ADAPTIVE MAXIMUM GLOBAL CLUSTERING BASED ON COARSE-FINE SCHEME FOR HUMAN IMAGES

Renu. D. S[*]

Mrs. K. S. Angel Viji[**]

Abstract

The main goal of segmentation is to partition an image into regions. Region-based segmentation is a technique for determining the region directly. Human segmentation in photo images is a challenging and important problem that finds numerous applications ranging from album making and photo classification to image retrieval. Previous works on human segmentation usually demand a time-consuming training phase for complex shape-matching processes. In this project, we propose a new method of segmenting an image into several sets of pixels with similar intensity values called regions. In this project, we propose a straightforward framework to automatically recover human bodies from color photos. Our method is made up of two procedures. First, we develop the adaptive global maximum Clustering .

In this procedure, we deal with an image histogram and automatically obtain the number of significant local maxima of the histogram. This number indicates the number of different regions in the image. Second, we detect a coarse torso (CT) using the multi cue CT detection algorithm and then extract the accurate region of the upper body. Then, an iterative multiple oblique histogram algorithm is presented to accurately recover the lower body based on human kinematics. The performance of our algorithm is high compared to conventional methods.

Keywords- Graph cuts, Adaptive global maximum clustering, haar cascades, multicue coarse torso detection

[*] PG Scholar, Department of computer science and engineering, Noorul Islam Center for Higher Education,Thuckalay

[**] Assistant Professor, Department of computer science and engineering, Noorul Islam Center for Higher Education,Thuckalay

I. INTRODUCTION

The project is concerned with the automatic recovery of human body regions with a detected face, with high segmentation accuracy. Automatic human body detection is a key enabler for applications in robotics, surveillance and Intelligent Transport System. In conventional methods, human body detection however suffers from variation of images such as clothing, lighting and shape morphing. Automatic human detection and body part localization are important and challenging problems in computer vision. The solution to these problems can be employed in a wide range of applications such as safe robot navigation, visual surveillance, human-computer interface, and performance measurement for athletes [9] and patients with disabilities, virtual reality, figure animation and also for search and rescue missions. However, retrieval by shape is still considered one of the most difficult aspects of content-based search since human bodies have articulated parts and deformable shapes. Therefore the research on human detection and body part localization is very active and it has produced a wide range of applications on general object detection and shape analysis.
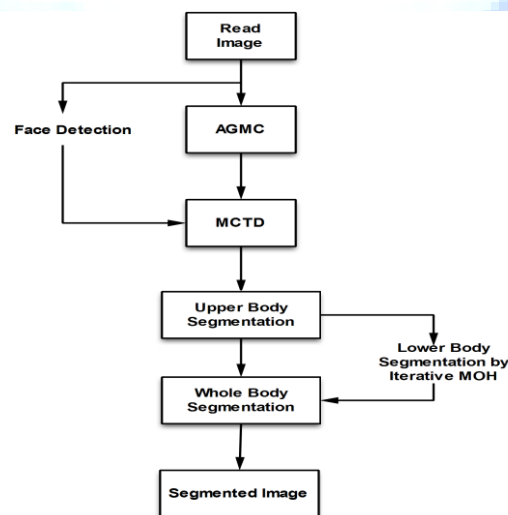

Figure 1: Block diagram of the proposed system

In this project, human body is recovered from photo images via integrating top down [3] body information and low level cues into graph cuts framework. A coarse to fine strategy is employed. The block diagram of the proposed scheme is given in Fig. 1. Given a photo image,

adaptive global maximum clustering is used to identify regions. A face detection method is used to locate the human face. The whole body extraction is subdivided in to two tasks, upper body segmentation and lower body segmentation.

The multicue coarse torso detection algorithm (MCTD) is utilized to segment the upper body that adjoins to head, in which the normalized cuts and global probability of boundary are effectively combined. The lower body is segmented based on iterative multiple oblique histogram.

