A Study for Extracting Nucleolus Candidate Zone in WBC

Ivan Vincent*, Suk-Hwan Lee**, Kwang-Seok Moon***, Ki-Ryong Kwon*

White Blood Cancer is one of the deadliest diseases for human being and its presence can be early diagnosed by detecting the existence of nucleolus in White Blood Cells (WBC) of peripheral blood smear images. In this paper, the author would like to compare the performance of nucleolus candidate zone extraction algorithm by utilizing Wavelet transforms and Curvelet transform. The result confirms that Curvelet transform shows a better performance than wavelet transform in extracting the nucleolus candidate zone. The authors hope that the result of this research can improve the accuracy in diagnosing white blood cells related diseases in the future.

Keywords: Nucleolus Candidate Zone, White Blood Cells, Curvelet Transform

1. Introduction

White Blood Cells (WBCs) hold a very important role in human body immunity system. Each cell contains nucleus as the main part of the cell and cytoplasm to provide support to the internal structures of a cell by being a medium for their suspension as shown in figure 2. White blood cells usually clustered into 2 major groups of myelogenic and lymphogenous based on their physical characteristic such as shape, size, edge and position [1].

Leukemia or white blood cancer is one of the deadliest diseases for human being which proliferate white blood cells very rapidly and in an uncontrollable manner. This proliferation process usually happen inside human bone marrow and involving peripheral blood. Leukemia can be clustered into 2 major categories of acute and chronic leukemia, and each category can be divided further based on the type of infected cell whether it is myeloid or lymphoid. By this form of clustering (figure 1), physicians acknowledge 4 types of Leukemia, they are Acute Myeloid Leukemia (AML), Acute Lymphoid Leukemia (ALL), Chronic Myeloid Leukemia (CML), and Chronic Lymphoid Leukemia (CLL).

Fig. 1. Leukemia Clusterization.

According to previous medical research, it was known that acute leukemia is characterized by proliferation of immature cells while chronic leukemia is characterized by proliferation of mature cells [1]. Because nucleolus only possibly exists in immature cells, in this paper our main focus is to detect the existence of nucleolus candidate zone in acute leukemia cell group. There have been various research conducted in this area especially for image segmentation technique such as color and shape features based segmentation and iteration algorithm such as snake algorithm.

In this paper, the author will present the performance comparison between wavelet and curvelet transform used in nucleolus candidate zone extraction algorithm. This algorithm initiated with performing nucleus segmentation procedure from peripheral blood smear image and followed by applying Wavelet and Curvelet transform to the nucleus segmented image to detect the existence of nucleolus inside the nucleus. The result from this experiment clearly shown that Curvelet transform has a better performance in segmenting the nucleolus candidate zone than wavelet transforms. The evaluation of this algorithm qualitatively performed by examining the accuracy of segmentation result and quantitatively done by using Hausdorff dimension value which measure perimeter roughness of the selected region of interest.

Fig. 2. Blood smear image features [11].

2. Related works

Over years many researcher already focus on Leukemia detection algorithm. There are some researchers who focus on automated Leukemia detection system, such as S.Mohapatra et.al. [6] who perform their research using various quantitative method i.e. features extraction using fractal dimension, contour signature, shape feature, and use SVM as classifier. Even more specific automated blast classification of acute leukemia using HMLP network is already researched by N.H. Harun et.al. [5].

But there are also other researchers who focus only in the improvement of feature extraction method such as R.Soltanzadeh et.al. [1] who improve the nucleolus extraction method using curvelet transform and proposed a new feature idea to use gradient difference in saturation channel to differentiate cell type. And other researcher who uses geometrical, texture, and statistical
3. Methodology

3.1 Nucleus Segmentation

Nucleus segmentation process is the first half of overall algorithm which utilizes a series of basic segmentation methods, start from read the target RGB image, and divides it into three channels: red, green, and blue to perform contrast adjustment in each individual channel. Median filter is also applied to red and green channel to maintain edge information as much as possible. The next step is to convert the color system from RGB into LUV and prepare for clustering process, as it is shown in figure 3. Peripheral blood smear image is mainly consists of 3 different objects: the red blood cells, white blood cells, and the background. Based on that fact, the clustering process can be simply done in three different classes representing each group in the image.

In this paper, the author uses K-means clustering method based on Euclidean distance and separates the object into three classes. Further processes are performed to the selected cluster by convert the image into grayscale; adjust its contrast level for determining the threshold value, and another conversion to binary image based on the automatically selected threshold value. The output of the process is a binary mask of the selected Region of Interest (ROI) as shown in figure 5b, which is applied to the original cell image (figure 5a) to segment the nucleus part of the image. Figure 5c displays result of this segmentation process.

