Iterative Histogram Matching Algorithm for Chromosome Image Enhancement Based on Statistical Moments

Seyed Pooya Ehsani, Hojjat Seyed Mousavi, Babak Hosseini Khalaj
Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran
Email: pooya.ehsani@gmail.com, hojjatseyedmousavi@gmail.com, khalaj@sharif.edu

***Abstract*—Vivid banding pattern of the chromosome image is a crucial part for diagnosis in karyotype medical test. Furthermore, thriving computer aided segmentation and classification depend on the initial image quality. In this paper, we propose an adaptive and iterative histogram matching algorithm for chromosome contrast enhancement especially in banding patterns which is one of the most important information laid in chromosome image. Objective histogram, with which the initial image needs to be matched, is created based on processes on the initial image histogram. Calculation of statistical moments of image histogram and determination of parameters in each step of iteration based on these moments form the algorithm structure. In order to assess the performance of the proposed algorithm, Density Profile (DP) and the contrast improvement ratio (CIR), are utilized. Experimental results indicate that the proposed method shows better results in both quantitative gauges of CIR and DP, and visual perception.**

***keywords*- chromosome karyotyping, contrast enhancement, histogram matching, statistical moments, contrast improvement ratio.**

I. Introduction

Contrast in chromosome images is of vital importance in many medical applications. Visual examination of medical images especially chromosome images is very essential in diagnosis of many diseases [1]. This is the importance of existence of the contrast in chromosome images. Linear contrast stretching and histogram equalization are two frequently used methods among various techniques for contrast improvement [2]. One of the problems of these algorithms is their globality which affect whole image regardless of local details. Moreover, global histogram equalization (GHE) may lead to overall noise increment [3].

Due to significance of details in chromosome images like other types of medical images, local contrast improvement is crucial. Adaptive contrast enhancement (ACE) [3]–[6], Adaptive histogram equalization (AHE) [7], [8] and Adaptive iterative histogram matching (AIHM) [9] are local enhancement methods which are used to assess the performance of our proposed algorithm.

In this paper, we propose an adaptive and iterative histogram matching algorithm for chromosome contrast enhancement. Objective histogram, with which the initial image needs to be matched, is created based on some processes on the initial image histogram. Hidden information laid in the histogram of initial image, makes the enhancement procedure more dependent to the input image and results in better contrast improvement. Moreover, continuous and moderate contrast enhancement of the chromosome image is the advantage of iterative procedure. The iteration steps may vary depending on the image characteristics and histogram. Calculation of statistical moments of image histogram and determination of parameters based on these moments in each step form the structure of the proposed algorithm. Utilizing statistical moments is the main difference between this and our previously proposed method in [9].

This paper is organized as follows: Section II describes the model presented for histogram of chromosome images. In section III we present our proposed algorithm for contrast enhancement. A contrast assessment is presented and the results are compared with conventional enhancement algorithms in Section IV. Conclusions are drawn in Section V.

II. Histogram Model

In this section, a model is presented for histogram of chromosome images, named h:

$$h(j) = \|\{p(x,y)|p(x,y) = j\}\| \qquad 0 \leq j \leq 255 \qquad (1)$$

Where $\|A\|$ is the length of set A. Since most samples are 8-bits *gray scale* images, it is assumed that $h(j)$ consists of 256 bins.

In the presented model a specific number of adjacent bins of conventional histogram are combined to each other to create a reshaped histogram with fewer bins and represent the coarse changes of histogram [9]. The reshaped histogram $H(j)$ is defined as:

$$H(j) = \sum_{i=(j-1)\times L}^{j\times L-1} h(i) \qquad 1 \leqslant j \leqslant \frac{256}{L} = J \qquad (2)$$

Where L is the number of adjacent bins combined into one bin.

According to visual perception, intensities of histogram are divided into three regions, *black*, *gray*, and *white* which are processed separately in the proposed algorithm [9]. All the processes are performed on images with white background. These regions are defined below:

Black region: Dark pixels, corresponding to dim regions in chromosome image, belong to this region. The region bounds are zero from the below and 63 from the above. Due to quality of the initial image number of pixels fallen in this category is much less than other two regions.

Gray region: Pixels in the range of [64,223] are fallen in this region. Most of the information laid on the inner structure of chromosome is in this region except the circumference pixels of the chromosome and dark or bright pixels of the chromosome.

