

염색체 영상의 재구성에 의한 형태학적 특징 파라메타 추출

장용훈·이권순·정형환·엄상희*·최옥환**·전계록***

= Abstract =

Morphological Feature Parameter Extraction from the Chromosome Image Using Reconstruction Algorithm

Y.H. Chang, K.S. Lee, H.H. Chong, S.H. Eom*, O.W. Choi**, K.R. Jun***

Researches on chromosome are very significant in cytogenetics since a gene of the chromosome controls revelation of the inheritance plasma. The human chromosome analysis is widely used to diagnose genetic disease and various congenital anomalies. Many researches on automated chromosome karyotype analysis has been carried out, some of which produced commercial systems. However, there still remains much room for improving the accuracy of chromosome classification.

In this paper, we propose an algorithm for reconstruction of the chromosome image to improve the chromosome classification accuracy. Morphological feature parameters are extracted from the reconstructed chromosome images. The reconstruction method from chromosome image is the 32 direction line algorithm. We extract three morphological feature parameters ; centromeric index(C.I.), relative length ratio(R.L.), and relative area ratio(R.A.), by preprocessing ten human chromosome images. The experimental results show that proposed algorithm is better than that of other researchers' comparing by feature parameter errors.

Key words : Chromosome, Morphological feature, Reconstruction, 32 direction line

서 론

인간의 염색체 분석은 임상 의학 및 세포 유전학(cytogenetics) 분야에서 여러 가지 선천성 질환(다운, 클라인펠터, 에드워드, 파타우 증후군 등)과 백혈병, 악성 종양, 방사선 장애, 변위유발소 선량법(mutagen dosimeter)의 연구 그리고 생물학적 연구 등을 위하여 매우 중요하고 필요한 연구이다.

염색체 수는 1956년 Tijo와 Levan에 의해 46개로 밝혀졌고, 그 후 1960년 국제 염색체 회의에서 염색체의 명명법과 분류법에 관하여 ISCN(An International System

for Human Nomenclature)이라는 규약이 설정되었다. 이 규약에 의하면 정상인의 염색체 수는 22쌍의 상염색체(autosome)와 1쌍의 성염색체(sex chromosomes)로 구분되며, 여성의 경우는 XX이고 남성은 XY인 성염색체로 구성된다. 염색체의 분류는 형태학적 특징인 동원체(centromere)의 위치, 부수체(satellite)의 유무 및 상대 길이비 등에 의하여 분류되며, 염색체의 군(group)과 염색체를 번호순으로 나열한 도형을 핵형 분류도(karyogram)라고 한다[1].

염색체 핵형 분석의 자동화에 관한 연구는 1964년 Ledly에 의해 수행된 이후[2], 지난 30여년 간 많은 학자들이

동아대학교 전기공학과

Dept. of Electrical Eng., Dong-A Univ.

* 부산대학교 전자공학과

* Dept. of Electronics, Pusan National Univ.

** 부산대학교 병원 산부인과

** Dept. of O.B. & G.Y., Pusan National Univ. Hospital

*** 부산대학교 병원 의공학과

*** Dept. of Medical Eng., Pusan National Univ. Hospital

통신저자 : 장용훈, (607-714) 부산광역시 사하구 하단동 840번지, Tel.(051)200-7739

의하여 수행되고 있으며, 근래에는 상용화된 시스템이 발표되어 사용되고 있지만 분류의 정확도가 낮아 염색체 핵형 분류의 정확도를 개선하기 위한 연구가 계속 진행되고 있다[3],[4].

컴퓨터를 이용한 염색체의 자동 분류는 염색체의 형태학적인 특징 파라메타(morphological feature parameters)를 추출하고 이를 패턴 분류기에 입력하여 염색체의 군(group)과 각각의 염색체를 분류하는 방법을 사용하고 있다. 일반적으로 사용되는 염색체의 형태학적인 특징 파라메타는 동원체의 위치 검출, 동원체 지수(centromeric index:C.I.), 상대 길이비(relative length:R.L.), 상대 면적비(relative area:R.A.) 등이 있다. 그러나 세포 배양 과정에서의 불균일, 각 염색체 형태구조의 비선형, 3차원적인 영상의 2차원적인 영상검출, 그리고 겹쳐진 염색체의 분리 등으로 인하여 동일한 번호의 염색체에서도 일정한 특징 파라메타를 추출하기가 용이하지 않다[5],[6]. 따라서 염색체의 형태구조학적인 특징 파라메타를 추출하기 위하여 윤곽선[3], MAT (medial axis transform) [5],[6],[7] 및 밴드 패턴 정보[8] 등을 이용하여 특징 파라메타를 추출하는 연구들이 수행되었다. 이러한 연구들은 개개의 염색체의 중앙 축을 검출한 후, 중앙 축의 각 화소에서 수직선을 그어 염색체 폭을 결정하였다. 그러나 염색체는 형태학적으로 굽어진 상태가 많기 때문에 굽어진 정도에 따라 추출된 특징 파라메타들은 많은 오차를 수반한다.

