VOL. 11, NO. 6, MARCH 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

AUTOMATED THRESHOLD DETECTION FOR OBJECT SEGMENTATION IN COLOUR IMAGE

Md. Akhtaruzzaman, Amir A. Shafie and Md. Raisuddin Khan
Department of Mechatronics Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia
E-Mail: akhter900@gmail.com

ABSTRACT

Object segmentation from single background colour image is important for motion analysis, object tracking, trajectory identification, and human gait analysis. It is a challenging job to extract an object from single background colour image because of the variations of colours and light intensity. Most common solution of the task is the uses of threshold strategy based on trial and error method. As the method is not automated, it is time consuming and sometimes a single threshold value does not work for a series of image frames of video data. In solving this issue, this paper presents an Automated Threshold Detection Algorithm, $H(\bullet)$. The algorithm is applied in segmenting human lower limbs from a series of image frames of human walking. The procedure starts with selection of optimal RGB channel. Then $H(\bullet)$ algorithm is applied for automated threshold detection to convert the image frames into grayscale image. In the next stage, Line Fill (LF) algorithm is applied for smoothing the edges of object and finally background is subtracted to extract the targeted object. Results of the applied procedure show that the algorithm is viable to extract object from single background color image and can be used in human gait analysis.

Keywords: object segmentation, image processing, automated threshold detection, RGB channel, gait analysis.

INTRODUCTION

Background subtraction is one of the common methods in segmenting moving objects from a sequence of image frame where a reference frame, mostly called Background Model, is subtracted from temporal images [1] [2]. Object segmentation from a certain image is really a challenging job as there exist variations of colours on the focused subject, complex pattern of background, light intensity on focused area, shadow of targeted subject, and variations in camera angle positioning [3] [4]. So, particular image properties are needed to be analyzed carefully to model an image frame to extract targeted object [5].

Human gait analysis is one of the interesting areas of research [6] where object segmentation strategy facilitates a wide range of opportunity. Saboune and Charpillet [7] presents foreground segmentation strategy where they have constructed silhouette image by applying threshold filter strategy and subtracting background (pixel by pixel) from current image. The main purpose of the study was to extract gait parameters to detect balance disorder and fall-tendency in elderly. Another study presents $2D$ marker-less gait analysis where threshold filtering method is used in foreground segmentation from colour image [8].

Common practice of threshold approach is consideration of only the intensity, not any relationship between or among the pixels. Most of the cases threshold value for a particular image frame is selected based on trial and error method. As the solution of the problem, this paper presents an Automated Threshold Detection Algorithm for foreground segmentation.

ORGANIZATION OF THE PAPER

Firstly, the paper presents experimental setup where image capture and boundary settings are introduced. Secondly, procedure for selection of optimum RGB channel is presented. Thirdly, greyscale conversion and Automated Threshold Detection Algorithm is demonstrated. Fourthly, implementation of LF algorithm is explained. Fifthly, overall algorithm block diagram and results are presented. Finally, the paper draws concluding remarks.

EXPERIMENTAL SETUP

As the project focuses on human lower limbs motion, original image frames are cropped to represent only lower extremity movement characteristics for forward walking. Video stream is taken using Samsung Galaxy Note II (GT-N7100) 8Mp camera with resolution of $(\text{width (w)} * \text{height(h)}) = (1280 * 720)$ and the image frames are cropped with resolution $(1010 * 471)$ or ≈ 0.4 Mp as presented in Figure-1.

Figure-1. Cropped image $(1010 * 471)$.

In Figure-1, focused subject (foreground) has distinct colour over blue background. Variations of light

intensity in setup environment effect background colour distributed with some gradient features. Gradient features are also varies depending on various positions of foreground over the background. Moreover, subject's shadow changes the luminosity of some particular areas of image frame. Image frames also contain some black regions most commonly in between lower limbs.

