韓國 컴퓨터情報學會 論文誌
第11卷 第6號, 2006. 12

2006-11-6-1-1

핵형 분류를 위한 퍼지 멤버쉽 함수의 처리

엄 상 희*, 남 재 현**

Computing of the Fuzzy Membership Function for Karyotype Classification

Sang-hee, Eom*, Jae-hyun, Nam**

요 약

많은 연구자들이 자동 염색체 핵형 분류와 해석을 연구하고 있다. 현미경상의 이미지를 개개의 염색체로 자동 분류하기 위해서는 이미지 전처리, 핵형 분류기 구현 등의 세부 절차가 필요하다. 이미지 전처리에서는 개개의 염색체 분리, 잡음 제거, 특징 파라미터 추출을 진행한다. 추출된 형태학적 특징 파라미터는 동원체 지수, 상대 길이비, 상대 면적비이다. 본 논문에서는 인간 염색체 핵형 분류를 위하여 퍼지 분류기가 사용되어졌다. 추출된 형태학적 특징 파라미터가 퍼지 분류기의 입력 파라미터로 사용되었다. 우리는 개개의 염색체 그룹에 대한 최적 퍼지 분류기를 위하여 멤버쉽 함수를 선택하는 것을 연구하였다.

Abstract

Many researchers have been studied for the automatic chromosome karyotype classification and analysis. For the automatic classify the each chromosome which is the image in microscope, it is necessary to process the sub-procedure, ie, image pre-processing, implementing karyotype classifier. The image pre-processing proceeded the each chromosome separation, the noise exception and the feature parameter extraction. The extracted morphological feature parameter were the centromeric index(C.I.), the relative length ratio(R.L.), and the relative area ratio(R.A.). In this paper, the fuzzy classifier was implemented for the human chromosome karyotype classification. The extracted morphological feature parameter were used in the input parameter of fuzzy classifier. We studied about the selection of the membership function for the optimal fuzzy classifier in each chromosome groups.

▸ Keyword : chromosome, karyotype, classification, fuzzy membership function

---

• 제1저자 : 엄상희 • 교신저자 : 남재현
• 접수일 : 2006.10.19, 심사일 : 2006.12.08, 심사완료일 : 2006. 12.20.

* 동주대학 사이버경찰과 조교수, ** 신라대학교 컴퓨터정보공학부 조교수

Ⅰ. Introduction

The analysis of chromosome is to classify and recognize the chromosome karyotype. This can be used at the diagnosis of a genetics disease. The human chromosome analysis is widely used to diagnose leukemia, malignancy, radiation hazard, and mutagen dosimeter as well as various congenital anomalies such as Down's, Klinefelter's, Edward's, Patau's syndrome. The karyotype analysis that was extracted from amniotic fluid or blood is standard test for pregnant woman in the whole world.

In 1956, Tjio and Levan discovered that the number of the human chromosome is 46. In 1960, the ISCN was instituted at the international chromosome conference. The regulation defines the human chromosome nomenclature method and classification method. According to ISCN, the chromosomes of a normal human consists of 22 pairs of autosomes and the sex chromosomes(X, Y). The sex chromosome is composed of XX in female and XY in male. Each pair of the chromosomes is classified by the centromeric index(C.I.), existence or nonexistence of satellite, the relative length ratio(R.L), and the relative area ratio(R.A.), is called morphological features. The number of chromosomes and their morphological features are called karyotype, and the tabulated diagram with the sequential numbers of the chromosome's groups and its numbered classification is called karyogram[1].

In 1964, the earliest research on the automated chromosome karyotype analysis was done by Ladly[2]. Since then, enormous amount of research on it has been done. In recently, Much research on automated chromosome karyotype analysis methods has been carried out, some of which produced commercial systems. However, there still remains much study to improve the accuracy of chromosome classification and to reduce the processing time in real clinic environments.

Chromosome classification is the most widely investigated stage of the automatic chromosome analysis. Their required morphological feature parameters extracted from preprocessing procedure. Over the years, classifiers have been tested, among them: distance and statistical, nearest neighbor, neural network algorithm, and fuzzy logic[3~5]. Particularly, The neural network method among an artificial intelligence method have been used much for the enhance of the recognition. But, they have some flaws, the learning time is many consume and many feature parameters are used. The fuzzy logic had been used to obtain successful results in pattern recognition. This can apply also to controls, economics, psychology, marketing, biology, politics, to any large complex system. The fuzzy logic have made membership function using significance and distribution of parameter, and the classification is processed by membership function. Therefore, this has some advantages that the time is shortly required to made the structure of the membership function, and the addition of feature parameter is easy[6].