본 연구에서는 중앙 축의 각 화소에 등 각도로 32 방향의 선을 연장하여 선분을 구하고 이들 중 가장 짧은 선분들로 염색체의 폭을 결정한 후 이를 이용하여 염색체를 재구성하는 알고리즘을 제안하였다. 그리고 제안된 알고리즘을 임상적인 관점에서 정상인으로 판명된 10명 즉, 460개의 염색체에 적용하여 염색체의 특징 파라메타인 C.I., R.L. 및 R.A.를 구하였다. 이러한 특징 파라메타 값들을 ISCN 및 HC에서 발표한 특징 파라메타 값들과 비교하였으며 아울러 기존의 기법 즉, 중앙 축에 수직선을 연장하여 염색체의 폭을 구한 후 특징 파라메타를 구하는 기법(Piper 등에 의해 수행된 MAT기법)을 사용하여 구한 특징 파라메타 값들과 비교·검토하였다[5],[6],[7]. 그리고 ISCN 및 HC에서 발표한 특징 파라메타 값들을 기준치로 설정하여 본 연구에서 제시한 알고리즘에 의한 특징 파라메타 값 및 기존의 기법을 적용하여 구한 특징 파라메타 값 사이에 발생하는 오차율을 비교·검토하였다[9],[10]. 검토 결과 본 연구에서 제시한 알고리즘에 의해 구하여진 특징 파라메타들이 비교적 오차가 적으며 안정적임을 알 수 있었다.

염색체의 형태학적 특징 파라메타

1. 동원체 지수(centromeric index : C.I.)

염색체의 특징 파라메타들 중 동원체 지수를 구하기 위하여 먼저 동원체의 위치를 검출한 후, 동원체를 중심으로 염색체의 장완(long arm)과 단완(short arm)의 끝 좌표를 구하고 두 좌표간의 유클리드 거리(euclidean distance)를 계산하여 계산된 길이를 개개의 염색체 길이로 정하였다. 이렇게 구한 염색체의 길이를 이용하여 식 (1)과 같이 C.I.를 계산할 수 있다.

$$C.I. = \frac{\text{염색체 단완의 길이}}{\text{염색체의 길이}} = \frac{\ell_s}{\ell_c} \leq 0.5 \quad (1)$$

여기서, ℓ_s 는 염색체 단완의 길이, ℓ_c 는 염색체의 길이를 나타낸다.

2. 상대 길이비(relative length ratio : R.L.)

염색체는 형태학적으로 서로 다른 길이를 가지고 있으므로 각 염색체의 길이를 비교하여 염색체를 분류할 수 있기 때문에 R.L.을 염색체의 분류 파라메타로 사용할 수 있다. 하나의 세포에서 추출된 46개의 염색체 길이를 구한 후, 44개의 상염색체들의 길이를 합한 총 길이로 각 염색체의 길이를 나누면 식 (2)와 같이 각 염색체의 R.L.을 구할 수 있다.

$$\ell_c = \text{상염색체의 총길이} = \sum_{c=1}^{44} \ell_c \quad (2)$$

$$R.L. = \frac{\text{염색체의 길이}}{\text{상염색체의 총길이}} = \frac{\ell_c}{\ell_c}$$

여기서, ℓ_c 는 44개의 상염색체의 총 길이를 나타내며, R.L.은 각 염색체의 상대 길이비를 나타낸다.

3. 상대 면적비(relative area ratio : R.A.)

각 염색체의 면적도 서로 다른 값을 가지므로, 염색체의 면적을 분류 파라메타로 사용할 수 있다. 염색체 하나의 면적은 개개의 염색체 영상 정보에서 화소 수를 계산(counting)하여 구할 수 있다. 따라서 식 (3)과 같이 염색체의 R.A.를 구할 수 있다.

$$S_c = \text{상염색체의 총면적} = \sum_{c=1}^{44} S_c \quad (3)$$

$$R.A. = \frac{\text{각염색체의 면적}}{\text{상염색체의 총면적}} = \frac{S_c}{S_c}$$

여기서, S_c 는 각 염색체의 면적, S_c 는 상염색체의 총면적,



그림 1. 염색체 필름을 스캐너로 획득한 염색체 영상
Fig. 1. The chromosome image detected from the chromosome film using scanner



그림 2. 문턱치 처리를 수행한 염색체 영상
Fig. 2. The chromosome image after threshold processing

R.A.는 염색체의 상대 면적비를 나타낸다.

염색체 영상의 전처리

1. 영상 획득 및 필터링

임상적인 관점에서 정상인으로 판명된 10명 즉, 460개의 염색체를 현미경상에서 촬영한 후 염색체 필름을 스캐너(AVR8800)를 사용하여 염색체 영상(800×600)을 획득하고 컴퓨터(IBM-PC Pentium-100)에 염색체 영상 파일로 저장하였다.

컴퓨터에 저장된 염색체 영상은 그림 1에서 나타난 것과 같이 반전된 농도치(negative value)의 정보를 가지므로, 전체 염색체 영상 정보의 농도치를 반전 필터링(inverse filtering)을 하였다.

염색체 영상의 배경과 정보를 분리하기 위해 반전된 염색체 영상 정보의 히스토그램(histogram)을 구하여 문턱치(threshold)처리를 수행하면 그림 2와 같은 영상을 구할 수 있다.