White region: This region consists of intensities in the range of [223,255]. In addition to bright regions of chromosome, the pixels of this region correspond to gradual whitening in circumference of the chromosome.

It is worthy to mention all bounds of the defined regions were determined experimentally.

III. Contrast Enhancement Based on Adaptive Histogram Matching

A. Presented Algorithm

Creating an objective histogram for iterative histogram matching based on unique characteristic of each region is the essence of proposed algorithm. In the remaining of this paper the following variables are defined as:

$h_k(j)$: Histogram of the input image at k^{th} step of iteration. $0 \leqslant j \leqslant 255$

$H_k(j)$: Reshaped histogram of the input image at the k^{th} step of iteration. $1 \leqslant j \leqslant J$
$l_k(j)$: Objective histogram used for histogram matching at the k^{th} step of iteration. $0 \leqslant j \leqslant 255$
For convenience we define:

$$l_k^j(i) = l_k(j \times L - \frac{L}{2} + i) \qquad -\frac{L}{2} \leqslant i \leqslant \frac{L}{2} \tag{3}$$

Im_k: Output image of the algorithm at the k^{th} step of iteration.
The algorithm creates the objective histogram, $l_k(j)$, in each three histogram regions individually, based on the reshaped histogram of Im_{k-1}. The histogram matching process is done on Im_{k-1} as the input of the algorithm at k^{th} step of iteration by assistance of $l_k(j)$, and output image, Im_k, is generated. The iteration process continues until output contrast reaches an acceptable level.
The critical step of creating $l_k(j)$ in each region of histogram is described below:
Black region: Considering relative lack of pixels in this region, $l_k(j)$ is determined in order to result in a gradual increment in density of pixels in *black region*. This process will also lead to piecemeal inclination of border intensities of *black region* and *gray region* to darker intensities. Thus there will be a contrast enhancement in spatial regions corresponding to these intensities. To do so, the value of each bin in $H_k(j)$ in *black region* is multiplied by a parameter named black multiple (BM_k), which varies in different steps of iteration:

$$l_k^j(0) = H_k(j) \times BM_k \qquad j \in BlackRegion \tag{4}$$

Gray region: Difference in the number of pixels in adjacent bins of $H_k(j)$ is the base principle for determination of $l_k(j)$ in *gray region*. Great accumulation of pixels in one bin and dramatic difference in height of adjacent bins in $H_k(j)$ depict that large amount of fundamental pixels of chromosome surface are limited in range. This phenomenon impairs image contrast. Thus for overcoming this issue, $l_k(j)$ must be created in a way to lessen the difference in adjacent bins of $H_k(j)$ by increment in bin with fewer pixels and decrement in bin with more pixels. This concept is more visible through (5) and (6).

$$\begin{cases} \text{diff} = (H_k(j+1) - H_k(j))/NP_k \\ l_k^j(+1) = H_k(j) + \text{diff} \qquad H_k(j) < H_k(j+1) \\ l_k^{j+1}(-1) = H_k(j) - \text{diff} \end{cases} \tag{5}$$

$$\begin{cases} \text{diff} = (H_k(j) - H_k(j+1))/NP_k \\ l_k^j(+1) = H_k(j) - \text{diff} \qquad H_k(j) \geqslant H_k(j+1) \\ l_k^{j+1}(-1) = H_k(j) + \text{diff} \end{cases} \tag{6}$$

$$j \in GrayRegion$$

In the equations above, Normalization parameter of step k NP_k, adjusts the influence of diff on creation of $l_k(j)$.
White region: Likewise to *black region* total number of pixels in *white region* has to be increased. A similar method is used to boost the number of pixels in this region regarding to a parameter, named white multiple WM_k, as in (7).

$$l_k^j(0) = H_k(j) \times WM_k \qquad j \in WhiteRegion \tag{7}$$

Another issue with which initial low contrast images face, is gradual whitening in border of chromosomes due to imperfections of imaging system. For overcoming this issue, the value of brightest intensity of objective histogram, $l_k(255)$ is calculated based on (8).

$$l_k^J(\frac{L}{2} - 1) = WS \times [\frac{L}{2} - 1] + l_k^J(0) \tag{8}$$

This equation implicates an interpolating line connecting $l_k^J(0)$,obtained from the brightest bin in H_k ($H_k(J)$), to value of the brightest bin in l_k, $l_k(255)$ according to adjustable slope named white slope, WS. This process leads to a steady inclination of surrounding pixels to brighter intensities.
After determining mid points of $l_k(j)$ through equations (4) - (8), linear interpolation is done to determine piecewise linear objective histogram $l_k(j)$ at the whole range of 0 to 255. This completes the process of creation of our objective histogram.