An RGB colour image frame can be represented by Equation. (1), where I presents a single image frame; i indicates series of pixels $(i = 1,2,3,\dots,n)$, where $n = (w * h)$ of the image frame; (x, y) presents horizontal and vertical positioning of image, $x = 1,2,3,\dots,w$ and $y = 1,2,3,\dots,h$. R, G, and B reflect Red Channel, Green Channel, and Blue Channel respectively.

$$I = I_i = I_{(x,y)} = I_{(R_i G_i B_i)} = I_{\left(R_{(x,y)} G_{(x,y)} B_{(x,y)}\right)} \quad (1)$$

RGB CHANNEL SELECTION

Purpose of RGB channel selection is to identify best possible channel that represents the foreground with maximum information. Figure-2 presents three channels distributions of a particular image frame with $2D$ as well as $3D$ front and rear views.

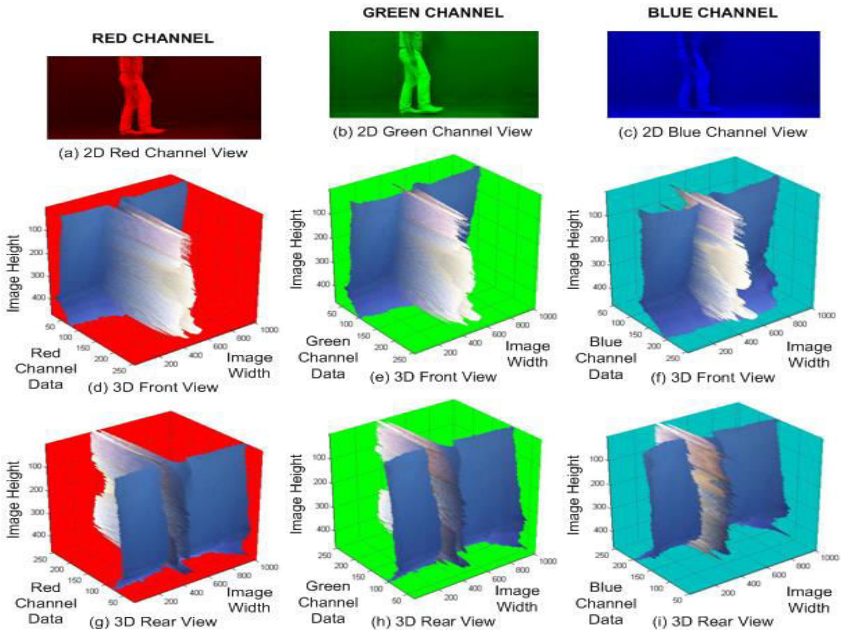

Figure-2. RGB channel distributions of a particular image frame and 3D representations of each channel information with front and rear view of the image frame.

Blue surfaces in the front view presentation of $3D$ channel data are representing backgrounds of input image. White rising portions are representing foreground of input image. By analyzing these three front views of three channels, it can be reported that Blue Channel data, in Figure-2(f), loses most of the information of foreground. Although, Green Channel data representation in Figure-2(e) shows better result than the Blue Channel, the best foreground data are presented by Red Channel with minimum loose of information as shown in Figure-2(d). Same result can be established through analyzing rear view representations of three channel data as demonstrated in Figure-2(g, h, i). The three channel representations of input image can be expressed by Equation. (2), where E represents entire pixel distributions of an image I and I_R, I_G, I_B represent Red, Green, Blue channel images respectively.

$$\left.\begin{aligned} I_R &= \mathrm{E}_{\mathrm{x=1,y=1}}^{\mathrm{x=w,y=h}}\left(\mathrm{I}_{(\mathrm{R}_{(\mathrm{x,y})},0,0)}\right) \\ I_G &= \mathrm{E}_{\mathrm{x=1,y=1}}^{\mathrm{x=w,y=h}}\left(\mathrm{I}_{(0,\mathrm{G}_{(\mathrm{x,y})},0)}\right) \\ I_B &= \mathrm{E}_{\mathrm{x=1,y=1}}^{\mathrm{x=w,y=h}}\left(\mathrm{I}_{(0,0,\mathrm{B}_{(\mathrm{x,y})})}\right) \end{aligned}\right\} \quad (2)$$