2. 중앙축에 수직선을 연장한 특징 파라메타 추출

Piper 등에 의해 수행되어진 염색체 추출에 관한 연구는 염색체 영상을 이치화 하고 세선화 알고리즘(thinning algorithm)을 이용하여 염색체의 중앙축을 구한 후 중앙축의 각 화소에 수직이 되는 선분들을 연장하여 염색체의 폭을 구하는 기법을 제시하였다. 그리고 이러한 방법에 의해 구한 염색체의 폭들 중 가장 짧은 폭을 염색체의 동원체(centromere)로 설정하여 염색체의 형태학적 특징 파라메타(C.I, R.L, R.A)를 구하였다. Piper 등이 제안한 기존의 기법에 의한 염색체의 형태학적 특징 파라메타를 구하는 수행 과정과 결과를 그림 3에 나타내었다. 그림 3(a)는 이치화한 염색체 영상, 그림 3(b)는 세선화 기법

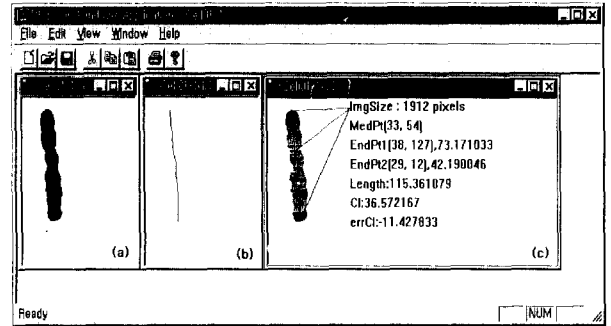


그림 3. 특징 파라메타 추출 과정
(a)이치화된 영상 (b)세선화 기법에 의한 결과 (c) 특징 파라메타 추출 결과

Fig. 3. The processing of feature parameters extraction (a)binary image (b)result image by thinning method (c)result image of feature parameter extraction

을 수행한 결과, 그리고 그림 3(c)는 C.I., R.L. 및 R.A.를 구한 결과를 나타내었다.

이와 같이 개개의 염색체 영상을 추출하는 방법은 상당히 많은 시간이 소요되며 지속 적인 운용자의 개입이 필요하다. 따라서 본 연구에서는 염색체 영상을 23개씩 두화면으로 나누어 46개의 염색체 영상의 이치화 처리와 특징 파라메타 추출을 연속적으로 수행하도록 하였다.

3. 염색체의 재구성 기법

염색체들은 형태학적인 구조가 비선형적으로 굽어진 상태로 존재하므로 특징 파라메타의 추출이 까다로우며, 동일한 번호의 염색체에서 구한 값들도 상당한 오차를 수반하게 된다. 그러나 본 연구에서 제안한 염색체 영상 재구성 방법을 적용하면 염색체들이 거의 선형적인 구조에 가까운 형태로 구성할 수 있다. 따라서 상대적인 길이를 사

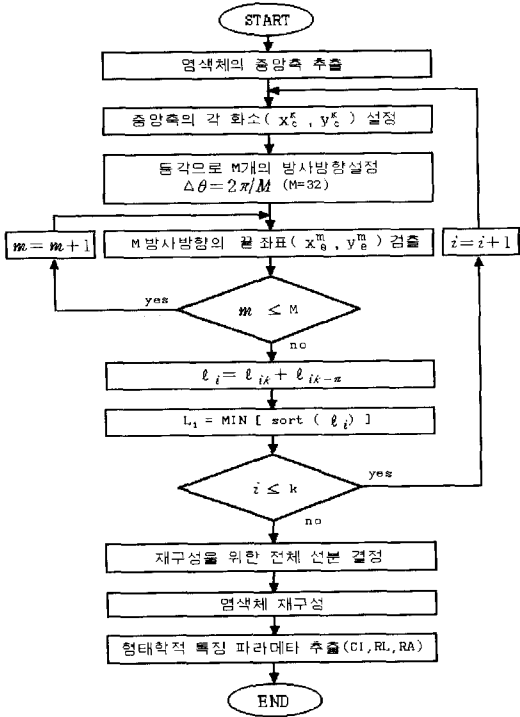


그림 4. 염색체 재구성을 위한 알고리즘의 순서도.
 Fig. 4. The flowchart of the algorithm for chromosome image reconstruction

용하는 특징 파라메타(동원체 지수, 상대 길이비)들은 더욱 정확한 값으로 추출할 수 있다. 또한 방사 방향과 동일한 염색체 농도 정보를 추출할 수 있으므로, 염색체의 농도에 관한 정보도 특징 파라메타로 사용할 수 있는 장점이 있다. 그러나, 재구성된 염색체 영상은 면적의 변화가 발생하여 오차가 다소 증가될 수 있지만, 상대적인 면적비를 계산하므로 전체 오차율은 무시할 수 있다.

1) 재구성을 위한 선분의 결정

염색체 영상을 재구성하기 위하여 각 염색체의 폭을 32방사 알고리즘에 의해 가장 짧은 폭을 구하여야만 한다. 개개의 염색체 선분들을 구하는 방법은 다음과 같다. 먼저 Zhang(11)등이 제안한 세선화 알고리즘을 이용하여 염색체의 중앙 축을 추출하고, 다음으로 중앙 축의 각 화소에 32 방향으로 선을 연장하여 염색체 영역의 가장 자리까지 32 방향의 누적된 길이를 구한다. 그 후 32 개의 길이들을 비교하여 이들 중에서 가장 짧은 선분의 좌표와 중앙축의 원점과의 유클리드 거리를 계산하고, 계산치와 π 라디안만큼 떨어진 선분의 합을 구하여 염색체를 재구성하기 위한 선분으로 결정하였다. 또한 이 선분에 있는 각 화소의 농도 정보를 합하여 선분에 있는 총화소의 수로 나누어 평균 농도를 구하였다. 그리고 이 값들을 이용하여 염색체 영상을 재구성하였다. 염색체 영상의 재

구성을 위한 32 방사 알고리즘은 김명남[12]등이 제안한 알고리즘을 활용하였다.