B. Determination of Parameters

The final problem remained unsolved is determinatin of appropriate parameters for (4) - (8). Here, a method is developed based on mean, 3^{rd} and 4^{th} order central statistical moments of the image. Normalized $h_k(j)$ can be considered as a discrete random variable with a probability density function as in (9).

$$P_k(x) = \begin{cases} \frac{h_k(x)}{\sum_{j=0}^{255} h_k(j)} & 0 \leq x \leq 255 \\ 0 & \text{O.W.} \end{cases} \tag{9}$$

Mean of the defined probability density function $P_k(x)$, is calculated by the following equation:

$$\mu_k = \int_0^{255} x \times P_k(x)dx \tag{10}$$

Similarly, 3^{rd} and 4^{th} order central statistical momentums of $P_k(x)$, named $skewness$ and $kurtosis$ respectively, can be defined as:

$$\lambda_k = E[(\frac{x-\mu_k}{\sigma_k})^3] = \frac{1}{\sigma_k^3}\int_0^{255}(x-\mu_k)^3 \times P_k(x)dx \tag{11}$$

$$\kappa_k = E[(\frac{x-\mu_k}{\sigma_k})^4] = \frac{1}{\sigma_k^4}\int_0^{255}(x-\mu_k)^4 \times P_k(x)dx \tag{12}$$

In the above equations σ_k, λ_k, and κ_k are used as abbreviation signs of $standard\ deviation$, $skewness$, and $kurtosis$ of Im_k respectively.
$Skewness$ of a probability function, is a measure of asymmetry of the random variable around its mean. Positive $skewness$ depicts more density over left side of the mean, while negative $skewness$ implies more concentration of values over right side of the mean. Similarly, $kurtosis$ is a descriptor of the shape of probability distribution which implies some information about peakedness of the random variable. Higher $kurtosis$ depicts sharper peaks of PDF of the random variable and more abrupt descending over its maximum value [10].
In addition to proper contrast improvement, some of the main characteristic features of modified image should be taken into consideration for selecting appropriate parameters. For instance, parameters should not be set such that they cause critical darkening or whitening of chromosome. One of the solutions for this problem is relative stability of the $mean$ and $skewness$ of image. Based on the amount of unbalanced darkness or brightness of the image at k^{th} step of iteration, BM_{k+1} and WM_{k+1} will be determined to neutralize this asymmetry. Growth in μ_k in comparison to μ_{k-1} implies overall whitening which necessitates increment in BM and decrement in WM and vice versa. In addition growth in λ_k depicts more aggregation in darker regions of histogram, requiring decrement in BM and increment in WM. These concepts can be more apprehensible through (13) and (14).

$$BM_{k+1} = BM_k \times (1 + \frac{\mu_k - \mu_{k-1}}{\mu_k})^{\alpha} \times (1 - \frac{\lambda_k - \lambda_{k-1}}{\mid \lambda_k \mid})^{\beta} \tag{13}$$

$$WM_{k+1} = WM_k \times (1 - \frac{\mu_k - \mu_{k-1}}{\mu_k})^{\alpha} \times (1 + \frac{\lambda_k - \lambda_{k-1}}{\mid \lambda_k \mid})^{\beta} \tag{14}$$

in the above equations, α and β are predefined constants representing the effect of each term on overall changes of BM and WM.