Though Red Channel shows better performance, this could be varied for various positions of foreground over the background. Therefore it is necessary to make the system automatic to select best channel. As the background occupies most of the area rather than the foreground, it is possible to select a range values (a and b) for background of particular channel data, where a indicates lowest and b indicates highest margin of background surface (similar strategy is explained latter in Equation. (5) to Equation. (9)). After determining background margin values, percent of Background Surface Volume (BSV) (pixels within margin values), Lower Pixel Volume (LPV) (pixels lower then a), and Higher Pixel Volume (HPV) (pixels higher than b) are calculated and compared. The larger percent value for HPV of a

particular channel comparing with other is considered as the best channel. This relation can be presented by Equation. (3), where X, Y, and Z illustrate dimensions of $3D$ representation of each channel data. Selection of best channel is implemented by using the rules shown in Equation. (4).

$$\mathrm{LPV}_{(X_{[x,y]},Y_{[a_1=0;b_1<a_2]},Z_{[x,y]})} < \mathrm{BSV}_{(X_{[x,y]},Y_{[a_2,b_2]},Z_{[x,y]})} < \mathrm{HPV}_{(X_{[x,y]},Y_{[a_3>b_2,b_3=255]},Z_{[x,y]})} \quad (3)$$

$$Cnl = \begin{cases} Red & if\ \%\ of\ R_{LPV} < G_{LPV}\ and\ B_{LPV} \\ Green & if\ \%\ of\ G_{LPV} < R_{LPV}\ and\ B_{LPV} \\ Blue & if\ \%\ of\ B_{LPV} < R_{LPV}\ and\ G_{LPV} \end{cases} \quad (4)$$

GREYSCALE CONVERSION

Selected channel image is converted into greyscale image (I_{grey}) using MatLab function, $rgb2gray(I)$. Figure-3 demonstrates greyscale image with its $3D$ surface data representation.

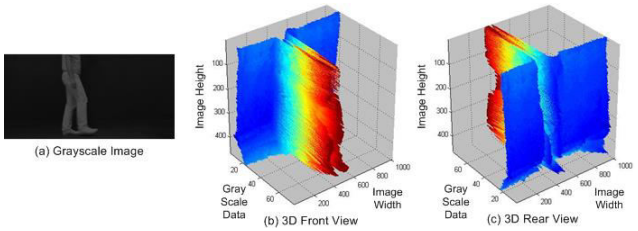

Figure-3. Greyscale image of a particular frame and $3D$ representation of grey level data.

AUTOMATED THRESHOLD DETECTION

Threshold value (Th) is calculated based on Equation. (5) to Equation. (9). Here E indicates all possible values for x (width), y (greyscale values), and z (height) axes of $3D$ representation of Greyscale Image (I_{grey}); $H(\bullet)$ is a function that returns array of position pointers ($p_{(y)}$). Margin values, a and b, calculation are shown in Equation. (6) where Mn and Mx are indicating lowest and highest margin range. Pointer position update factor, δ_1, is calculated based on Equation. (7) [9] where $p_{1(y)}$, $p_{2(y)}$, and δ are presenting current pointer position, new pixel position, and y distance between two positions respectively. Pixel mass property (p_m) for each updated pointer increases by σ, where $\sigma = (1\ or\ 0)$ as shown in Equation. (8). Optimum values for Mn and Mx could be selected as half of the maximum depth of greyscale data.

