r, \theta) Q_n(r) \exp(-jm\theta) rdrd\theta
\]

\[
= \frac{1}{2\pi} \int_1^{r_1} f(r, \theta) Q_n(r) \exp(-jm\theta) dr
\]

\[
= \frac{1}{2\pi} \int_{r_2}^{r} f(r, \theta) Q_n(r) \exp(-jm\theta) dr + \frac{1}{2\pi} \int_{r_1}^{r_2} f(r, \theta) Q_n(r) \exp(-jm\theta) dr
\]

(6)

We can see from the picture, there is no pixel value in the S1-S2 region which has no value to the moments. Therefore, the integral of this part is zero. Then the exponential moment of the moment can be expressed as:

\[
\phi_{nm} = \frac{1}{2\pi} \int_0^r f(r, \theta) Q_n(r) \exp(-jm\theta) dr
\]

\[
= \frac{1}{2\pi} \int_0^r f(x, y) Q_n(r) \exp(-jm\theta) dx dy
\]

(7)

Based on the above reasoning, the approximate reconstruction image \( \hat{f}(r, \theta) \) can be obtained by summing up the exponential moments of finite images as:

\[
\hat{f}(r, \theta) \approx \sum_{n=0}^{N} \sum_{m=0}^{M} \phi_{nm} J_n(r) \exp(jm\theta)
\]

(8)

Also, we can draw the conclusion as follows:

\[
\hat{f}(x, y) \approx \sum_{n=0}^{N} \sum_{m=0}^{M} \phi_{nm} J_n(r) \exp(jm\theta)
\]

(9)

Among which, \( r = \sqrt{x^2 + y^2} \), \( \theta = \arctan \left( \frac{y}{x} \right) + k\pi, 0 \leq \theta < 2\pi \), when the pixels are in the II, III, I, IV quadrants, \( k \) equals to 0, 1, 1, 2 respectively. So far, we have derived the formula for calculating the exponential moment in the rectangular coordinate system, which replaces the traditional interpolation method, reduces the amount of computation and simplifies the algorithm.

2.2. Construction of multi-distorted invariant features

The most important feature of the squamous carcinoma cells when compared with the normal squamous cells is that the structure of the nucleus is abnormal. The size of the nucleus increases, the karyoplasm proportion is abnormal and the nucleus accounted for more than one-third of the entire cell, but the volume of the whole cell generally does not increase. Diagnostically, the change of the cytoplasm is not important, but the state of the cytoplasm must be considered when distinguishing the type and the degree of differentiation of the cancer cells. Therefore, we use an exponent—the Fourier moment as the feature to the image in recognition of the squamous carcinoma cells, which is because the Fourier moments has the best description performance and the fastest calculation speed in various image moments.
The exponential moment itself is not a distortion invariant, but it can obtain translation, the gray scale, the size and the rotation invariance. The normalization method is as follows:

Calculate the center of gravity first when calculate the image moments. For the two-dimensional image, the center of gravity of the image can be obtained by the following formula:

$$x_c = \frac{m_{01}}{m_{00}}, \quad y_c = \frac{m_{10}}{m_{00}}$$

Among which,

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y)$$

Then we carry on the coordinate transformation to shift the origin of the coordinates to the center of the image. The exponent--Fourier Moments we get by using this method in the new coordinate has the feature of translational invariance.

At the same time, due to the angle function of the exponential Fourier Moments is $e^{j\phi}$ in the new coordinate system, same phase factor $e^{j\phi}$ will be added to all the moments $E_{nm}$ after rotating $\phi$ degrees and the mold of the exponent--Fourier Moments $|E_{nm}|$ is invariable, so the exponential Fourier moments calculated in the new coordinate system also has the feature of rotation invariance. (Peng and Zhang, 2011; Bi and Qiu, 2005)

Comparing to the rotation and translational invariance, we will obtain scale invariance by adopting the following method:

Put the target in a circle. The center of this circle is the gravity center of the image, that is to say, it is the origin of the new coordinate system, and the radius of the circle is the distance from the farthermost image pixels to the origin. Then normalize this circle to a unit circle. The image moments calculated from this unit circle has the feature of scale invariance. Adopting this method, we need to find the radius of the circle. Detailed procedures are as follows:

1). Calculate the border and the gravity center of the image $(x_c, y_c)$
2). Assume $r_{max}^2 = 0$
3). Perform the following operation to each row of pixels in the image:
   Find the leftmost and rightmost pixels that are not zero in this row: $x_1, x_2$. If all the pixels are zero, then jump to the next line.
   $$\begin{align*}
   x'_i &= |x_i - x_c|, \quad i = 1, 2 \\
   x_0 &= \max(x_1, x_2) + 0.5 \\
   y_0 &= |y - y_c| + 0.5 \\
   r^2 &= x_0^2 + y_0^2
   \end{align*}$$
   If $r^2 > r_{max}^2$, then $r_{max}^2 = r^2$, continue to calculate the next line.
4). Use $r_{max}$ as the radius of the circle and as the scaling factor. The image moments obtained at this time has the feature of scale invariance. (Kass and Witkin, 1998)

3. EXPERIMENTAL PROCEDURES

Based on above the above-mentioned theoretical model for feature extraction, this paper uniformly uses 111 cell microscopic images of medical technology as the research object, which is BMP 8bit gray image and whose size is $64 \times 64$. The hardware tested that the host environment is Intel Core i5 4590 processor, 8G ram, ASUS GTX960-DC2OC-4GD5-SI graphics card and the memory is 4G. The software testing platform is MATLAB R2014a. Because the research focus of this paper is the feature extraction of squamous cell carcinoma cells and image moments construction, we don’t consider the image segmentation as the main research contents but directly calculate the image moments of the image training set of 50 segmented single squamous carcinoma cells and identify them by using the SVM method. Finally, by calculating the normalized image reconstruction error (NIRE) of the reconstructed images, this paper proves that this method has a significant effect on the accurate identification of the squamous carcinoma cells.
Since the exponential moment is mathematically equivalent to the trigonometric moment, the Normalized Image Reconstruction Error when reconstructing the fixed image can be defined as:

\[
\varepsilon^2 = \frac{\iint_{-\infty}^{\infty} \left[ f(x, y) - \hat{f}(x, y) \right]^2 dxdy}{\iint_{-\infty}^{\infty} f^2(x, y) dxdy}
\]  

(13)

The calculation results as shown:

**Figure 4. The Reconstruction Errors of the Normalized Images**

4. CONCLUSIONS

This paper introduced a new image recognition algorithm which use the MEFMs as image moment. We mainly used the exponential moments as the image recognition features. We used the exponential moment in the polar coordinates to map directly to the rectangular coordinate system, and used the method of the center edge circle to obtain the image moment with rotation, translation and scale invariance. In this recognition algorithm, the using image moment as an image feature avoids the problem of using different methods to different cells to measure the features of different images. Therefore, the recognition algorithm introduced in this passage is a
general algorithm, which can be promoted to other cell image recognition systems and lay a foundation for the future development of an automatic cell recognition system.

Acknowledgements

This work was supported by YSFP61308051, also supported by the Science and Technology development plan project of Department of Education of Jilin Province ([2015] NO. 426), China.

REFERENCES