Moreover, increase in κ_k to κ_{k-1} represents more concentrated values around its maximum. Decrement in κ can be a sign of more uniformity in aggregation of values in histogram and higher dynamic range of image. Thus ideal behavior of κ is descending behavior which has great correlation with changes in NP. If κ_k is increased in comparison to κ_{k-1}, NP_{k+1} must be reduced to emphasize the process of equalization of heights of bins in $H_k(j)$. This change in NP increases diff in (5) and (6) and will cause quicker decrement in height of greater bins and increment in height of smaller bins. This concept is described briefly in (15).

$$NP_{k+1} = \begin{cases} NP_k \times (1 - \frac{\kappa_k - \kappa_{k-1}}{|\kappa_k|})^\gamma & \kappa_k > \kappa_{k-1} \\ NP_k & \kappa_k \leq \kappa_{k-1} \end{cases} \quad (15)$$

γ determines dependency of NP to changes in κ.
It is worthy to mention for our sets of samples, we found the typical values for the parameters can be L = 8, 16, 32, BM_1 = 1.5...3.5, WM_1 = 1.5...2.5, NP_1 = 5...7, WS = 1...2, α = 5, β = 1, and γ = 2 experimentally.

IV. Results and Comparisons

A. Contrast Enhancement Measurement

Every image contrast enhancement algorithm needs an evaluation method to compare each algorithm results and specify the more efficient one. in This paper we utilize two different assessment algorithms. The first one is called Contrast Improvement Ratio (CIR) proposed in [11], [12] and the latter is called Density Profile [13].

1) CIR: In CIR method, a local contrast is defined for each pixel of the image:

$$c(x,y) = \frac{|\, p - a \,|}{|\, p + a \,|} \quad (16)$$

where p and a are the average values of gray levels in two concentric windows of size 3×3 and 7×7 for each pixel. Here $c(x,y)$ is the contrast measurement and is in the range of $[0,1]$. Using local contrasts of original and processed images the CIR is computed.

$$CIR = \frac{\sum_{(x,y)\in R} \left(c(x,y) - \hat{c}(x,y)\right)^2}{\sum_{(x,y)\in R} c^2(x,y)} \quad (17)$$

Where c and $\hat{c}$ are the local contrasts in original and processed images respectively. R is the region of interest which we may want to evaluate the contrast improvement. R is assumed to be the whole image of chromosome in our simulations.
As it is mentioned in section III our algorithm is an iterative method and iterations continue until the image quality reaches an acceptable level. At first more iterations may seem more efficient but it is not the case. The diagram of CIR as a function of iteration steps is depicted in figure 1 in which one can see CIR shows an S-shaped behavior. It is an increasing function and has a maximum, which corresponds to iterations where chromosome image is degraded and the image tends to become dark. The greater CIR means more contrast enhancement, and image contrast may be improved but the image may be degraded. So our Proposed algorithm uses CIR Diagram and stop the iteration before the CIR reaches to its maximum. By visual perception, the best point to stop iterating is near the inflection point of the diagram. Around inflection point the normalized increment of the CIR is very high and we can stop the iteration when the normalized increment of the CIR exceeds a Threshold. This point will be very close to the inflection point.

2) Density Profile: Another method which is introduced according to specific properties of chromosomes is Density Profile (DP) [13]. In this method, first the axis of the chromosome is obtained and then for each pixel of the axis the mean value of grayscales belonging to chromosome in a perpendicular segment to main axis is computed. Density Profile is obtained by drawing these mean values as a function of spatial location of the points on chromosome axis.

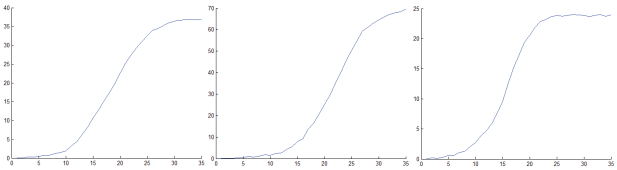


Fig. 1: CIR diagram as a function of iteration number for three different samples of chromosomes, our proposed method stops the iterating when diagram meets an inflection point. For instance the most left image iteration will stop after about 17 or 18 iterations.

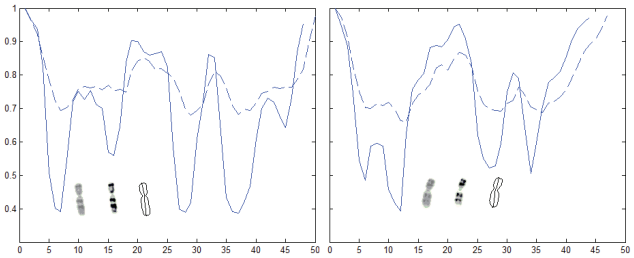


Fig. 2: DP for two different samples of chromosomes. Dashed line correspond to the original image and continuous line to processed image. The original image, the processed image using our algorithm and its corresponding skeleton and main axis are sketched on the diagram from right to left, respectively.