II. ADAPTIVE GLOBAL MAXIMUM CLUSTERING

In this method, according to kim and Kang [2] histogram of the image is obtained, from which the number of significant local maxima is automatically obtained. Then 2-means clustering is applied over the local maxima obtained from the histogram. A random cluster number is assigned and based on that k means clustering is applied. The constraint values are checked for. Iteration is performed until efficient clusters are obtained. The goal of this process is

$$[1,L] = \left(\bigcup_{i=1}^{n} I_i \right) \bigcup R \qquad (1)$$

where each subinterval $I_i = [a_i, b_i]$ is a cluster containing the ith adaptive global maximum, which is the global maximum of the histogram. R is the domain of the (n+1)th histogram such that $\max_{l \in R} h(l)$ is very small compared with the original histogram, R or is empty. Such very small histogram values are usually useless in the detection of regions. To find a cluster that is a subinterval with an adaptive global maximum at each iteration, we fix k=2 and repetitively implement the standard k-means clustering under the rules described below. The standard k-means clustering method is a process to solve the following minimization problem.

$$\arg_{I_{1,\ldots\ldots I_k}} \min \sum_{j=1}^{k} \sum_{x \in I_j} \left| x - d_j \right|^2 \qquad (2)$$

where $\left| x - d_j \right|^2$ is a distance between a data point x and the jth cluster center d_j. The algorithm consists of a simple re-estimation procedure as follows. Initially, the data points are assigned at random to the k sets. For step 1, the centroid is computed for each set. In step 2, every point is assigned to the cluster whose centroid is closest to that point. These two steps are alternated until

a stopping criterion is met, i.e., when there is no further change in the assignment of the data points.

III. FACE DETECTION

Face detection is concerned with finding whether or not there are any faces in a given image (usually in gray scale) and, if present, return the image location and content of each face, a fully automatic system that analyzes the information contained in faces. Haar cascade features are is used. This approach to detecting objects in images combines four key concepts:

- Simple rectangular features, called Haar features.
- An Integral Image for rapid feature detection.
- Haar cascade machine-learning method.
- A cascaded classifier to combine many features efficiently.

The features used are based on Haar wavelets. Haar wavelets are single wavelength square waves (one high interval and one low interval). In two dimensions, a square wave is a pair of adjacent rectangles - one light and one dark.

Rectangle features can be computed very rapidly using an intermediate representation for the image called as the integral image [7]. The integral image at any location x, y contains the sum of the pixels above and to the left of x, y.

After an integral image is represented, haar cascades are used to select a small number of visual features from a very large set of possible features. Haar cascades for frontal face detection is used to analyse the features of the face if present in the given image. Following this is a method for combining classifiers [7] in a cascade which allows background regions of the image to be quickly discarded while spending more computation on promising face like regions.

IV. UPPER BODY SEGMENTATION

A. Coarse Torso detection algorithm:

Torso represents the human body without head and limbs. It includes the chest and abdomen. Coarse Torso means not fine in appearance, here it represents a rough appearance of torso.

The multi cue coarse torso detection algorithm is used to segment the upper body that adjoins to head, in which the normalized cuts [4], [5] and global probability of boundary are effectively combined. After the face is detected the center of the head is identified. Then based on the center of the head region , according to the height and width of the head an appropriate

torso region is recognized,based on which the rest of the segmentation follows. Normalized Cuts segments are grouped into a torso candidate region based on the bounding box along different orientations where the bounding boxes are generated according to face priori[1]. In the combining procedure, three cues are employed to select the coarse torso: area probability, location probability, contour probability [1], [8].

According to Fang et.al [1] based on the above cues, the coarse torso can be detected with the MCTD algorithm. Given a bounding box region, all segments that are overlapped with the bounding box region is found out without considering the head region. For each such segments, the area and location probability is computed. The closer to the center of the bounding box region, the segment is likely to be a component of torso. Once a segment is added to the torso region, the contour probability is computed and recomputed to constrain the unlimited increase in coarse torso.