In this paper, we proposed chromosome karyotype classifier using fuzzy logic, and performed research to select a suitable membership function for the classifier. The membership function in fuzzy classifier was constructed by standard deviation of morphological feature parameters that published by an international system for human cytogenetic nomenclature(ISCN) and human cytogenetics(HC) [7,8]. The feature parameter to be used at the experiment was extracted from human chromosome in clinically normal state.

II. Preprocessing for human karyotype classification

1. The structure of karyotype classification system

The framing and analysis of the chromosome karyogram needs experimental cytogenetic knowledge. It requires various processing works such as cell culture, s taining, slide production, suitable chromosome group choice, photography, and the isolation of individual chromosomes by scissors, etc. Therefore, for the purpose of reduce the burden of these processing works, much research has been conducted on the automated chromosome karyotype classification method, using computer.

In this paper, the structure of pattern classifier for Human karyotype classification is given Figure 1. The cell images that acquired from amniotic fluid or blood of pregnant woman, were obtained using CCD camera connected to a light microscope. We extracted each chromosome image from stored image files, obtained histogram, eliminated its background image with the lowpass filter, and made image sub-files and stored them in the PC system. In a previous works, global features like the histogram of gray levels or the two-dimensional(2D) Fourier transform components have been used. But, in this study, we have experimented with three type of chromosome morphological features : the centromeric index(C.I.), the relative length ratio(R.L.), and the relative area ratio(R.A.).

The preprocessing is applied the images for input parameters of classifier using thinning and MAT. But. the chromosome have been the irregular nonlinear structure morphologically, the un-uniformity in the cell culture procedure, the inaccuracy by the 2D image detection in the 3D image, and the separation of the chromosome image which lies on upon another. The extraction of identical feature parameter is not easy though it has same chromosome number. The extraction of a feature parameter which can classify the chromosome accurately is very important.

Therefore, in this paper, the appropriate feature parameter was extracted by image reconstruction method that proposed at former time[9]. The extraction parameter is used to input parameter in the fuzzy classifier.

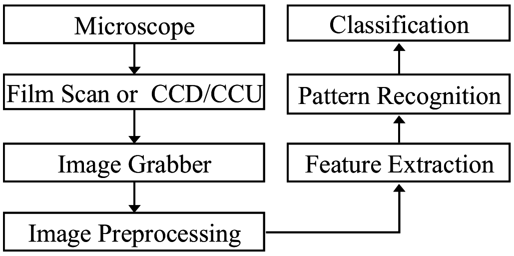


그림 1. 핵형분류시스템의 블록 다이어그램
Fig. 1 The block diagram of karyotype classification system

2. The chromosome group and the feature parameters

The nucleus of all somatic cells normally contains 23 chromosome pairs, Further that subdivided into 22 autosomal chromosome pairs and a pair of sex chromosomes(an X and another X in females, and an X and a Y chromosome in males). The chromosome groups are classified using a location of centromere, the existence of the satellite, and the density and the feature of stain band by the cytogenetic expert. The table 1 shows the chromosome groups including each chromosome number.

표 1. 염색체 그룹과 번호
Table 1. The chromosome group and number

Group	Number	Group	Number
A	1~3	E	16~18
B	4~5	F	19~20
C	6~12, X	G	21~22, Y
D	13~15		

We obtained the medial axis of each chromosome with thinning method in order to detect the chromosome centromere. One of the best way to find centromere proves to calculate the width of each chromosome. We measured the width of each pixel on medial axis in the direction of 16 radiation angles($2\pi/16$ radians). We found the shortest width and determined the centromere of the chromosome. A chromosome is divided into two arms by the centromere. The centromeric index(C.I.) is defined as Table 2.

According to the chromosome morphological structure, the length and the area of each chromosome are unique. We can classify each chromosome by comparison of each chromosome length and area. The relative length ratio(R.L.) is defined as a percent of the total autosomal length or mean length of each chromosome. Each chromosome area is obtained by counting the number of pixels in each chromosomes extracted from original images. The relative area ratio(R.A.) is defined as Table 2.

표 2. 염색체 특징 파라미터 계산
Table 2. The calculation method of chromosome feature parameter

Feature parameter Calculation method
centromeric index ; C.I.
$C.I. = \dfrac{length\ of\ short\ arm}{total\ chromosome\ length} = \dfrac{\ell_s}{\ell_c} \le 0.5$
relative length ratio ; R.L.
$\ell_t = total\ chromosome\ length = \sum_{c=1}^{44} \ell_c$ $R.L. = \dfrac{each\ chromosome\ length}{total\ chromosome\ length} = \dfrac{\ell_c}{\ell_t}$
relative area ratio ; R.A.
$S_t = all\ chromosome\ area = \sum_{c=1}^{44} S_c$ $R.A. = \dfrac{chromosome\ area}{all\ chromosome\ area} = \dfrac{S_c}{S_t}$

III. Implementation of fuzzy classifier

1. The structure of fuzzy classifier

In this study, Figure 2 is show the structure of fuzzy classifier that is implemented by fuzzy theory. The standard parameter is used from the ISCN and the HC. The membership function generator that is generated the 3 membership function such as triangular type, trapezoidal type, and gaussian type is programmed.