$$E_{x=1,z=1}^{x=w,z=h} p_{(y)_{(z,x)}} = E_{x=1,y=a,z=1}^{x=w,y=b,z=h} \left(H\left((p_{(y)} + \delta_1), (p_m + \sigma) \right)_{(z,x)} \right) \quad (5)$$

$$a = p_{new_{(y)}} - Mn; b = p_{new_{(y)}} + Mx \quad (6)$$

$$\delta_1 = \frac{p_{2(y)} * (\mp\delta)}{(p_{1(y)} + p_{2(y)})} \quad (7)$$

$$\sigma = \begin{cases} 1 & if \quad \delta_1 \neq 0 \\ 0 & if \quad \delta_1 = 0 \end{cases} \quad (8)$$

$$Th = \frac{\sum_1^h max\left(\left(p_{(y)} \right)_1^w \right)}{h} \quad (9)$$

A conceptual model of $H(\bullet)$ function is presented in Figure-4(a), where $p_{(y)}$ is pointer head; a, and b are searching ranges of next pixel. An example pixel model of greyscale depth information for a particular row of I_{grey} is presented in Figure-4(b). As the model presents depth information, there exists only one pixel for each column. At initial condition, pixel mass properties (p_m) are assigned as natural series starting from 1 and incremented by 1 for each existing pixel position of a particular row of y. All the other empty pixels are set to 0. In Figure-4(b), δ is y distance between p_1 and p_2 where p_1 actually the initial pointer position, $p_{(y)}$, at first column and p_2 is just a pixel position at second column. δ_1 is also y distance between p_1 and p'_1 where p'_1 is the second pointer position of second iteration at column 2. For each iteration, pointer position on X and Y axes are updated by 1 and δ_1. Pointer mass property is updated based on σ.

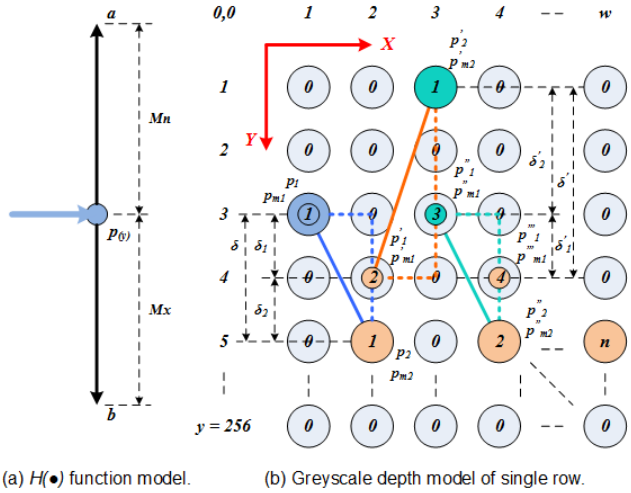

Figure-4. Function, $H(\bullet)$, model applied on greyscale depth model of a single row from I_{grey}.

Figure-5 shows the examples of pointer variations (threshold level variations) for depth information graphs of three different rows (row 30, row 60, and row 90) in I_{grey} comparing with three different values (30, 50, and 70) for Mx and Mn. It is observed that,

all values (≥ 50) for Mx and Mn produce nearly same maximum values of pointer level variations.

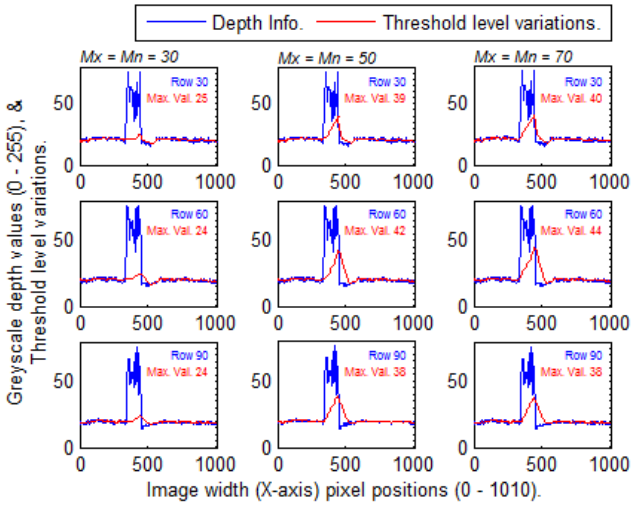

Figure-5. Threshold level variations over greyscale depth information for range values 30, 50, and 70.