A. Wavelet Transform

Wavelet transform has been known for its good performance in image processing applications such as image denoising, compressing, and features extraction. However, there are still some limitations in representing geometric properties of the structure [3]. In wavelet transform, a signal or function \( f(t) \) can be analyzed if expressed as a linear decomposition by

\[
f(t) = \sum a_i \psi_i(t)
\]

where \( l \) is the integer index, \( a_i \) are the real-valued expansion coefficients, and \( \psi_i(t) \) are expansion set (wavelet basis). In this paper, the author uses two different type of wavelet basis, Haar and Daubechies, apply decomposition in four-level wavelet transformation. Haar wavelet scaling function and mother wavelet function requirement are as follows :

\[
\phi(t) = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \quad \psi(t) = \begin{cases} 1, & 0 \leq t < 1/2 \\ -1, & 1/2 \leq t < 1 \\ 0, & \text{otherwise} \end{cases}
\]

(2)

While Daubechies scaling function coefficients are formulated as :

\[
h_0 = \frac{1+\sqrt{2}}{4\sqrt{2}}, \quad h_1 = \frac{3+\sqrt{2}}{4\sqrt{2}}, \quad h_2 = \frac{3-\sqrt{2}}{4\sqrt{2}}, \quad h_3 = \frac{1-\sqrt{2}}{4\sqrt{2}}
\]

(3)

And the mother wavelet function are a result of mirroring operation of the scaling function coefficients resulting as follows:

\[
g_0 = h_3, \quad g_1 = -h_2, \quad g_2 = h_1, \quad g_3 = -h_0
\]

(4)

The result of wavelet decomposition is a one row matrix coefficient (C) which constructed as a result of concatenation of each subband as follows:

\[
C = |c_{A_N}|c_{H_N}|c_{V_N}|c_{D_N}| \cdots, |c_{A_1}|c_{H_1}|c_{V_1}|c_{D_1}|
\]

(5)

\( N \) is the number of wavelet decomposition level, \( H, V, \) and \( D \) represent wavelet horizontal, vertical, and diagonal direction respectively [8]. Then the coefficients (C) are modified using the following threshold suggested in [1] and [3] except for the scaling function. Using the new coefficient value and powering the approximation coefficient by tenth (enhance the low resolution function). Using the new coefficient value and powering the approximation coefficient by tenth (enhance the low resolution function), reconstruction process (inverse wavelet transform) is now ready to be performed. Please refer to [8] for further information about wavelet transform.

B. Curvelet Transform

Curvelet transform is a multi-scale directional transform that allows an almost optimal non-adaptive sparse representation of objects with edges [2] [4] so it should capable to perform better than wavelet transform to detect 2D singularities in image.

Continuous-time Curvelet transform use the coordinate system in two dimension with spatial variable \( x \), frequency domain variable \( \omega \), and with \( r \) and \( \theta \) as polar coordinates in frequency domain. It also has “mother” curvelet in the sense that all curvelets at scale \( 2^{-j} \) are obtained by rotations and translations of \( \varphi_j \), which form is stated as :

\[
\varphi_{j,k}(x) = \varphi_j(R_\theta(x - x^{(l,j)}/2^j))
\]

(6)

Where \( R_\theta \) is the rotation by \( \theta \) radians, \( x_k \) is the position, \( l \) is the angle, and \( j \) is the scale variable. From this knowledge we now can define the Curvelet coefficient as the inner products between element \( f \) in \( L^2(\mathbb{R}^2) \) and a Curvelet \( \varphi_{j,k} \) as stated in (6) which formula is shown in (7).

\[
c(j,l,k) := \langle f, \varphi_{j,l,k} \rangle = \int_{\mathbb{R}^2} f(x) \overline{\varphi_{j,l,k}(x)} \, dx
\]

(7)
And for Digital Curvelet Transforms, equation (7) will be change to:

\[ c^D(j, l, k) := \sum_{t_1, t_2} f[t_1, t_2] \psi^D_{j, k}[t_1, t_2] \]  

(8)

In this paper, Fast Discrete Curvelet Transform (FDCT) via wrapping method with 4 scales and 8 angles is applied to target images. We also modify the Curvelet coefficient using the following threshold suggested in [1] and [3] except the coarsest level coefficient. Using the modified coefficient value and powering the coarsest level coefficient by tenth (enhance the low resolution information), inverse fast discrete curvelet transform (IFDCT) via wrapping is performed. Please refer to [4] for further information about curvelet transform.

After acquiring the inverse transform result from each transformation mention above, some simple image enhancement methods should be performed to all of the image to enhance the quality of the extraction such as contrast stretching. Further image processing technique such as Otsu thresholding (to automatically select the threshold value from a greyscale image), canny edge detector (to omit the edge of nucleus), morphology filter, and finally apply the mask (figure 6) to target image. The results of this comparison are shown in figure 7.

![Transformation of Grayscale Image](image)

Fig. 4. Nucleolus candidate zone extraction block diagram.

4. Experimental Results

This experiment conducted in simulation environment, using computer specification: Intel® Core™ i3 CPU, 4GB of RAM, and Windows 7 32-bit Operating System.

