Feature-based Thresholds on Alpha Matting for Images for Natural Image Dataset

Dr.T. Velumani[1], Assistant Professor,
Department of Computer Science, Rathinam College of Arts and Science College (Autonomous), Coimbatore, TamilNadu, India.

***Abstract*- The aim of the research is to separate the foreground and background in natural images. The objects separation is performed by analysing the boundary area between foreground and background or the unknown. The analysis of unknown area is used to determine the threshold value to separate definitive foreground and background in alpha matting. The process begins with defining a sub-image of the grayscale image dataset with Region of Interest (ROI). Furthermore, the features of each sub-image consisting of contrast, correlation, energy and entropy are extracted using the Grey Level Co-Occurrence Matrix (GLCM) in angles of 0°, 45°, 90°, and 135°. Local extraction results are averaged and normalized and then, it is treated as a threshold for alpha matting. The result is evaluated using Peak Signal Noise to Ratio (PSNR) and shows a significant increase in performance.**

Keywords: alpha matting, threshold, region of interest and GLCM.

I. INTRODUCTION

In recent decades, researches related to the extraction of objects are massively performed as its accuracy will greatly affect the quality of the image or video editing. In addition, object extraction for image or video editing is very important in multimedia applications due to the dramatic increase of the growth of the network multimedia industry. In the beginning, Porter and Duff [1]introduce alpha channels as a function to control linear interpolation of foreground and background colours for anti-aliasing purposes when the foreground joins the background. This is called "pulling matte" or "digital matting" technique in object extraction.

"Pulling matte" is performed by combining the semi-transparent colours of the foreground with the background to produce a new blend colour. The degree of gradation of the foreground colour ranges from full black to white. Mixed colour will be foreground if it is full white and be the background for black. Mixed colours are the measured average of foreground and background colours. The accuracy of separation of foreground and background within the boundary of the object determines the success of its process.

A qualified matte extraction result should have an even colour distribution, neither too white nor black. If the pixel is too white, it will be dominant in the area correlated with extracted foreground, and it will be the contrary if it is too black which will be the background. The object separation of segmentation and matting are different in how to treat foreground and background pixels. In the hard segmentation technique, the process is firmly performed against the pixels, so that the pixel becomes part of the foreground or background only. In contrast to the matting technique, the unknown region pixel (α = alpha) becomes part of the foreground and background.

The withdrawal process of alpha values (alpha matting) is performed by differentiating the pixels as part of the foreground and background. Alpha matting is a convex combination of two colours allowing the transparency effect in computer graphics. Alpha values range is 0.0 - 1.0 with a full transparent value of 0.0 and full opaque value of 1.0, and the unknown region is determined by defining a threshold value.

Initially Levin et al[2]use the alpha threshold was defined as 0.17 - 0.15 assuming that the noise value in an image is within the range. Threshold definition is performed based on user perception by considering the characteristics of the image extracted, so certain expertise is needed in determining the threshold in order to get a qualified matte. User error in determining the threshold will affect the quality of matte. To overcome this problem, an adaptive threshold-based algorithm is proposed to determine the threshold value referring to image characteristics used as alpha threshold.

Computation of alpha channel based on global adaptive threshold is performed by using Fuzzy C-Means algorithm [3] and linear optimization [4]. However, the change in illumination causes certain parts to be brighter and darker on another. To overcome the problem, local adaptive threshold is applied by dividing the image into several sub-images, which then compute the threshold value by normalizing feature extraction. Determination of the threshold value is performed by calculating the Region of Interest (ROI) as a block based processing model of the image for initialization in determining the background and background areas. The purpose of selecting an ROI area is intended to divide the image into sub-images. Furthermore, the sub-image is used as a basis for analysing and testing extracted features from the selected image area. Threshold value is determined from the normalization of the average feature extraction value which consists of contrast, correlation, energy and entropy.

II. RELATED WORKS

Research related to alpha matting-based image segmentation has been carried out in recent decades. J. Wang and M.F. Cohen applies trimap[5][6][7][8][9][10]which is a pre-segmented image to distinguish the foreground, background and unknown areas. Limitation of the approach is misclassification of colour samples in complex scenes. Levin et al [11] define user-specific-constraints by applying manual scribble as sampling-sets to define foreground (F) and background (B) to overcome this problem, while unknown or alpha (α) is determined by the specified threshold value of 0 and 1. The use of trimap in separating the foreground and background areas results a visually near perfect[12]. However, the thickness level of manual scratches in the process of defining scribble requires a high level experience, especially in complex and complicated images such as the image of hair, feathers and falling snow [13].