이에 대한 플로우차트는 그림 4와 같으며, 알고리즘의 수행 과정은 다음과 같다.

(a) 세선화 알고리즘을 수행하여 염색체의 중앙 축을 구한다. 이때 염색체 양끝에서 잘린 부분(양끝의 약 10%씩)은 맨 끝점의 좌표와 다음 점의 좌표와의 각도를 이용하여 연장한다.

(b) 설정된 염색체 중앙 축의 화소(x_c^k, y_c^k)를 원점으로 선택한다.

(c) 원점을 (x_c^k, y_c^k)로 하여 Δθ=2π/M의 등 각도로 M개의 길이를 계산한다. 여기서 k는 중앙 축의 화소에 대한 알고리즘의 반복 수행 횟수이며, M은 32이다.

(d) 길이 성분 i에 대하여 식(4)를 사용한 문턱치 에지 검출을 행하여 H_edg값이 문턱치 H_th보다 크면 길이 성분 i에 대한 수행을 종료하고 이의 좌표 (x_e^m, y_e^m)을 저장한다.

$$H_{edg}(r^i) = \frac{g(r^i+1) + g(r^i) - g(r^i-1) - g(r^i-2)}{2} \tag{4}$$

여기서 r^i는 길이 성분 i에 대한 방사 방향의 증분을 나타내며, g(r)는 염색체 각 화소의 농도 정보치를 나타낸다.

(e) M개의 방향에 대해 (d)를 반복 수행한다.

(f) (e)가 종료되면 식(5)와 같이 원점의 좌표 (x_c^k, y_c^k)와 저장된 좌표 (x_e^m, y_e^m)의 유클리드 거리를 계산한다.

$$l_{ik} = \sqrt{(x_e^m - x_c^k)^2 + (y_e^m - y_c^k)^2} \tag{5}$$

(g) (f)의 수행이 완료되면 식(6)과 같이 l_ik와 π간격만큼 떨어진 l_ik-π의 길이를 합하여 재구성에 대한 길이 l_i를 구한다.

$$l_i = l_{ik} + l_{ik-\pi} \tag{6}$$

(h) 길이 l_i중에서 가장 짧은 길이 L_i를 식(7)과 같이 구한다.

$$L_i = \text{Min} [\text{Sort}(l_i)] \tag{7}$$

(i) 중앙축의 총화소 수 k개를 만족할 때까지 (d)에서 (h)까지 반복 수행한다.

이상과 같이 수행하면 염색체를 재구성하기 위한 전체 선분이 결정된다.

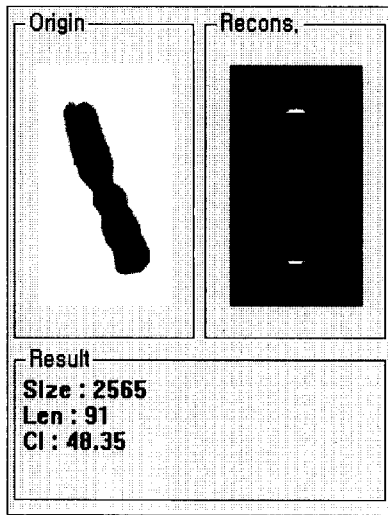


그림 5. 원염색체 영상과 재구성된 염색체 영상
Fig. 5. The original and reconstructed chromosome image

2 재구성된 염색체 영상의 특징 파라메타 추출

염색체 영상의 재구성을 위하여 3.1절에서 설명한 알고리즘을 먼저 수행하여 염색체 재구성을 위한 선분을 결정한다. 결정된 선분 상에 있는 각 화소의 농도 정보를 합하고, 그 선분상의 총화소 수로 나누면 평균 농도 정보를 구할 수 있으며, 이 농도 정보를 사용하여 염색체를 재구성하였다.

재구성된 염색체 영상에서 양끝의 약 10%씩을 제외하고, 나머지 부분의 폭을 정렬(sort)하여 가장 짧은 폭을 동원체의 위치로 설정하며, 이를 중심으로 염색체의 장관, 단완의 길이 및 염색체의 길이를 구한다. 동원체의 위치와 염색체의 길이를 이용하면 염색체의 형태학적 특징 파라메타인 C.I.와 R.L.을 구할 수 있다. 또한 재구성된 염색체 영상의 총화소 수를 계수하여 염색체의 면적비인 R.A.를 구한다. 원염색체 영상과 재구성된 염색체의 영상 및 특징 파라메타는 그림 5와 같이 된다. 이러한 방법을 임상적인 관점에서 정상인으로 판명된 10명 즉, 460개의 염색체에 적용하여 염색체의 중앙 축에 수직선을 사용한 방법과 재구성 알고리즘에 의해 추출한 특징 파라메타의 결과를 비교·검토하였다. 추출한 염색체의 특징 파라메타의 정확도를 알기 위하여 전체 평균값을 구하였으며, 동원체 지수와 상대 길이비는 1986년 ISCN에서 발표한 자료와 비교하였고, 상대 면적비는 HC(Human Cytogenetics)의 데이터와 비교하였다.

