Principal Objects Detection Using Graph-Based Segmentation and Normalized Histogram

Bui Ngoc Nam and Pham The Bao

**Faculty Mathematics and Computer Science, University of Science
Ho Chi Minh City, Vietnam**

Abstract

In this paper, we introduce a new method to distinguish the principal objects in image datasets using graph-based segmentation and normalized histogram (PODSH). Unlike the usual object detection systems which require the input objects, we propose a new approach to recognize objects one might focus on when taking images. Motivated by the habit of taking picture, we suppose that the position of a main object is located near the image centre and this object always holds a large area. The normalized histogram is added to increase the effect of our system. In the experiment, we used images which consist of objects to test the precision of PODSH. Our system is implemented by Matlab.

Keywords: *objects detection, graph-based segmentation, normalized histogram, principal objects.*

1. Introduction

Locating principal objects plays an important role for many learning systems. In recent years, researchers has explored in this section. However, most projects focus on the object recognition problem in which the interested objects is initially given. There are many algorithms have got a significant result such as SIFT [1], SURF [2], PCBR [3] and Discriminatively Trained Deformable Part Models [4]. Nevertheless, consider a problem that one wants to find the set of considerable objects in complex image dataset. These methods have just listed cannot solve this problem because there are no initial required objects. Therefore, we conduct a new object measurement in order to identify the important objects.

Motivated by the high performance of the graph based-segmentation algorithm developed by Pedro F. Felzenszwalb and Daniel P. Huttenlocher [5] as mentioned in [6], we proposed the new quantity to characterize the 'importance' of an object. Based on the target when taking a picture, we realize that the main objects always located near the centre of image and occupy a large square on an image. But using just two quantities to describe the 'importance' of image causes many drawbacks will be discussed later. Therefore, we imposed new rules on the normalized histogram of each object to correct these mistakes. In addition, due to the requirement of graph based-segmentation algorithm [5], we use neural network model to train parameters of this method.

2. Determine parameters in graph-based segmentation

The use of graph based algorithm proposed by Pedro F. Felzenszwalb and Daniel P. Huttenlocher to the segment problem due to its best result and low time performance. The segments or objects in image are nonlinear data due to the fact that their area can be large or small and cannot be predicted. Therefore, this algorithm entails three input parameters to extract fine segmentations. However, these parameters are also nonlinear which is not proper to locate objects automatically. Due to that reason, we conduct a method that find out the suitable nonlinear parameters for each image by using neural network training. This model is processed by using Levenberg-Marquardt training algorithm with 20 neurons in hidden layer [7]. We used the Matlab neural network toolbox to train our model.

Pedro and Daniel used three parameters which are the variance of Gaussian function, the variable threshold function and the minimum square of component. These factors are symbolized by σ, k and min [5] respectively. Given an input image, our purpose is characterizing the (σ, k, min) vector which is also the output vector of the neural network model. The input vector is the features of an input image. Due to experiences, we recognize the relationship between an image and these aspects below:

- The size of images.
- The average value in each Red, Green, Blue and Gray color space.

- The standard deviation value in each Red, Green, Blue and Gray color space.

One might wonder what target to evaluate the fitness of an output vector. The answer is in the requirement of our problem; the most important objects need to be extracted. After testing 500 colour images, we proposed the proper number objects are from 5 to 15. Fortunately, this number is also the quantity to estimate the fitness of (σ, k, min) vector. That means the (σ, k, min) vector must satisfy the condition; the number of components must be between 5 and 15 after segmenting.

The 500 colour images are used again to test our model. This 70% images have the number of components in range 5 and 15.

3. Determine measurement for detecting principal objects

The main idea for extracting principal objects automatically by using segment method is each image component can be considered as a real life object. However, we should constraint many rules to locate the objects that one concentrates on. The first two factors come from the scenario taking picture. Because the main object is targeted, they will be situated near the image centre and take a large areas. Based on these factors, the two elements of importance vector are the square and position of an object. However, this two issues are not sufficient because in some cases the background focal point is close the image midpoint than that of the main object. Moreover, background always occupies a large square in image. This problem is illustrated by the figure 1.

Because of those disadvantages, we propose another constraint relating to the normalized histogram. In order to make this idea clearly, we use the car image above to illustrate our view. We first compare the different between the two normalized histogram respects to the car and background object in figure 2.

If we limit bins have small frequencies, the energy range of car is over 200 wider than that of background (under 100). In this paper, we propose the energy range measurement to distinguish the main objects and background. This supposition based on our experiment result.

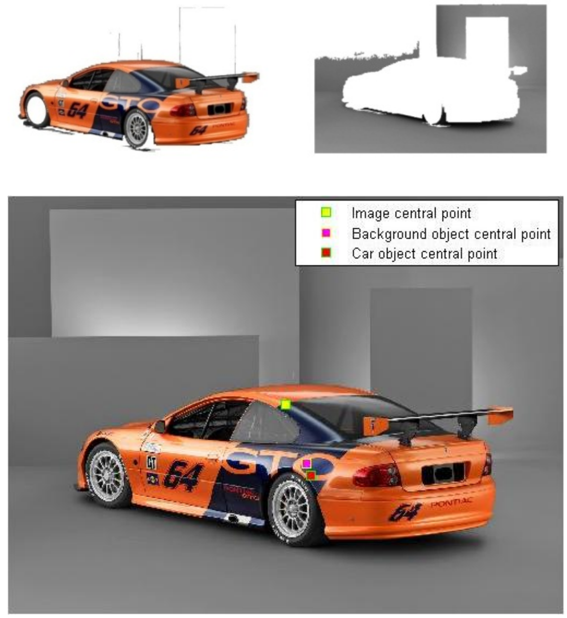

Fig. 1 A comparision between the car object and the backgroud object in terms of square and focal point. The top left image is the car object which has 603878 pixels and the focal point is the red spot on the bottom image. The top right image is the backgroud object which is compoed by 1160178 pixels and its central point is the magenta spot. The yellow spot is the image central point.