The determination of threshold values in the range of 0-1 for alpha (α) values is adapted to Fuzzy C-Means by Basuki, et al [3]. The threshold value is obtained by calculating the average maximum value in the class having smallest and lowest middle value among others, where the pixel value uses three class concepts in Fuzzy C-Means. Thresholding in this method is calculated by collecting clusters in the areas having the same level of similarity and proximity to each other, by developing grey-level similarity of grey based on inter-class and intra-class so that the separation of background, foreground, and alpha is hoped to be more expressive. This technique is then used repetitively for segmenting semi-automatic video objects [14][15].

Before, threshold determination in images was globally calculated. P. Case and H. R. Rana use local thresholding for alpha matting to improve the level of threshold quality in image segmentation [16]. Combinations of image segmentation techniques are performed to obtain optimal results, including Edge and Region Based Segmentation. The Grey-Histogram Technique and Gradient-Based Method are used to define alpha matting of Edge Base Segmentation, and Thresholding method (Local and Global) and Region Operating are for Region Based Segmentation. From the experiments, it is concluded that each segmentation technique in image matting has both advantages and disadvantages that lie in the homogeneity of the natural image dataset used, spatial character structure, and image texture. Therefore, they propose KNN Matting because it can integrate both.

In general, the threshold is divided into two: global and local threshold. The problem of using global threshold is that there is a change in illumination compared to the same T value for the entire pixel. This will cause certain parts to be brighter and darker in others (for example, shadows of objects in the original image). However, these problems is overcome by local thresholding [17] adaptively applied to several techniques such as Niblack's Techniques, Sauvola's Technique, Bernsen's Technique, Yanowitz and Bruckstein's Method, and Maximum Entropy [18]. Local adaptive threshold is able to produce optimal values in the segmentation of an image because the threshold is calculated by dividing the image into new sub-images and not from the entire surface as one.

Based on the previous research, the alpha threshold will be calculated in the local adaptive threshold, where the maximum entropy threshold value obtained from the Region of Interest (ROI) area is normalized as the input of alpha threshold. The feature extraction from each sub-image of ROI produces a successful result for determining the threshold [19]. Feature extraction applied is the GLCM (Grey Level Co-Occurrence Matrix) which is a feature extraction method using statistical analysis of using grayscale images. In addition, GLCM is also able to examine textures by considering the spatial relationship of pixels in an image.

III. RESEARCH FRAMEWORK

The proposed feature-based object extraction system is illustrated in Fig. 1. The first step in the extraction process began with image acquisition as a source of data analysed. Each image in the RGB domain was transformed into the grayscale domain to simplify the computational process. The next grayscale image was cut and divided into 16 sub-image blocks using the Region of Interest (ROI) method. It was performed in order to locally calculate the feature.

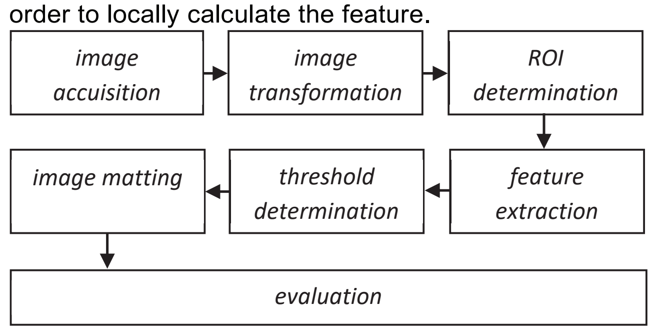

Fig. 1. Research framework

Feature extraction was performed over the sub-images generated from ROI by the parameters of contrast, correlation, energy and entropy in angles 0°, 45°, 90° and 135° using the GLCM (Grey Level of Co-Occurrence Matrix) method. The feature extraction results were then normalized and treated as a threshold value (α) in the image matting. The accuracy of the extracted object was evaluated using PSNR (Peak Signal Noise to Ratio) by comparing the matte produced by the system with the ground truth (matte reference from the dataset).