B. Upper-Body Segmentation on CT:

The CT and the detected face region directly provide strong hard constraints for upper-body segmentation, and the t-links connecting to the upper body can be constructed by adopting kernel density estimation (KDE). Given a pixel x, the similarity between the pixel and the torso region {xi} and the face region {xj} is defined as

$$f_t(x) = \frac{1}{mnh^2} \sum_{x_i \in C} K\left(\frac{x - x_i}{h}\right) \qquad (3)$$

where m is the number of pixels in the face region S_f, n is the number of pixels in CT, and K is the kernel function. Here, we use the Gaussian kernel with mean zero and variance h as follows:

$$K(x_1 - x_2 h) = \frac{1}{\sqrt{2\pi}} e^{-\frac{d^2(x_1, x_2)}{2h^2}} \qquad (4)$$

Where $d^2(x_1, x_2)$ is the Euclidean distance between x1 and x2. According to the anthropometry, we initialize a circle region centered at the neck for upper-body segmentation. Therefore, we can set background seeds outside the initial circle region. In addition, the pixels in the lines close to the waist and head belonging to the background or not can be determined according to the CT and detected face region, by comparing the similarity of the pixels to the torso and face regions, thereby much noise removed in segmentation.

V. LOWER BODY SEGMENTATION

Lower body segmentation follows the upper body segmentation. The accurate segmentation of the lower body depends primarily on the segmentation accuracy of the torso segmentation.

To perform the lower body segmentation, the center of the segmented torso is first identified. Let this point be represented as (x,y). Then extend the y axis below, draw rectangles at regular intervals depending on the height and width of the torso. Set these points as foreground points. Set the background by extending the x axis, based on the approximate width of the torso.

In lower body segmentation, multiple foreground and background labels are assigned. Similar intensity values are assigned to each labels. Segmentation is based on these similar intensity values. Also foreground and background probabilities are assigned based on the intensity values. Graph cuts [8] are used to determine the probability of the foreground and background.

Lower body segmentation is more challenging than upper body segmentation, because the poses of legs are unpredictable. We separate the lower body from the scene, so the segmented upper body can be set to the background. An iterative MOH algorithm is used to obtain fine results. MOH is used to describe the projection information of the coarse lower body, which can be used to find the false negatives. Each bin of MOH represents multiple cues of coarse segment results: accumulation, span, number of line segments, and boundary points of figure/ground on each projection line. The accumulation refers to the number of all segmented pixels that divide the projection line into multiple segments in a given bin; and the span is defined as the length of a line segment. MOH can obtain the missed parts and judge the integrity of the lower body, so that it is used to update Graph Cuts seeds.

VI. EXPERIMENTAL RESULTS

The proposed scheme is experimented on various images with various pose, appearance and different intensity backgrounds. For a given input image in figure 2(a), the results of the segmentation scheme is shown.

Input Image

Figure 2(a). Input image.

The AGMC and face detection are performed on the input image followed by the coarse torso detection, from which the upper body is segmented and then the lower body segmentation is carried out.
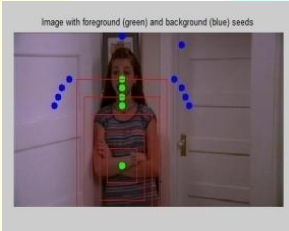

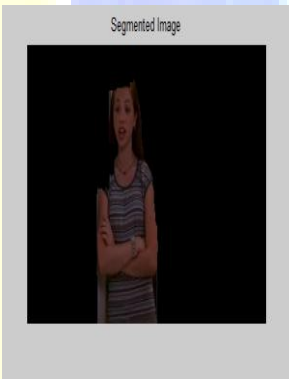Figure 2(b). Face detection

Figure 2(c). Coarse torso detection

Figure 2(d). Segmented image

VI. PERFORMANCE ANALYSIS

The performance of the above proposed concept is calculated using two parameters, accuracy and specificity. The segmented image is compared with the ground model and the following are calculated. Hit rate, given in equation (5) defines the success rate of the pixels that match between the segmented image and the groud truth image. False Acceptance Rate given in equation (6) represents the incorrectly matched pixel rate between the segmented image and

ground truth image. The test outcome positive and the test outcome negative are defined as follows in equation (7) and (8) respectively. The accuracy and specificity are calculated using the formula as given in equation (9) and (10).