In the figure 2 DPs of two different samples of chromosomes are presented. As it can be seen, the Profile can show the banding pattern of the chromosome. When the chromosome band is dark PD diagram consits of a local minimum and when the band is bright, the PD consits of a local maximum. The contrast enhancement method which strengthen the distance between maximums and minimums of the Density Profile is considered to be a desired contrast enhancement algorithm specifically for chromosome applications. The PD diagram can be quantized by its dynamic range, that is, the difference between global minimum and maximum. It is worthy to mention that the two tails of the PD are always 1 because they belongs to background graylevels and this two maximums are excluded from the PD's dynamic range. In order to evaluate our new proposed algorithm, we compare it with two conventional enhancement algorithms and our method proposed in [9]. These conventional algorithms are CGT and LSD ACE algorithms [3]–[6]. the CGT and ACE parameters used in [5] are 5, 12 and 7 for gain constant in CGT and LSD and window size respectively. The parameters used in AIHM [9] are $L = 32$, $BM = 2.2$, $WM = 1.1$, $WS = 1.5$, and $NP = 7$. The proposed algorithm has evaluated with parameters $L = 32$, $BM = 2.8$, $WM = 1.9$, $WS = 1.5$, and $NP = 7$.

B. Compared Results

In the evaluation of the proposed method, a set of human single chromosomes were tested. For some of our database samples CIRs are measured for different methods and tabulated in table I and PD dynamic ranges are tabulated in table II. As it can be realized from the results, the proposed method yields the highest CIRs and DP among the tested methods. Here we compare the visual enhancement of conventional algorithms such as CG, LSD, AIHM in [9] and our newly revised method in Figure 3. One can see that

TABLE I: Comparison of the CIRs obtained from different methods for enhancing single chromosome images.

Image	CG	LSD	AIHM	our method
1	1.9521	1.0774	15.2651	16.1388
2	2.3995	1.4994	7.7643	8.4452
3	1.8574	0.8967	10.3944	21.3419
4	2.932	1.301	3.9898	7.6793
5	2.185	0.9299	5.4817	7.5477
6	2.3715	1.0402	4.3583	5.8406
7	1.1926	2.3774	16.4825	17.6476
8	1.3961	2.8296	6.8985	7.6871
9	1.7816	2.7262	7.0807	8.358
10	1.001	2.0532	9.1336	8.2337
11	1.0232	2.1099	5.1633	11.8407
12	1.079	2.0884	4.5688	10.6711
average	1.76425	1.744108333	8.048416667	10.95264167

TABLE II: Comparison of the PDs dynamic range from different methods .

Image	Initial image	CG	LSD	AIHM	our method
1	0.1316	0.2049	0.2154	0.4196	0.4378
2	0.1381	0.2058	0.2014	0.2801	0.3018
3	0.1798	0.2444	0.2681	0.3406	0.3381
4	0.1356	0.1969	0.2199	0.2376	0.2987
5	0.1206	0.1764	0.2033	0.3031	0.3611
6	0.1679	0.2034	0.22	0.2879	0.3321
7	0.2495	0.2757	0.2955	0.5652	0.5674
8	0.2232	0.2692	0.29	0.4037	0.4377
9	0.1318	0.1891	0.2058	0.3993	0.399
10	0.2347	0.2446	0.2688	0.4944	0.4425
11	0.1662	0.2231	0.2539	0.337	0.4271
12	0.1312	0.1698	0.197	0.3541	0.4682
average	0.16751	0.21694	0.23659	0.36855	0.40667

our revised proposed method produces the best visualization effect on banding pattern after enhancing the image. In addition to better image enhancement result, one of important features of current algorithm is low complexity computations in comparison with the method in [9]. That is, for the same image our algorithm concludes the iteration process in considerably less steps in comparison with the method in [9]. The other advantage of proposed algorithm over [9] is that the problem of intense darkening in some samples has been solved and the percentage of failure has been decreased.

It is worthy to mention that sample images are obtained through (*Q-banding*), and inverted for better visualization. The pattern of bands in *Q-banding* is very similar to that seen in *G-banding* [14].