IV. DISCUSSION

The matte extraction test was performed using a public matting dataset consisting of teddy.bmp, kid.bmp, teddy_ear.bmp, fire.bmp and hair.bmp (as well as shown in Fig. 2.), in following stages:

- Image transformation: image conversion from RGB to Grayscale domain aimed to simplify the computational process. The image composition with 1 (one) channel would be simpler compared to 3 (three) channels with a value range of 0 - 255. The conversion process was performed, following the Equation (1).

$$GrayImage = 0.2989R + 0.5870G + 0.1140B \quad (1)$$

In which R is red, G is green and B is blue channel of the intensity value.In addition, the computational process was also simplified for the efficiency of processing time by considering the threshold value so that the intensity value became binary (0 and 1) so that it was able to be used in the extraction process.

- Region of Interest (ROI):Region of Interest is a part of an image identified for a particular purpose. In this research, the determination of ROI was performed using the block processing method [20] where the converted image in the grayscale domain was divided into 16 blocks of the same size as shown in Fig. 3. Furthermore, ROI was treated as a sub-image in which each block was labelled as initialization to distinguish the foreground and background regions.
- Feature extraction:GLCM (Gray Level Co-0ccurrence Matrix) was used to extract features which based on the two order texture calculation to calculate the relationship of pairs of two pixels in the original image. Sub-images were used as data sources tested. The GLCM features analyzed were: Contrast, Correlation, Energy, and Entropy. Feature extraction was performed by taking a window of 5x5 pixels from each sub-image. The window was taken from the upper left corner of each sub image as shown in Fig. 3 and calculated to determine the value of each feature as follows:

1) Contrast

Contrast was a measure of the existence of variations in the level of gray pixel images, calculated using Equation (2).

$$\textbf{Contrast} = \sum_{n=1}^{L} n^2 \left\{ \sum_{|i-j|=n} GLCM(i,j) \right\} \quad (2)$$

in which L was the number of levels used in computing, nwas the number of pixels, iwas the smallest pixel intensity, and jwas the largest pixel intensity.

2) Correlation

Correlation is a measure of linear dependency between gray levels in an image which is calculated using the following Equation (3).

$$\textbf{Correlation} = \frac{\sum_{i=1}^{L}\sum_{j=1}^{L}(ij)(GLCM(i,j) - \mu_i'\mu_j'}{\sigma_i'\sigma_j'} \quad (3)$$

in which L was the number of levels used in the computation proses, n was number of pixels, i was the smallest pixel intensity, and j was the largest pixel intensity. Then, $\sigma_i'\sigma_j'$was deviation standard for all pixel intensity in matrix GLCM.

3) Energy

Energy was the intensity measure of region area variation, calculated by the Equation (4).

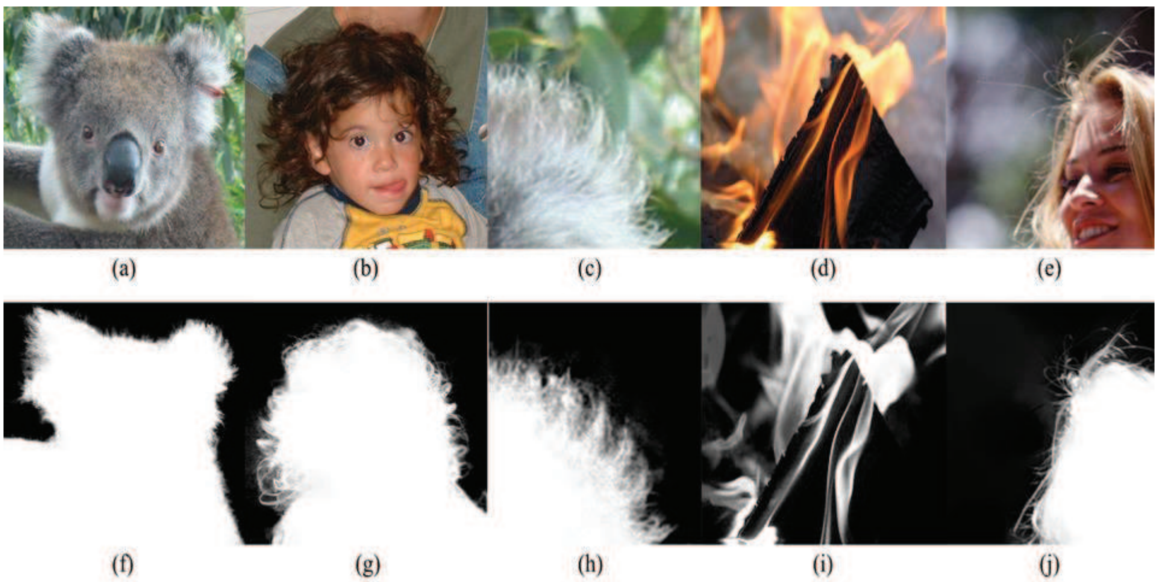

Fig. 2. Dataset of image matting, while figure (a) to (e) showing original image and figure (f) to (j) showing image reference

$$Energy = \sum_{i=0}^{G-1}\left[GLCM(i)\right]^2 \quad (4)$$

in which i was the pixel intensity average in matrix GLCM.