APPLYING LINE FILL (*LF*) ALGORITHM

For smooth distribution of foreground region (white region) of Threshold Image (I_{th}) shown in Figure-6(a, b), Line Fill (LF) algorithm is applied in two steps. First, applying Horizontal Line Fill (HLF) and second, applying Vertical Line Fill (VLF) as presented in Figure-6 (c, d). The algorithm search for minimum existing pixel rows (l_{min}) of background and fills with white colour if $l_{min} \leq l_{th}$. Threshold length (l_{th}) is chosen carefully as small as possible, because larger value of l_{th} could reduce or lose foreground information.

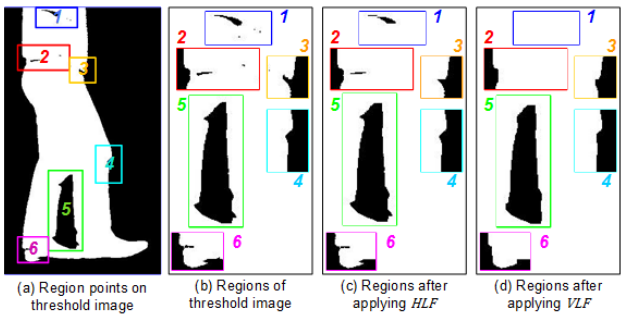

Figure-6. Regions of threshold image and results after applying HLF and VLF algorithm.

Output of LF algorithm concludes that the algorithm not only fills the expected regions, but also smoothen the edges. The algorithm (HLF and VLF) can be presented by Equation. (10) and Equation. (11). Here E stands for all possible values of y (image height 1 to h) or x (image width 1 to w). Function $minlength(\bullet)$ determines minimum length (l_{min}) with start point (s) and end point (e) of a particular axis line of background. p_x and p_y present pixels of a single line for X and Y axes. l_{th} is threshold line length.

$$E_{y=1}^{y=h}([l_{min}]_s^e)_y = E_{y=1}^{y=h}(minlength([p_x]_s^e))_y$$
$$E_{y=1}^{y=h}([p_x]_s^e)_y = \begin{cases} 255 & if\ l_{min} \leq l_{th} \\ p_x & otherwise \end{cases} \quad (10)$$

$$E_{x=1}^{x=w}([l_{min}]_s^e)_x = E_{x=1}^{x=w}\left(minlength\left([p_y]_s^e\right)\right)_x$$
$$E_{x=1}^{x=w}\left([p_y]_s^e\right)_x = \begin{cases} 255 & if\ l_{min} \leq l_{th} \\ p_y & otherwise \end{cases} \quad (11)$$

ALGORITHM BLOCK DIAGRAM AND RESULTS

Foreground region is determined based on the output of FL algorithm and the region is extracted from original color image. The whole process can be presented by block diagram as shown in Figure-7.

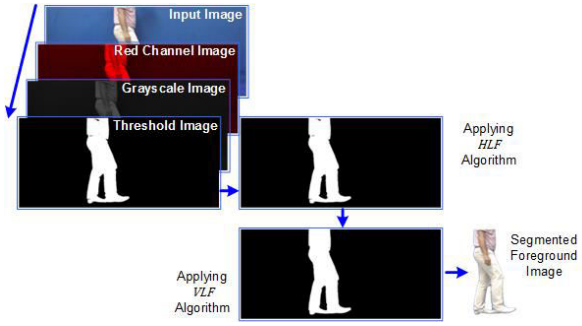

Figure-7. Block diagram of foreground segmentation process from colour image.

Segmented foregrounds of series of image frames for walking sequence are presented bellow. Here Figure-8(a) shows segmented objects for gait initiation sequence (1-8) and Figure-8(b) shows segmented objects for gait continuation sequence with right leg stance phase (1-10) pattern and left leg swing phase (3-8) pattern.