V. Conclusion

In this paper, we proposed a new iterative algorithm for contrast enhancement of medical images of chromosomes, based on adaptive histogram matching. We first proposed a model for histogram of chromosome images. With assistance of this model a reference histogram created and used for histogram matching in each iteration. There are some parameters in the presented model which are adaptively changed in the iteration process and their weighting could be chosen to meet different requirements and acquire different results. In this method in addition to visual enhancement of the images according to conventional algorithm and our previous method in [9] (AIHM) the complexity decreased a lot and iteration process concludes faster. Moreover, the overall rate of success is increased. Eventually, detailed simulations were carried out using various sets of single chromosomes, showing that the proposed method enhances details satisfactorily. Research in this field is in progress.

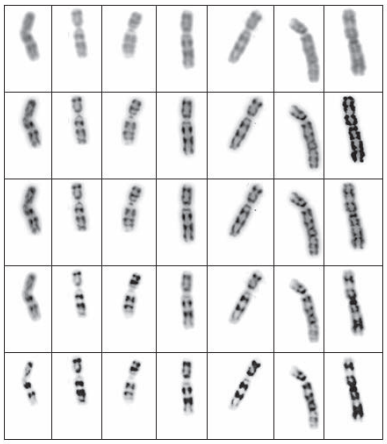


Fig. 3: A comparison of different methods for enhancing single chromosome images. Columns from top to bottom: initial image, CG transform , LSD transform, previous AIHM and our proposed method. Samples are obtained by inverting the Q-banding images of chromosomes.

Acknowledgement

The authors would like to thank M. Fotuhi from CE department of Sharif University of Technology, Tehran, for his helpful support in providing valuable suggestions and medical images.

References

[1] Offner S., "A Plain English Map of the Human Chromosomes." *The American Biology Teacher*, pp. 87-91, Feb. 1992.

[2] R. C. Gonzalez and R. E.Woods, Digital Image Processing. New York: Addison-Wesley, 1992.

[3] T.-L. Ji, M. K. Sundareshan, and H. Roehrig, "Adaptive image contrast enhancement based on human visual properties", *IEEE Trans. Med. Imag.*, vol. 13, pp. 573-586, Aug 1994.

[4] Sijing Chen, Chenyang Ge, "Adaptive Contrast Enhancement Based on Detection of Histogram Feature", *Chinese Conf. on Patt. Recog.*, 2008.

[5] D.-C. Chang and W.-R. Wu, "Image Contrast Enhancement Based on a Histogram Transformation of Local Standard Deviation", *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 518-531, Aug 1998.

[6] Ch. H. Ooi and N. A. M. Isa, "Adaptive Contrast Enhancement Methods with Brightness Preserving", *IEEE Trans, Consumer Electronics.*, vol. 56, no.4, Nov. 2010.

[7] J. A. Stark, "Adaptive Image Contrast Enhancement Using Generalizations of Histogram Equalization", *IEEE Trans. Imag. Proc.*, vol. 9, pp. 889-896, May 2000.

[8] T. Kim and J. Paik, "Adaptive Contrast Enhancement Using Gain-Controllable Clipped Histogram Equalization", *IEEE Trans, Consumer Electronics.*, vol. 54, no. 4, Nov 2008.

[9] S. P. Ehsani, H. S. Mousavi, B. H. Khalaj, "Chromosome image enhancement using adaptive, iterative histogram matching", accepted to be published on *Proc.* 7^{th} *iranian Conf. on Machine vision and Image processing.*, Nov. 2011.

[10] Wikipedia, the free encyclopedia, Kurtosis, [Online]. Available: http://en.wikipedia.org/wiki/Kurtosis

[11] R. Gordon and R. M. Rangayyan, Feature enhancement of film mammograms using fixed and adaptive neighborhoods, *Appl. Opt*, vol. 23, no. 4, pp. 560-564, 1984.

[12] Y.-P. Wang, Q. Wu, K. R. Castleman, and Z. Xiong, Chromosome Image Enhancement Using Multiscale Differential Operators, *IEEE Trans. Med. Imag.*, vol. 22, no. 5, pp. 685-693, May 2003.

[13] E. Poletti, E. Grisan and A. Ruggeri, "Automatic classification of chromosomes in Q-band images", *IEEE EMBS.*, pp. 1911-1914, 2008.

[14] Wikipedia, the free encyclopedia, Karyotype, Jul 2011. [Online]. Available: http://en.wikipedia.org/wiki/Karyotype