4) Entropy

Entropy was used to express the size of gray level irregularities in the image calculated by applying the Equation (5).

$$Entropy = -\sum_{i=1}^{L}\sum_{j=1}^{L}(GLCM\left(i,j\right)\log(GLCM(i,j) \quad (5)$$

Feature value determination of contrast, correlation, energy and entropy at rotation angles of 00, 450, 950 and 1350was applied in all windows in each sub image. The average value in eachfeature value was summed and divided by the number of features analyzed.

- Alpha matting: A qualified matte extraction result should have an even color distribution, neither too white nor black. If the pixel is too white, it will be dominant in the area correlated with extracted foreground, and it will be the contrary if it is too black which will be the background.The object separation of segmentation and matting were different in how to treat foreground and background pixels. In the hard segmentation technique, the process was firmly performed against the pixels, so that the pixel would be part of the foreground or background only. In matting technique, the unknown region pixel (α = alpha) would be part of the foreground and background. Thus, the threshold value in this area would decide the quality of separation result.The object separation process was performed by the assumption that dominant pixels with the white (α = 1) would correlate with foreground and black (α = 0) with background. The accuracy will meet the problem in the unknown region which is located at the edge of an object. Image matting operations were performed by specifying a threshold value between 0-1 to define the unknown region value.The threshold determination was previously performed by Levin and Lischinski where the alpha threshold was defined between$0.1^7 - 0.1^5$, assuming that the noise value in an image ranges within. Threshold definition was performed based on user perception by considering characteristics of the image extracted, so users need certain expertise in determining the threshold in order to obtain the best result. User error in determining the threshold will affect the quality of matte [2]. Basuki et al, propose an adaptive threshold-based algorithm by applying Fuzzy C-Means to determine threshold

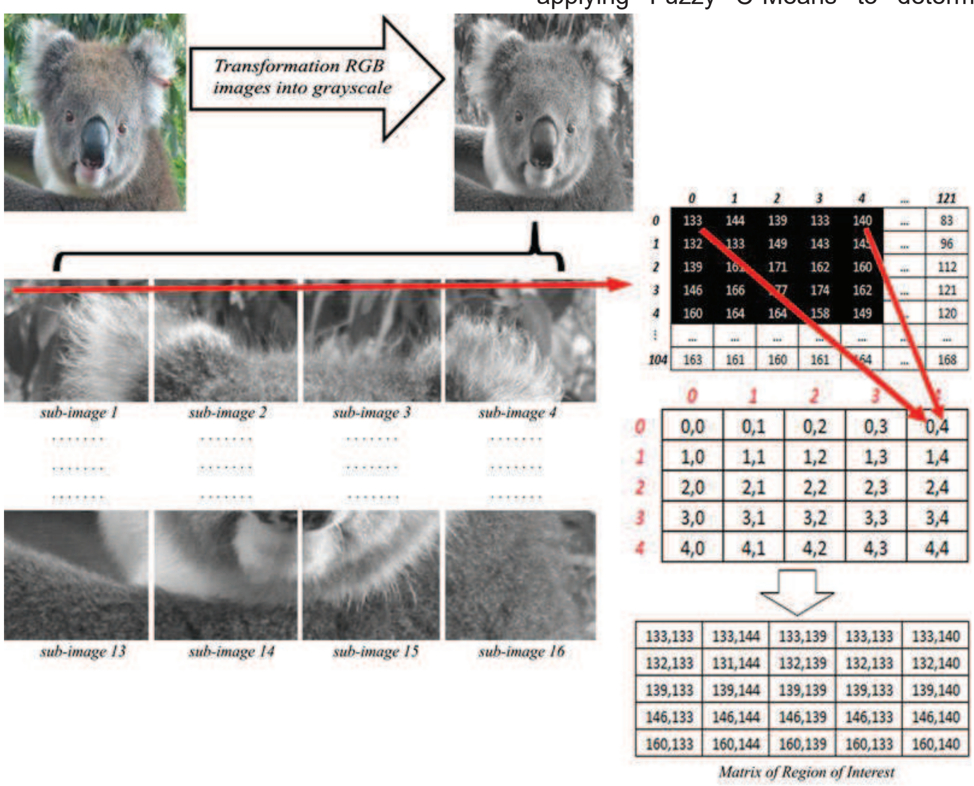

Fig. 3. Image transformation and ROI determination

values referring to the image characteristics used as thresholds in the unknown region to overcome this problem [3][21].In this research, the threshold value was calculated in each sub-image resulting from the local determination of ROI. The average value of feature extraction from GLCM (contrast, correlation, energy and entropy) in each sub-image (Equation 2 - 5) was normalized by dividing the entropy value in each sub-image by the average number of GLCM values in an image as I as shown in Equation (6).