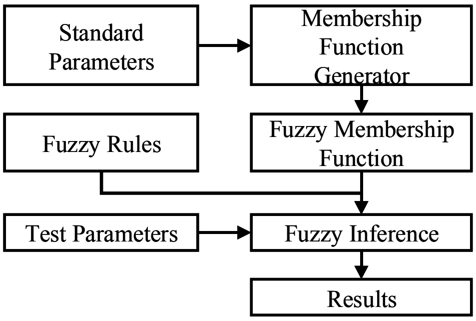


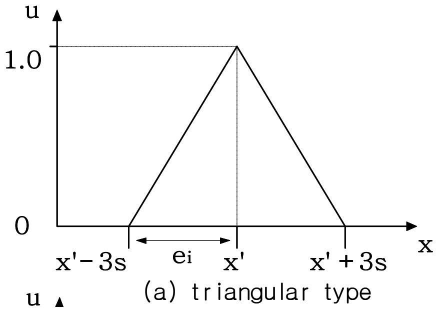
그림 2. 퍼지 분류기 구조
Fig. 2 The structure of fuzzy classifier

The feature parameter is extracted from chromosome is used the examination input parameter of fuzzy classifier. The fuzzy rule in fuzzy inference part is combined all input parameter by the Min operation which is proposed by Mamdani.

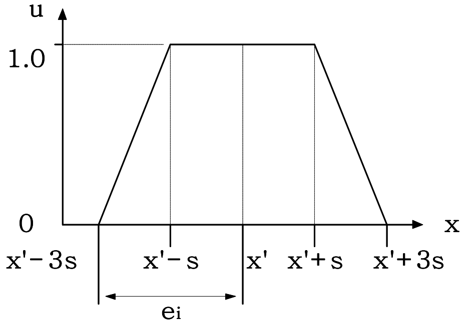
2. Computing the membership function

The membership function by fuzzification has to define for reproduction of chromosome karyotype classification.

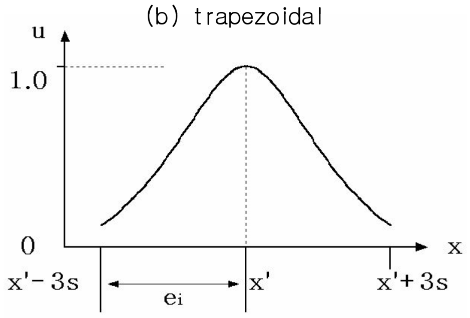
In this study, we used the 3 membership function such as triangular type, trapezoidal type, and gaussian type by the membership function generator. These membership functions are easy to program and to calculate of the defuzzification. The figure 3, ei which is the variable width of membership function is show limitation of maximum and minimum about the average value for pattern classification.



(a) triangular type



(b) trapezoidal



(c) gaussian type

그림 3. 퍼지 변수의 멤버쉽 함수
Fig. 3 The membership function of fuzzy variable

x′ is an average value of variable and $\overline{X}'$ is an average value of input parameter population. s is standard deviation and u is the result of input x. The x′, s, ei parameters which generate membership function are equation (1).

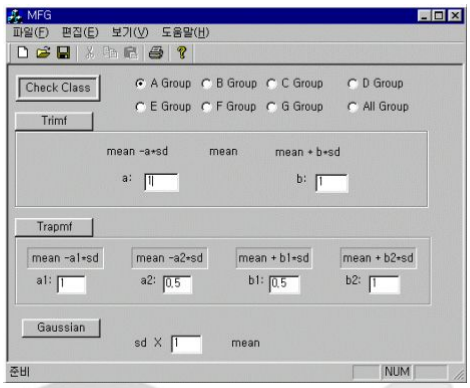
$$x' = \overline{x}' + k\,\frac{\sum(f\,z)}{N} \qquad \cdots\cdots (1)$$

$$s = \sqrt{\frac{k^2}{N-1}\left[\sum(f\,z^2) - \frac{(\sum f\,z)^2}{N}\right]}$$

$$e_i = x' \pm 3s$$

where, N is the number of measure, $\overline{X}'$ is the pre-average value, k is the range of class, f is the number of measure in class and z is the variation of class from the pre-average value.