실험결과 및 고찰

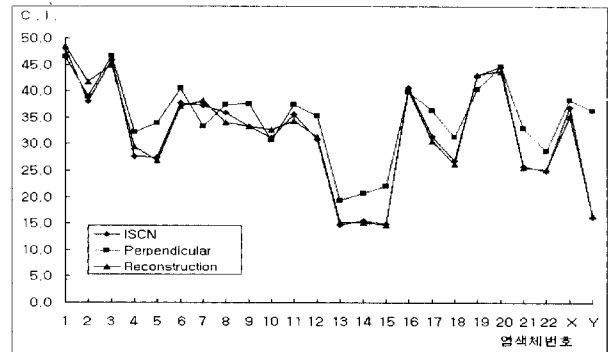
염색체의 특징 파라메타는 염색체를 자동 분류하기 위

해 사용되는 여러 가지 파라메타 중 가장 중요한 요소이며, 염색체 자동 분류에 사용되는 여러 가지 인공지능 기법의 학습 데이터 및 입력 데이터로 이용되고 있다. 그러므로 염색체의 특징 파라메타들이 안정된 값을 유지하며 적은 오차율을 가질 수 있도록 하여야 한다.

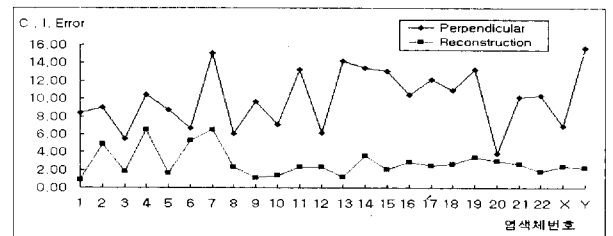
본 연구에서는 임상적인 관점에서 정상인으로 판명된 10명의 염색체 즉, 460개의 염색체 영상들에 본 연구에서 제안한 알고리즘을 사용하여 염색체 영상을 재구성하였으며, 재구성된 염색체 영상의 특징 파라메타 C.I., R.L. 및 R.A.의 값들을 추출하였다.

표 1은 염색체 영상 재구성 알고리즘에 의해 추출된 특징 파라메타 값을 ISCN 및 HC에서 발표한 특징 파라메타 값들과 비교하였으며 아울러 Piper 등에 의해 수행된 기존의 기법에 의해 구해진 특징 파라메타 값들과 비교하였다. 그리고 ISCN 및 HC에서 제시한 특징 파라메타 값 및 기존의 기법을 사용하여 구한 특징 파라메타 값 사이에 발생하는 평균 오차율을 구하여 비교하였다.

이와 같이 비교한 결과를 그래프형태로 도시하면 그림 6, 그림 7, 그리고 그림 8과 같이 된다. 그림 6(a)는 ISCN, 기존의 기법, 그리고 본 연구에서 제시한 영상 재구성 알고리즘에 의해 구한 C.I.값들을 비교한 결과이며.



(a) 동원체 지수(C.I.)
(a) centromeric index



(b) 동원체 지수의 오차
(b) error of the C.I.

그림 6. ISCN, Perpendicular data 및 염색체 재구성에 의해 추출된 C.I.의 비교
Fig. 6. The comparison of extracted C.I. data among ISCN, Perpendicular data and reconstruction algorithm

표 1. ISCN, HC와 기존의 기법과 제안한 알고리즘의 수행에 의한 염색체의 형태학적 특징 파라메타의 비교.

Table 1. Morphological feature parameters comparison for ISCN & HC data, perpendicular method and proposed algorithm. Per : Perpendicular, Rec : Reconstruction, Ori : Original image