$$Threshold = \frac{Average\ (Contras_s + Correlation_s + Energy_s + Entropy_s)}{\sum average image\ feature} \quad (6)$$

in which s was feature value on sub-image. The result of the Equation (6) was treated as threshold value for alpha matting [11]as input for Equation (7).

- Experiment result and evaluation: Testing of the image matting dataset was performed by input images in the form of original, terrible and reference images as a reference for testing comparisons between proposed method and intended results. Each image (teddy.bmp, teddy_ear.bmp, kid.bmp, fire.bmp and hair.bmp) was converted into gray image and divided into 16 blocks of sub-image. Each sub-image is taken by a 5x5 pixel window which was calculated by GLCM to obtain the

feature value and normalized to the threshold value treated as an input alpha value (α) in the matting image as shown in the Equation (7).

$$I_i = \alpha_i F_i + (1 - \alpha_i) B_i \quad (7)$$

in which $F_i$was definitive foreground value, B_i was definitive background and α_i was unknown region.

Then, matte was extracted using closed-form matting [3][11][14][15]. The result was replaced by original image in order to obtain object desired as described in Fig. 4.The trial result conducted on each of the analyzed images showed that the proposed method had a significant increase in the performance compared to closed-form solution with FCM as alpha threshold [3],[14],[15]. Evaluation of each image was performed using

PSNR (Peak Signal Noise to Ratio) as in the Equation 9 and the results are shown in Fig. 5.

$$PSNR = 10 log_{10} \left[\frac{255^2}{MSE} \right] \quad (8)$$

in which

Fig. 4. Extraction stage : (a). original image, (b). scribble image, (c). matte reference from dataset, (d). matte extraction from the proposed approach, and (e). object extraction the result from the proposed approach

$$MSE = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} \left(\frac{abs(grdImg - matteExt)}{M \times N} \right)^2 \quad (9)$$

With $grdImg$ was the ground truth image as the reference image, $matteExt$ was a matte extracted by the system and $M \times N$ was the size of the executable image.

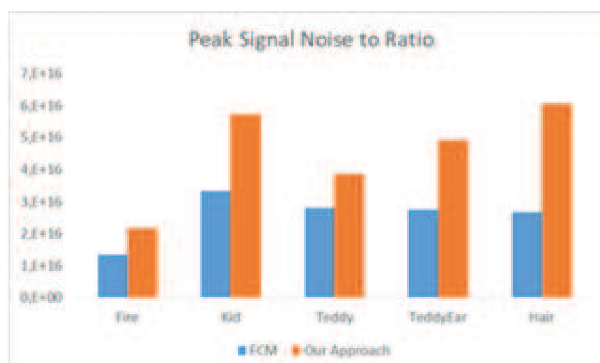

Fig. 5. Performance evaluation using PSNR

V. CONCLUSION

Alpha matting based of object extraction was tested on 5 images from the matting image dataset. Initially, the image was transformed into the grayscale domain and divided into 16 sub-images and using ROI for each. Each sub-image feature was extracted (local threshold) using GLCM with contrast, correlation, energy and entropy parameters, then results were normalized and treated as threshold values in alpha matting.

The trial results showed that the object extraction process with a feature-based threshold for natural images worked well in extracting matte. The visually matte which correlates with foreground pixels was able to correlate more accurate. In addition, an increase in accuracy was quantitatively shown by the results of evaluations using Peak Signal Noise to Ratio (PSNR) which showed an average increase up to 63% from the previous [3][14][15][21].

REFERENCES

[1] T. Porter and T. Duff, "Compositing Digital Images," *Computer Graphics,* vol. 18, no. 3, pp. 253 - 259, 1984.

[2] A. Levin, A. Rav-Acha and D. Lischinski, "Spectral Matting," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 30, pp. 1699 -1712, 2008.