![Fig. 5](image)

Fig. 5. (a) Lymphoblast sample image, (b) resulting mask from nucleus segmentation process, and (c) nucleus segmented result.

The authors use a sample image that representing lymphoblast cell group. Each image has a fixed resolution of 500 x 500 pixels and displayed in RGB color system as shown in figure 5(a). The image source are gathered from ALL-IDB 1 by Professor Fabio Scotti from Acute Lymphoblastic Leukemia Image Database for Image Processing under supervision of Professor Fabio Scotti from Department of Information Technology - Università degli Studi di Milano [9][12].

As seen in figure 5, the result gathered after applying the mask to the original image shown a quite accurate segmentation on the WBC nucleolus. Figure 6 shown the shape of extracted Region of Interest (RoI) mask where the nucleolus candidate zones most likely exist inside this boundary.

![Fig. 6](image)

Fig. 6. (a) Edge boundary mask extracted using Haar wavelet transform, (b) edge boundary mask extracted using Daubechies D4 Wavelet transform, and (c) Edge boundary mask extracted using Curvelet transform.

![Fig. 7](image)

Fig. 7. Normal cell image (a), performance comparison: using Haar wavelet (b), Daubechies D4 wavelet (c), and Curvelet transform (d).

Figure 7 shows qualitative performance evaluation while Hausdorff Dimension value shows quantitative performance evaluation for each transformation. Figure 7(a) shows a normal cell image which result shows no detected nucleolus candidate zone. In figure 7(b) we can see the result from Haar wavelet transform forming square shapes around the edge of candidate zone with Hausdorff dimension value of 0.8050, while figure 7(c) shown the result form Daubechies wavelet transform forming more flexible boundary compare to Haar wavelet with Hausdorff dimension value of 1.0513. Figure 7(d) shown the result from Curvelet transform forming even more flexible boundary which shaped like “islands” and precisely located the nucleolus candidate zones inside WBCs nucleus with Hausdorff dimension value of 1.1671.

All the results shown in figure 7 are confirmed based on ground truth method and this comparison confirm the fact that nucleolus
detection algorithm using Curvelet transform performs better than Daubechies and Haar wavelet, while the Hausdorff dimension value confirms that the larger value indicates more roughness in the perimeter of extracted candidate zone which means a better accuracy in detecting the nucleolus candidate zone.

The results shown in this experiment is very useful for further studies about leukemia cells classification algorithm. Curvelet transform could become an option as one of the feature extraction method to detect and classify the leukemic diseases.

5. Conclusions

The development of leukemia detection algorithm shows a good usage possibility and high expectation to help hematologist in performing a better diagnosis. Among various possible extraction methods, this paper use nucleolus candidate zone extraction method utilizing Wavelet and Curvelet transformations and shows a very good performance comparison.

The result shown that Curvelet transform had a better performance than Wavelet transform in a matter of accuracy to extract nucleolus candidate zone from peripheral blood smear image. The author realize that this extraction method is not good enough to form a fully automated leukemia classification system, there are various other factor that need to be developed and improved such as robustness against intensity level of the RGB images and better clustering methodology. The authors hope that this algorithm could be useful to improve the future works.

Acknowledgement

This work was supported under the framework of international cooperation program managed by National Research Foundation of Korea (2012K2A1A203297) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2011-0023118).

References

(9) ALL-IDB, http://homes.di.unimi.it/scotti/all/, last accessed on December 2013

Ivan Vincent received the B.S. degree in electrical engineering from Institut Teknologi Bandung, Indonesia, in 2012 and currently pursuing his master degree majoring in IT Convergence and Application Engineering in Pukyong National University, Korea.

Suk-Hwan Lee received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from Kyungpook National University, Korea in 1999, 2001, and 2004 respectively. He worked at Electronics and Telecommunications Research Institute in 2005. He is currently an Associate Professor in Department of Information Security at Tongmyong University, which he started in 2005.

Kwang-Seok Moon received the B.S., M.S., and Ph.D degrees in Electronics Engineering in Kyungpook National University, Korea in 1979, 1981, and 1989, respectively. He is currently a professor in department of Electronic engineering at Pukyong Nationa University. His research interests include digital image processing, video watermarking, and multimedia communication.

Ki–Ryong Kwon received the B.S., M.S., and Ph.D. degrees in Electronics Engineering from Kyungpook National University in 1986, 1990, and 1994 respectively. He worked at Hyundai Motor Company from 1986 to 1988 and at Pusan University of Foreign Language from 1996 to 2006. He is currently a Professor in Department of IT Convergence and Application Engineering at the Pukyong National University. He visited University of Minnesota in 2000–2001 and Colorado State University in 2011–2012 with visiting professor in USA. He is currently the General Affair and Organizing Committee/Vice President of Korea Multimedia Society, also is a Director of IEEE R10 Changwon section.