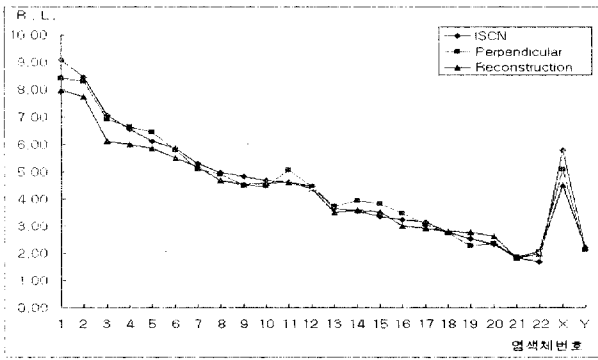
| NO | Centromeric Index | | | | | Relative Length | | | | | Relative Area | | | | |
|----|-------------------|-------|-------|-------|-------|-----------------|------|-------|------|-------|---------------|------|-------|------|-------|
| | ISCN | Per | | Rec | | ISCN | Per | | Rec | | HC | Ori | | Rec | |
| | | mean | error | mean | error | | mean | error | mean | error | | mean | error | mean | error |
| 1 | 48.0 | 46.42 | ±4.21 | 48.41 | ±0.43 | 9.08 | 8.42 | ±1.26 | 7.97 | ±0.69 | 8.44 | 8.70 | ±1.92 | 8.45 | ±1.88 |
| 2 | 38.1 | 38.90 | ±4.50 | 41.70 | ±2.42 | 8.45 | 8.28 | ±1.58 | 7.74 | ±0.78 | 8.02 | 8.79 | ±1.12 | 8.50 | ±1.75 |
| 3 | 45.9 | 46.64 | ±2.72 | 45.01 | ±0.91 | 7.06 | 6.92 | ±1.37 | 6.12 | ±0.37 | 6.83 | 6.86 | ±1.77 | 6.32 | ±0.84 |
| 4 | 27.6 | 32.16 | ±5.21 | 29.44 | ±3.25 | 6.55 | 6.65 | ±3.05 | 6.01 | ±0.35 | 6.30 | 6.66 | ±1.09 | 6.48 | ±1.55 |
| 5 | 27.4 | 33.85 | ±4.35 | 26.88 | ±0.81 | 6.13 | 6.43 | ±1.39 | 5.86 | ±0.38 | 6.08 | 6.45 | ±0.40 | 6.05 | ±1.13 |
| 6 | 37.7 | 40.50 | ±3.32 | 37.11 | ±2.65 | 5.84 | 5.79 | ±0.60 | 5.51 | ±0.50 | 5.90 | 6.00 | ±1.79 | 6.02 | ±1.49 |
| 7 | 37.3 | 33.30 | ±7.54 | 38.22 | ±3.24 | 5.28 | 5.08 | ±0.82 | 5.17 | ±0.88 | 5.36 | 5.18 | ±1.02 | 5.07 | ±1.07 |
| 8 | 35.9 | 37.36 | ±3.02 | 34.12 | ±1.16 | 4.96 | 4.91 | ±1.03 | 4.67 | ±0.57 | 4.93 | 5.05 | ±0.56 | 4.94 | ±0.48 |
| 9 | 33.3 | 37.56 | ±4.84 | 33.23 | ±0.56 | 4.83 | 4.50 | ±1.49 | 4.52 | ±0.52 | 4.80 | 4.64 | ±1.04 | 4.62 | ±0.68 |
| 10 | 31.2 | 30.69 | ±3.57 | 32.69 | ±0.70 | 4.68 | 4.46 | ±0.53 | 4.60 | ±0.50 | 4.59 | 4.61 | ±0.42 | 4.48 | ±1.70 |
| 11 | 35.6 | 37.41 | ±6.65 | 34.38 | ±1.14 | 4.63 | 5.06 | ±1.27 | 4.63 | ±0.72 | 4.61 | 5.02 | ±0.78 | 4.56 | ±1.47 |
| 12 | 30.9 | 35.30 | ±3.09 | 31.35 | ±1.15 | 4.46 | 4.44 | ±0.59 | 4.38 | ±0.46 | 4.66 | 4.80 | ±0.69 | 4.41 | ±0.86 |
| 13 | 14.8 | 19.31 | ±7.11 | 15.14 | ±0.60 | 3.64 | 3.75 | ±1.27 | 3.49 | ±0.82 | 3.74 | 3.89 | ±0.94 | 3.40 | ±1.93 |
| 14 | 15.5 | 20.73 | ±6.70 | 15.18 | ±1.81 | 3.55 | 3.95 | ±0.69 | 3.59 | ±0.32 | 3.56 | 3.79 | ±1.03 | 3.49 | ±0.72 |
| 15 | 14.9 | 22.13 | ±6.55 | 14.65 | ±1.02 | 3.36 | 3.81 | ±1.89 | 3.52 | ±0.29 | 3.46 | 3.48 | ±0.69 | 3.35 | ±0.47 |
| 16 | 40.6 | 39.94 | ±5.22 | 40.26 | ±1.47 | 3.23 | 3.46 | ±1.97 | 3.01 | ±0.13 | 3.36 | 2.94 | ±0.80 | 2.92 | ±0.71 |
| 17 | 31.4 | 36.36 | ±6.09 | 30.48 | ±1.23 | 3.15 | 3.07 | ±0.56 | 2.92 | ±0.19 | 3.25 | 3.02 | ±1.03 | 2.81 | ±0.31 |
| 18 | 26.9 | 31.25 | ±5.47 | 26.18 | ±1.34 | 2.76 | 2.73 | ±0.35 | 2.80 | ±0.26 | 2.93 | 2.70 | ±0.43 | 2.73 | ±0.53 |
| 19 | 42.9 | 40.29 | ±6.63 | 43.09 | ±1.72 | 2.52 | 2.26 | ±0.40 | 2.77 | ±0.93 | 2.67 | 2.02 | ±0.77 | 1.95 | ±0.90 |
| 20 | 44.6 | 44.69 | ±1.93 | 43.75 | ±1.51 | 2.33 | 2.36 | ±0.37 | 2.61 | ±0.25 | 2.56 | 2.09 | ±0.43 | 2.20 | ±0.56 |
| 21 | 25.7 | 32.95 | ±5.08 | 25.54 | ±1.34 | 1.83 | 1.86 | ±1.43 | 1.81 | ±0.30 | 1.90 | 1.56 | ±1.06 | 1.35 | ±0.56 |
| 22 | 25.0 | 28.63 | ±5.17 | 25.23 | ±0.88 | 1.68 | 2.07 | ±0.67 | 1.96 | ±0.27 | 2.04 | 1.77 | ±0.69 | 1.63 | ±0.33 |
| X | 36.9 | 38.37 | ±3.45 | 35.21 | ±1.19 | 5.80 | 5.10 | ±0.51 | 4.54 | ±0.62 | 5.12 | 5.30 | ±1.50 | 4.79 | ±2.40 |
| Y | 16.3 | 36.37 | ±7.82 | 16.60 | ±1.13 | 2.22 | 2.13 | ±0.51 | 2.20 | ±0.34 | 2.15 | 1.80 | ±1.11 | 1.76 | ±0.12 |