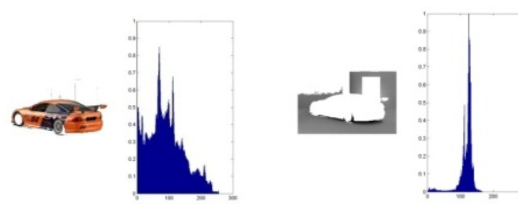

Fig. 2 The normalized histogram of car and background object.

The proven of this hypothesis relies on the probability theory as we did in [6]. Given 326 images, we extracted 447 principal objects by using PODSH. There are 347 precise main objects. Based on the central limit theorem, we gain the following result:

$$\frac{f-p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1) \qquad (1)$$

where n is 447, f is 347/447 and p is the probability this method isolates main objects.
With the 90% confidence interval in the Eq. (1), we get the formula:

$$-1.6449 \le \frac{f-p}{\sqrt{\frac{p(1-p)}{n}}} \le 1.6449 \qquad (2)$$

Solving the Eq. (2), we have p is in between 0.7422 and 0.807. That means the probability that PODSH produced the main objects is from 74% to 80% with 90% confidence interval.

After contributing the principal object measurement, the objects are filtered based on this quantity. However, the role of each value in this vector is different. Particularly, the object area is the most important factor because the small object can cause noisy in both focal point and normalized. Therefore, we should limit the small object first. After that, the sum of two left issues is taken and their average value is used to be the threshold. Objects has the summation larger than the threshold are considered as the principal objects.

4. Experiments

The image dataset consists of 2258 objects extracted from 1648 images in two categories: vehicles and train-planes. Despite of clustered background, PODSH only preserves main objects, particularly trains, planes and vehicles. The 2258 objects extracted from this dataset are used to test the precision of PODSH. The precision is the percentage of main objects account for in 2258 objects. The average result and time consuming is given by table 1. Figure 3 gives some example results of this system.

Table 1: Average precision and time consuming of principal object detection method

Average precision	Average time consuming
79.6722%	6.2453s

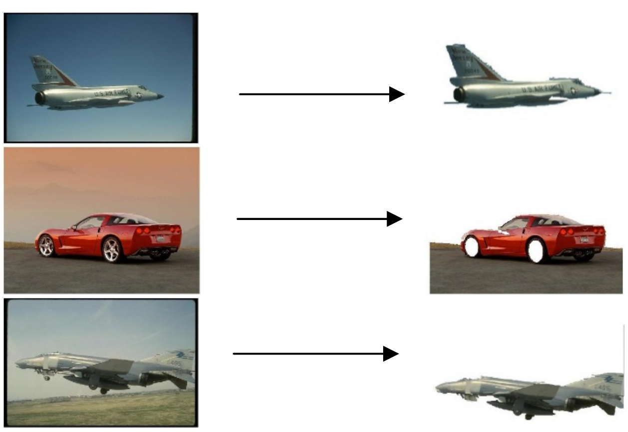

Fig. 3 The planes and cars are extracted from images with clusterd background using PODSH. The images on the left side is the original image as the input parameters of PODSH and on the other side is the principal objects extracted by using PODSH.

5. Conclusions

Beside the encouraging result, this approach still has many shortcomings due to the lack of information of edge and corner when using histogram. The results can growth if the edge, corner and frequency factors are considered. In many cases, the focal point of background always located in the image center. Therefore, more rules should be added to improve our system.

Acknowledgments

We would like to thank DOLSOFT Inc for providing us a large-scale image database with complex background.

References

[1] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, Springer Netherlands, Vol. 60, No. 2, 2004, pp. 91-110.
[2] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, 2008, pp. 346-359.
[3] H. Deng, W. Zhang, E. Mortensen, T. Dietterich, and L. Shapiro, "Principal Curvature-based Region Detector for Object Recognition", in IEEE Conference on Computer Vision and Pattern Recognition, 2007.
[4] P. Felzenszwalb, D. McAllester, D. Ramaman, "A Discriminatively Trained, Multiscale, Deformable Part Model", Proceedings of the IEEE CVPR, 2008.
[5] P.F. Felzenszwalb, D.P. Huttenlocher, "Efficient Graph-Based Image Segmentation", International Journal of Computer Vision, Springer Netherlands, Vol. 59, No. 2, 2004, pp. 167-181.
[6] N.D. Anh, P.T. Bao, B.N. Nam, and N.H. Hoang, “A New CBIR System Using SIFT Combined with Neural Network and Graph-Based Segmentation", in Proc. ACIIDS (1), 2010, pp. 294-301.
[7] M.T. Hagan and M. Menhaj, "Training feedforward networks with the Marquardt algorithm", IEEE Trans. Neural Networks, Vol. 5, No. 6, 1994, pp. 989–993.

Bui Ngoc Nam holds MSc degree in Computer Science from the University of Nottingham in 2011. He received the BSc degree in Department of Computer Science from University of Science, Ho Chi Minh City, Vietnam in 2009. He is currently a lecturer in Faculty of Computer Science of University of Science Ho Chi Minh City, Vietnam. His research interest includes computer vision, image processing and parallel computing.

Pham The Bao has obtained his Master degree and PhD in Computer Science at University of Science, Ho Chi Minh City, Vietnam. He is currently a Vice-Dean in Faculty of Computer Science and Head of Computer Science Department at University of Science Ho Chi Minh City, Vietnam. His interested research field includes computer vision, image process, parallel computing, soft computing and recognition.