[3] R. S. Basuki, M. Hariadi and R. A. Pramunendar, "Fuzzy C-Means Algorithm for Adaptive Threshold on Alpha Matting," in *Conference on Information Technology and Electrical Engineering*, Yogyakarta, 2012.

[4] S. Chen, Z. Li, J. Liu and X. Tang, "Image Matting Using Linear Optimization," in *Proc. 15th International Conference Multimedia*, 2007.

[5] J. Wang and M. F. Cohen, "Optimized Color Sampling for Robust Matting," in *Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, 2007.

[6] N. E. Apostoloff and A. W. Fitzgibbon, "Bayesian Video Matting Using Learn Image Priors," in *Proc. Fifth LASTED International Conference Visualization Imaging and Image Processing*, 2004.

[7] J. Sun, J. Jia, C. Tang and H. Shum, "Poisson Matting," *ACM Transaction on Graphics,* vol. 23, no. 3, pp. 315 - 321, 2004.

[8] Y. Y. Chuang, A. Agarwala, B. Curless, D. H. Salesin and R. Szeliski, "Video Matting on Complex Scenes," *ACM Transaction on Graphics,* vol. 21, no. 3, pp. 243 - 248, 2001.

[9] Y. Y. Chuang, B. Curless, D. H. Salesin and R. Szeliski, "A Bayesian Approach to Digital Matting," in *Proc. IEEE Conference Computer Vision and Pattern Recognition*, 2001.

[10] M. A. Ruzon and C. Tomasi, "Alpha Estimation in Natural Images," in *Proc. IEEE Conference Computer Vision and Pattern Recognition*, 2000.

[11] A. Levin, D. Lischinski and Y. Weiss, "A Closed-Form Solution to Natural Image Matting," *IEEE Transaction on Pattern Analysis and Machine Intelligence,* vol. 30, no. 1, pp. 228 - 242, 2008.

[12] J. Wang and M. F. Cohen, "Image and Video Matting : A Survey," *Foundation Trends Computer Graphics and Visual,* vol. 3, pp. 97 - 175, 2007.

[13] H. Liu, L. Ma, X. Cai, Z. Chen and Y. Shen, "A Closed-Form to Video Matting of Natural Snow," *Information Processing Letter,* vol. 109, no. 18, pp. 1097 - 1104, 2009.

[14] R. S. Basuki, M. A. Soeleman, R. A. Pramunendar, A. F. Yogananti and C. Supriyanto, "Video Object Segmentation Aplying Spectral Analysis and Background Subtraction," *Journal of Theoretical and Applied Information Technology,* vol. 72, no. 2, pp. 208 - 214, 2015.

[15] R. S. Basuki, M. Hariadi, E. M. Yuniarno and M. H. Purnomo, "Spectral-Based Temporal-Constraint Estimation for Semi-Automatic Video Object Segmentation," *International Review on Computer and Software,* vol. 10, no. 9, pp. 959 - 965, 2015.

[16] H. R. Rana and S. Kumar, "Color Tree Image Extraction Using KNN Matting," *International Journal of Advanced Research in Computer Engineering and Technology,* vol. 2, no. 6, pp. 57 - 62, 2014.

[17] H. J. Vala and A. Baxi, "A Review on Otsu Image Segmentation," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET),* vol. 2, no. 2, pp. 387 - 389, 2013.

[18] N. Senthilkumaran and S. Vaithegi, "Image Segmentation by Using Thresholding Techniques fo Medical Images," *Computer Science & Engineering : An International Journal,* vol. 6, no. 1, pp. 1 - 13, 2016.

[19] N. Zayed and H. A. Elnemr, "Statistical Analysis of Haralick Texture Features to Discriminate Lung Abnormalities," *Journal of Biomedical Imaging,* vol. 12, 2015.

[20] G. R. Jothilakshmi, J. R. Christilda, A. Raaza, Y. SreenivasaVarma and V. Rajendran, "Extracting Region of Interest Using Distinct Block Processing Method in Sono-mammogram Images," in *International Conference on Computer, Communication and Signal Processing (ICCCSP)*, Chennai, India, 2017.

[21] R. S. Basuki, M. A. Soeleman, M. Hariadi, M. H. Purnomo, R. A. Pramunendar and A. F. Yogananti, "Spectral-Based Video Object Segmentation Using Alpha Matting and Background Subtraction," in *Fourth IIEEJ International Workshop on Image Electronic and Visual Computing*, Koh Samui, Thailand, 2014.