Hit rate = True Positive / Test Outcome Positive　　　　　　　　　　　　　　　　(5)

False Acceptance Rate= False Negative /Test outcome Negative　　　　　　(6)

Test Outcome Positive= True Positive + False Positive　　　　　　　　　　　(7)

Test Outcome Negative=True Negative + False Negative　　　　　　　　　　(8)

Accuracy=True Positive Rate　　　　　　(9)

Specificity = 1- False Positive Rate.　　(10)

True Positive Rate is the fraction of true positives out of the positives. False Positive Rate is the fraction of false positives out of the negatives.

A ROC (receiver operating characteristics) curve is plotted to evaluate the segmentation process. The True Positive Rate is plotted against the False Positive Rate. Two curves are obtained. The slope of the accuracy curve will give the accuracy rate.
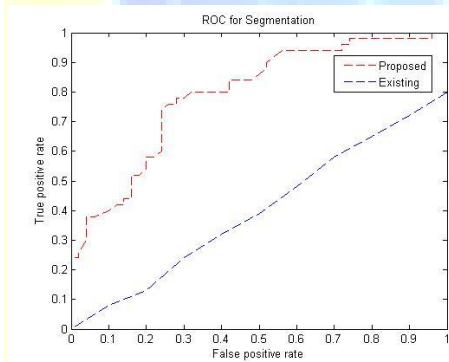

Figure 3. ROC curve

After performing segmentation with different images and comparing them with their ground model, the accuracy is found to be higher.

The above figure plotted shows that the segmentation scheme of the proposed method is much faster and efficient than the earlier conventional methods.

VII. CONCLUSION

This work aims to bring in more accuracy to the entire segmentation process. In this paper we have proposed a framework that will effectively segment the human body from color photos. The segmentation accuracy is greater compared to the existing schemas. The work can be extended to segmenting human images even when the frontal face is not available.

A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories
Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A.

International Journal of Management, IT and Engineering
http://www.ijmra.us

409

VIII. REFERENCES

[1] Huchuan Lu,Guoliang Fang, Xinqing Shao, and Xuelong Li (2012), "Segmenting human from photo images based on a coarse-to-fine scheme," *IEEE Trans. Syst., Man, Cybern. B, Cybern*., Vol. 42, No. 3.

[2]Sunhee Kim and Mayungjoo Kang(2012), "Multiple-Region segmentation without supervision by adaptive global maximum clustering," *IEEE Trans. Image Process.,* Vol. 21, No. 4.

[3] M. Kumar and P. Torr (2010), "OBJCUT:Efficient segmentation using top down and bottom-cues," *IEEE Trans.Pattern Anal. Mach. Intell*., Vol. 32, No. 3, pp. 530–545.

[4] W.Tao, H. Jin, and Y. Zhang (2007), "Color image segmentation based on mean shift and normalized cuts," *IEEE Trans. Syst., Man,Cybern. B, Cybern*., Vol. 37,No. 5,pp. 1382–1389.

[5] J.Shi and J.Malik (2007), "Normalized cuts and image segmentation," *IEEE Trans. Pattern Ana Mach. Intell*., Vol. 22, No. 8, pp. 888–905.

[6] C. Huang, H. Ai, Y. Li, and S. Lao (2005), "Vector boosting for rotation invariant multiview face detection," *in Proc. ICCV*, pp. 446-453

[7] Paul Viola and Michael J. Jones(2004), "Robust Real-Time Face Detection," International Journal of Computer Vision 57(2), 137-154,

[8] Y. Boykov and M. P. Jolly (2001), "Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images," *in Proc. ICCV*, pp. 105–112.

[9] Z. Liu and S. Sarkar(2006), "Improved gait recognition by gait dynamics normalization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 6, pp. 863–876.