영상 재구성 알고리즘에 의한 측정치와 ISCN에서 발표한 값들이 거의 일치함을 알 수 있었다. 그림 6(b)는 ISCN에서 발표한 C.I.값을 기준치로 할 때 기존의 기법 및 영상 재구성 알고리즘에 의해 구한 C.I.값의 오차율을 비교한 결과이며, 기존의 기법보다 영상 재구성 알고리즘에 의해 구해진 C.I.값이 오차율이 적음을 알 수 있었다. 그림 7(a)는 ISCN, 기존의 기법, 그리고 영상 재구성 알고리즘에 의해 구한 R.L.값들을 비교한 결과이며, 염색체의 길이가 긴 경우에는 ISCN 및 기존의 기법보다 다소 적게 나타났으나 염색체의 길이가 짧을수록 ISCN에서 발표한 특징 파라메타 값에 접근함을 알 수 있었다. 이러한 결과는 세션화 기법을 사용할 때 염색체의 장완과 단완의 끝부분에서 영상정보의 소실이 발생하기 때문이다. 소실된 정보는 보간법에 의해 복구되나 극히 적은 양의 정보가 원상 복구되지 않기 때문이다. 그림 7(b)는 ISCN에서 발

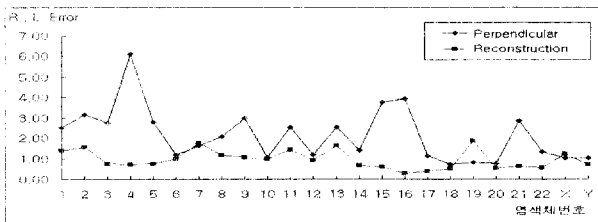
표한 R.L. 값의 오차율을 비교한 결과이며, 기존의 기법보다 영상 재구성 알고리즘에 의해 구해진 R.L. 값이 오차율이 적음을 알 수 있었다. 그림 8(a)는 HC, 염색체의 원영상, 그리고 영상 재구성 알고리즘에 의해 구한 R.A.값들을 비교한 결과이며, 비교적 일치함을 알 수 있었다. 그림 8(b)는 HC에서 발표한 R.A.값을 기준치로 할 때 염색체의 원영상 및 영상 재구성 알고리즘에 의해 구한 R.A.값의 오차율을 비교한 결과이며, 재구성 알고리즘에 의해 구해진 R.A.값이 원영상의 R.A.보다 오차율에 차이가 없음을 알 수 있다.

결론

염색체 분석은 임상의학 분야, 세포 유전학분야 등에서 많은 연구가 진행되고 있다. 염색체의 핵형 분류도 작성



(a) 상대 길이비(R.L.)
(a) relative length



(b) 상대 길이의 오차
(b) Error of the R.L.

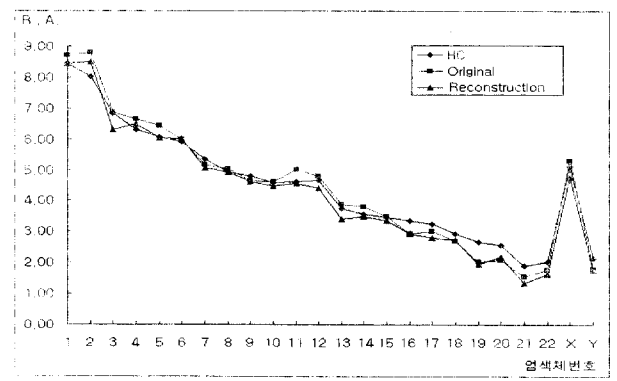
그림 7. ISCN, Perpendicular data 및 염색체 재구성에 의해 추출된 R.L.의 비교

Fig. 7. The comparison of extracted R.L. data among ISCN, Perpendicular data and reconstruction algorithm

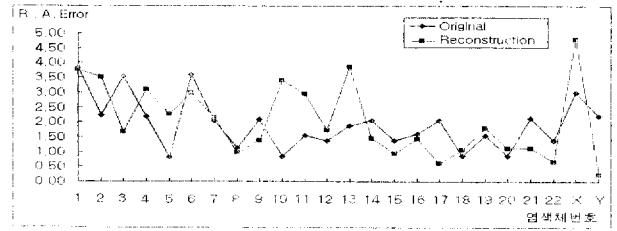
시 현미경상에서 적절한 염색체 군의 탐색, 사진촬영 및 현상, 사진 상에서 각 염색체의 선택 등, 많은 수 작업의 수행으로 인하여 하나의 핵형 분류도 작성에 시간이 많이 소요되고 있는 실정이다. 이로 인하여 염색체를 자동 분류하는 시스템이 요구되어 왔으며 현재까지 많은 연구가 진행되고 있다. 영상분야에서 패턴인식에 사용되는 특징 파라메타는 영상의 전처리과정에서 원영상이 보유하고 있는 영상정보를 소실 없이 추출하는 것이 가장 중요하다. 염색체 영상은 세포배양과정에서의 불균일, 염색체 형태의 비선형, 3차원적인 영상의 2차원적인 검출 등으로 인하여 동일한 번호의 염색체에서도 형태학적 특징 파라메타가 다르게 나타난다. 따라서, 오차율이 적고 안정한 특징 파라메타의 추출이 요구된다.