In this paper, we programmed the membership function generator for calculating membership function about variable input data easily. The generator is easy to append and to change input data. It generates three types of membership functions. It also generates the proper membership function of each chromosome group. The membership function in fuzzy classifier was constructed by standard deviation of morphological feature parameters that published by an international system for human cytogenetic nomenclature(ISCN) and human cytogenetics(HC). The figure 4 shows the membership function generator, which is programmed by VC++.



● A Group ○ B Group ○ C Group ○ D Group
○ E Group ○ F Group ○ G Group ○ All Group

mean -a*sd mean mean + b*sd
a: 1 b: 1

mean -a1*sd a1: 1
mean -a2*sd a2: 0.5
mean + b1*sd b1: 0.5
mean + b2*sd b2: 1

sd X 1 mean

그림 4. 멤버쉽 함수 발생기
Fig. 4 A membership function generator

3. The fuzzy classification rule and inference

The fuzzy classification rule of fuzzy pattern classifier represents multi-input single-output. The fuzzy inference shows equation (2). Where, A_i is fuzzy set of input X, B_i is fuzzy set of input Y, C_i is fuzzy set of output Z. x_o, and y_o are the each elements of X and Y

$$\begin{array}{llll} (rule) & A_i \text{ and } B_i & then\ C_i & i=1,2,\ldots,n \cdots\cdots (2) \\ (fact) & x_o \quad y_o & & \\ (conclusion) & C' & & \end{array}$$

In this paper, we use the MIN operation inference among the min operation, product operation, linear operation. The equation (3) is show membership function C'.

$$\mu C'(z) = \vee\{[\mu A_i(x_o) \wedge \mu B_i(y_o)] \wedge \mu C_i(z)\} \cdots\cdots\cdots (3)$$

The state of defuzzification is procedure to calculate the only one output Z_o that used fuzzy set C' that is calculated by result of inference.

In general, the defuzzification uses the maximum method, the maximum average method and the center of gravity method. In this paper, however, output Z_o used the center of gravity method as equation (4).

$$z_o = \sum_i \frac{z_i \cdot \mu C'(z_i)}{\mu C'(z_i)} \quad \cdots\cdots \quad (4)$$

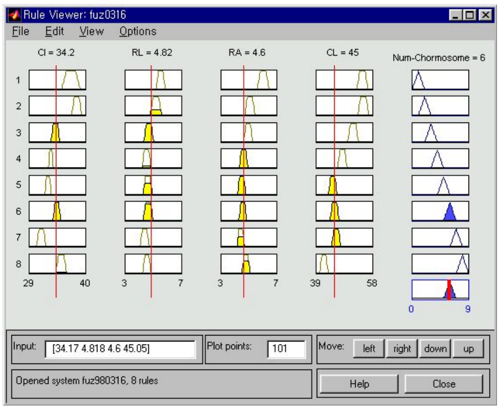
Ⅳ. Experimental Result

1. The configuration of human karyotype classification

We acquired ten chromosome images using CCD camera connected to a light microscope, which provided to be clinically normal state in the chromosome morphological structure. These chromosome images were stored as the image files in the PC system. We extracted three chromosome morphological feature parameters(C.I., R.L. R.A.) by preprocessing ten chromosome images. The feature parameters are used to classify the chromosome in each chromosome groups. We performed research to select a suitable membership function to configure the part of fuzzy inference in the fuzzy classifier. A membership function which is used here are triangular type, trapezoidal type, and gaussian type.

2. The experimental results

we configured three membership functions to select the correct membership function from standard parameter in ISCN and HC. The configured membership function is applied the each fuzzy inference and the each chromosome is classified by fuzzy inference. The figure 5 shows the input/output relation used trapezoidal membership function about chromosome group C.



Input: [34.17 4.818 4.6 45.05]

Plot points: 101

Opened system fuz980316, 8 rules

그림 5. 사다리꼴 멤버쉽 함수의 입출력 관계
Fig. 5 The I/O relation used trapezoidal membership function(group C)

Where, 'Num-chromosome=6', left above side shows the chromosome 6 in group C, this can estimate the chromosome 11. The figure 6 shows the error rate of classified chromosome group about 3 membership functions. Where, In chromosome group A, the 2 membership functions except for gaussian type have no error rate. Rest of groups, the triangular type and trapezoidal type have a few error rate. The gaussian type has high error rate.