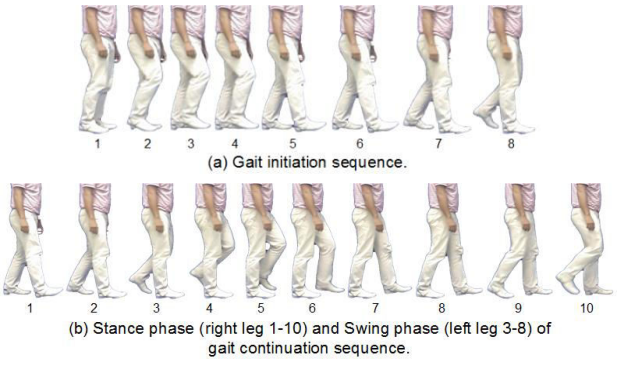

Figure-8. Segmented objects as outputs of the algorithm are presenting, (a) Gait Initiation (GI) sequence, and (b) stance and swing phases of right and left leg for Continuous Gait (CG).

CONCLUSIONS

Purpose of this study is to identify human gait sequence from a series of image frames. The study presents object segmentation strategy from colour image through automated threshold selection procedure. Results show that the procedure could be used in characterizing human walking pattern and recognition of gait. Joint motion trajectories and angle variations may possible to represent which could lead in designing and planning of gait for robotic or therapeutic systems. Automatic gait recognition and person identification for video surveillance system is another domain of research where this strategy could be adopted.

ACKNOWLEDGEMENT

Authors would like to express their appreciation to the Ministry of Higher Education (MOHE), Malaysia, in funding the project through the Fundamental Research Grant Scheme (FRGS). They also would like to express their gratitude to Research Management Centre (RMC), International Islamic University Malaysia (IIUM), Kuala Lumpur, Malaysia.

REFERENCES

[1] Piccardi M. 2004. Background subtraction techniques: a review. 2004 IEEE International Conference on Systems, Man and Cybernetics, 10-13 Oct. 2004, Vol. 4, pp. 3099-3104.

[2] Horprasert T., Harwood D. and Davis L.S. 2000. A Robust Background Subtraction and Shadow Detection. In Proc. Asian Conf. on Comp. Vision, 8 Jan. 2000, pp. 983- 988.

[3] Lee L. and Grimson W. E. L. 2002. Gait Analysis for Recognition and Classification. Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR 2002), Washington, DC, USA, 20-21 May, 2002, pp. 148-155.

[4] Rosenhahn B., Kersting U. G., Smith A. W., Gurney J. K., Brox T. and Klette R. 2005. A System for Marker-Less Human Motion Estimation. Proceedings of 27th DAGM Symposium, Vienna, Austria, August 31 - September 2, 2005, pp. 230-237.

[5] H. Peng, Y. Yang, J. Zhang, X. Huang, J. Wang. 2014. A Region-based Color Image Segmentation Method Based on P Systems. Romanian Journal of Information Science and Technology. Vol. 17, No 1, 2014, 63-75.

[6] Akhtaruzzaman M. and Shafie A. A., 2011. Joint Demeanors of an Anthropomorphic Robot in Designing the Novel Walking Gait. The 8th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI 2011), Nov. 23-26, 2011, Songdo ConventiA, Incheon, Korea, pp. 563-567.

[7] Saboune J. and Charpillet F. 2005. Markerless Human Motion Capture for Gait Analysis. Proceedings of the 3rd European medical and biological engineering conference 2005, Prague.

[8] Goffredo M., Carter J. N. and Nixon M. S. 2009. 2D Markerless Gait Analysis. 4th European Conference of the International Federation for Medical and Biological Engineering IFMBE Proceedings. Vol. 22, 2009, pp. 67-71.

[9] Akhtaruzzaman M. and Shafie A. A., 2011. A Novel Gait for Toddler Biped and its Control Using PIC 16F877A. 2011 4th International Conference on Mechatronics (ICOM'11), 17-19 May 2011, Kuala Lumpur, Malaysia.