본 연구에서는 안정한 특징 파라메타를 구하기 위하여 타 연구에서 수행되지 않은 염색체 영상의 재구성 알고리즘을 적용하여 수행한 결론은 다음과 같다.

염색체의 특징 파라메타를 추출하기 위하여 기존에 사용되는 기법을 적용시 C.I.의 오차율은 약 10.2%이었으나 본 연구에서 제안한 염색체 영상 재구성 알고리즘을 적용할 경우 C.I.의 오차율은 약 2.81%이었다. 그리고 기



(a) 상대 면적비(R.A.)
(a) relative area



(b) 상대 면적비의 오차
(b) Error of the R.A.

그림 8. HC, 염색체의 원 영상 및 32 염색체 재구성에 의해 추출된 R.A.의 비교

Fig. 8. The comparison of extracted R.A. data among ISCN, Perpendicular data and reconstruction algorithm

존의 기법을 적용시 R.L.의 오차율이 약 2.13%로 나타났으나 본 연구에서 제안한 알고리즘을 적용할 경우 R.L.의 오차율이 약 0.95%이었다. 또한 R.A.에서는 제안한 알고리즘과 원 영상에서의 오차율은 각각 약 2.03%와 1.92%로 나타났으며, 기존의 기법은 영상정보를 이치화 하는 등 운용자의 개입이 필요한 반면 제안한 알고리즘은 운용자의 개입 없이 특징 파라메타를 추출하는 전 과정을 수행할 수 있었다.

이상과 같이 특징 파라메타의 결과 값은 C.I. 및 R.L.에서 오차율이 적고 안정한 값을 가짐을 알 수 있었고, R.A.는 염색체의 원영상에서 추출된 값과 거의 유사하여 안정된 결과 값을 나타낸다. 또한 재구성된 염색체 영상에서 염색체의 특징 파라메타를 추출하여 염색체 자동 분류에 활용하면 염색체 자동분류의 정확도를 더욱 높일 수 있으리라 생각되며, 이에 대한 연구가 진행 중이다. 형태학적 특징 파라메타이외의 농도정보도 정량화하여 사용할 수 있어 염색체 자동 분류시 정확도가 우수한 패턴 인식 시스템을 구성할 수 있으리라 생각되며, 이러한 영상 정보를 사용하여 ISCN에서 염색체를 도식화하여 발표한 형태로 구성하면 염색체의 자동분류뿐만 아니라 선천성 질

환에 관한 진단시스템의 응용에도 사용 가능하리라 생각된다.

참 고 문 헌

1. Robert S. Ledly, "High-speed automatic analysis of biomedical picture," Science, vol. 146, pp. 216-223, 1964.
2. J. M. Cho, and D. H. Hong, "Computer-assisted karyotyping system of Giemsa-stained chromosomes (II)," Proc. of 1989 Korea-Japan Joint Conference on MBE, pp. 19-23, Sep. 21-22, 1989.
3. Lucas J. van Vliet, Ian T Young, and Brian H. Mayall, "The athena semi-automated karyotyping system1," Cytometry, vol. 11, pp. 51-58, 1990.
4. Brian H. Mayall, James D. Tucker, Mari L. Christensen, Lucas J. van Vliet, and Ian T Young, "Experience with the athena semi-automated karyotyping system," Cytometry, vol. 11, pp. 59-72, 1990.
5. Lerner, B., Levinstein, M., Rosenberg, B., Guterman, H., Dinstein, I., and Romem, Y., "Feature selection and chromosome classification using a multilayer perceptron neural network", IEEE International Conference on Neural Networks, vol. II.2/7, pp. 3540-3545, Jun. 28-Jul. 2, 1994.
6. 조종만, 신경회로망을 이용한 염색체의 핵형 분류에 관한 연구. 인하대학교 박사학위논문, 1994.
7. J. Piper, E. Granum, D. Rutovitz, and H. Ruttledge, "Automation of chromosome analysis," Signal Processing, vol. 2, pp. 203-221, 1980.
8. Jens Gregor and Erik Granum, "Finding chromosome centromeres using band pattern information," Comput. Biol. Med., vol. 21, No. 1/2, pp. 55-67, 1991.
9. An International System for Human Cytogenetic Nomenclature(ISCN), KARGER, 1985.
10. D. E. Rooney, B. H. Czepulkowski, Human Cytogenetics Essential Data, John Wiley & Sons, 1994.
11. Rafael C. Gonzalez and Paul Wintz, Digital image processing, Addison-Wesley, pp. 398-403, 1987.
12. 김명남, 조진호, "좌심실에 대한 2D 단축 심초음파도에서 기하학적 모델을 이용한 심내벽 윤곽선의 자동 검출", 대한의공학회지, Vol. 15, No. 4, pp. 447-453, 1994