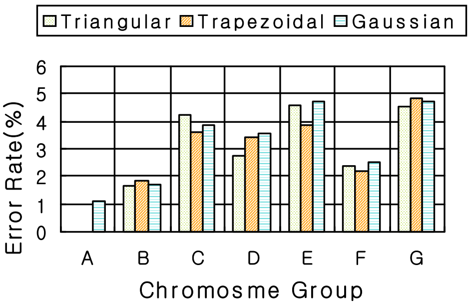


그림 6. 3가지 멤버쉽 함수에 따른 분류 결과
Fig. 6 The classification result according to three type membership function

As a results, we can get more excellent result if the classifier which has membership function of lower error rate is configured. Therefore, The table 3 wha t we configured shows the selected membership functions in the fuzzy inference part. we classified the chromosome by adapted fuzzy classifier for each chromosome group, which that is used the optimally selected membership function. The classification results have 2.66% error rate that were not find 6 chromosome among 230 chromosomes which is 10 persons.

표 3. 멤버쉽 함수의 선택 결과
Table 3. The selected results of membership functions

Membership Function / Group	Triangular	Trapezoidal
Group A	○	
Group B	○	
Group C		○
Group D	○	
Group E		○
Group F		○
Group G	○	

V. Conclusion

The fuzzy logic had been widely used to obtain successful results in pattern recognition, controls, economics, and any large complex systems. The fuzzy logic have made membership function using significance and distribution of parameter, and the classification is processed by membership function. The fuzzy classifier presents various results according to selection any membership functions.

In this paper, we processed research to choose a membership function, that are apply to optimal classifier in human karyotype classifier using fuzzy logic. we programmed the membership function generator for calculating membership function about variable input data easily. The generator is easy to append and to change input data. Three membership functions are experimented to select the correct membership function of each chromosome group from standard morphological feature parameter.

As a results, we improved that the time is shortly required to made the structure of the membership function, and the addition of feature parameter is easy. we also can get more excellent result that classified the chromosome by adapted fuzzy classifier for each chromosome group, which that is used the optimally selected membership function.

참고문헌

[1] J. R. Gosden, Chromosome analysis protocols, Humana press, 1994.

[2] Robert S. Ledly, "High-speed automatic analysis of biomedical picture", Science, vol. 146, pp. 216-223, 1964.

[3] J. Piper and E. Granum, "On fully automatic Feature measurement for banded chromosome classification", Cytometry, Vol. 10, pp. 242-255, 1989.

[4] E. Granum and M. G. Thomason, "Automatically inferred markov network models for classification of chromosomal band pattern structures", Cytometry, Vol. 11, pp. 26-39, 1990.

[5] B. Lerner, M. Levinstein, B. Rosenberg, H. Guterman, I. Dinstein, and Y. Romem, "Feature selection and chromosome classification using a multilayer perceptron neural network", IEEE International Conference on Neural Networks, Vol. II, pp. 3540-3545, Jun. 28-Jul. 2, 1994.

[6] Y. H. Chang, K. S. Lee, H. H. Chong, S. H. Eom, O. H. Choi, K. R. Jun, "Morphological feature parameter extraction from the chromosome image using reconstruction algorithm", Journal of Biomedical Engineering Research ; 1996 ; 17 ; 545-552.

[7] Y. H. Chang, K. S. Lee, K. R. Jun, "Reconstructional Morphological feature parameter extraction from the chromosome image using reconstruction algorithm", Journal of Biomedical Engineering Research ; 1996 ; 17 ; 545-552.

[8] S. H. Eom, Y. H. Chang and K. R. Jun, "Comparison of Karyotype Classification Algorithm using Fuzzy Inference and Neural Network", The 4th Asia-Pacific Conference on Medical & Biological Engineering, Vol. 2, p.345, 1999.

[9] ISCN, An international system for human cytogenetic nomenclature, Karger, 1985.

[10] D. E. Rooney, and B. H. Czepulkowski, Human cytogenetics essential data, John Wiley & Sons, 1994.

저 자 소 개

엄 상 희

1993년 동아대학교 전기공학과 (공학사)
1995년 동아대학교 전기공학과 석사
2000년 부산대학교 전자공학과 박사
2000년 3월 ~ 현재 동주대학 사이버경찰과 조교수
※관심분야 : 영상처리, 의공학, 인공지능, 패턴인식

남 재 현

1989년 부산대학교 컴퓨터공학과 (공학사)
1992년 부산대학교 컴퓨터공학과 석사
2002년 부산대학교 컴퓨터공학과 박사
1993년 3월 ~2002년 10월 동주대학 네트워크전자계열 조교수
2002년 11월 ~현재 신라대학교 컴퓨터정보공학부 조교수
※관심분야 : 컴퓨터네트워크, VoIP, 패턴인식