

# Main Subject Detection of Image by Cropping Specific Sharp Area

FOTIOS C. VAILOULIS<sup>1</sup>, MARIOS S. POULOS<sup>1</sup>, GEORGE D. BOKOS<sup>1</sup> and NIKOLAOS ALEXANDRIS<sup>2</sup>

Department of Archives and Library Science  
Ionian University<sup>1</sup>  
Palaia Anaktora, Corfu, PO Box 96  
GREECE

Department of Informatics  
University of Piraeus<sup>2</sup>  
80 Karaoli & Dimitriou str., Piraeus 18534,  
Greece

*Abstract:* - In this paper we faced the problem of main subject detection of an image. We used a method called Composition-Guided Image Acquisition in order to get a mask around the main subject of the image. Then we cropped an area of the original image creating a new much smaller image. This area is defined as a rectangle into the mask of the main subject. Finally we use the new image as input to the Normalized Cuts method in order to get a new image with a red line around the main subject of the original image. We show that our method improves dramatically the task of detecting the main subject of an image comparing it with the use only of the Normalized Cuts method.

*Key-Words:* - Main subject detection, image segmentation

## 1 Introduction

In this paper we study the problem of main subject detection, using realistic digital photos. In particular, our aim is focused on the markup of a dominant subject of the original photo, using main subject detection and image segmentation techniques.

When someone takes a photo, usually he focuses on a subject and usually there are some other objects which are insignificant. These objects constitute the background of the photo. When a human sees a photo, immediately can decide which the main subject of the photo is and which are the objects that consist the background of the photo.

There is a method for automated main subject detection which is named Composition-Guided Image Acquisition [1] and works the following way: first, the strong edges of the original image are detected (figure 1), then with the help of the Snake algorithm [2] a mask around the main subject is created (figure 2) and finally a new image is created where the area in the original image out of mask is motion-blurred (figure 3).

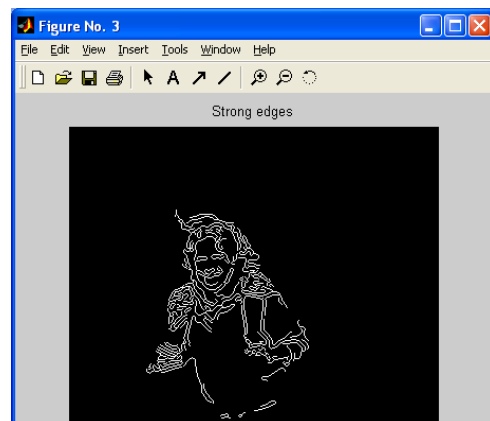


Fig.1. The strong edges of the original image

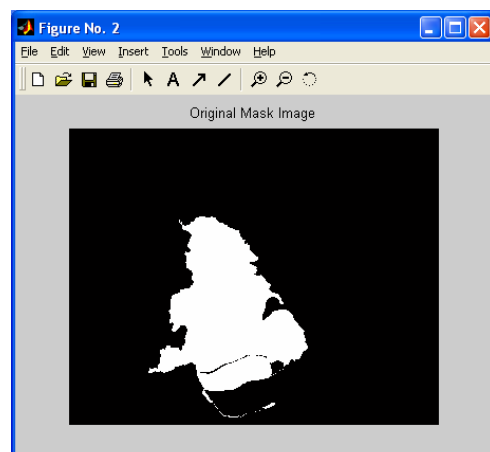


Fig.2. The mask of the main subject

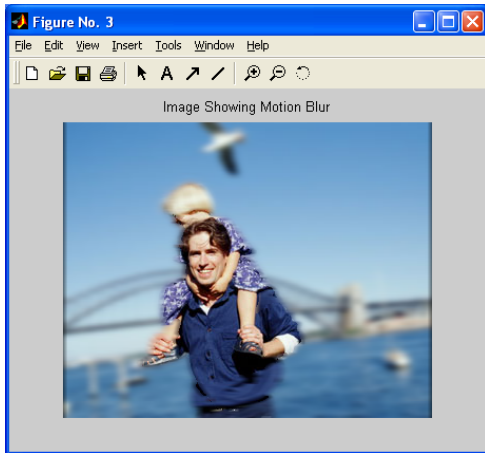


Fig.3. The image with the blurred background

The image segmentation techniques are targeted on grouping together the pixels of the image with the same characteristics such as color, contrast, brightness. A recent method for that reason is the Normalized Cuts which was provided by Stella X. Yu and Jianbo Shi [3]. This method takes the original colored image (figure 4) and transforms it to a grayscale image. Next, a similarity matrix is created with the adjacent pixels that have the same degree of gray. After that, the edges computation starts and the edges are created (figure 5). Finally the image with a red line over the discrete segments appears (figure 6).



Fig.4. The original image of a baby

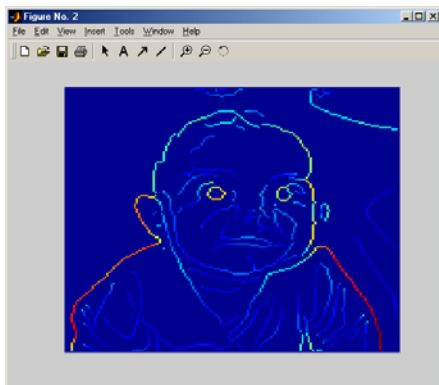


Fig.5. The edges of the image

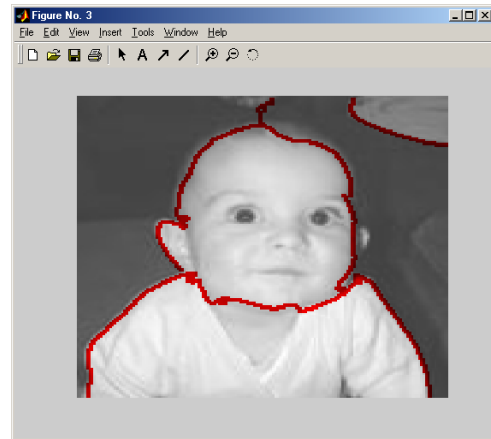


Fig.6. The segmented image

## 2 Problem Formulation

Luo, Etz, Singhal, and Gray [5] developed a computational approach to main subject detection by using Bayes neural network. Their algorithm is performance-scalable so that it need not be reconfigured for different sets of images, and involves (a) region segmentation, (b) perceptual grouping, (c) feature extraction, and (d) probabilistic reasoning and training. An initial segmentation is obtained based on the homogeneous properties of the image such as color and texture. False boundaries are removed with perceptual grouping of identifiable regions such as flesh tones, sky, and tree. Then, geometric features are extracted, including centrality, borderness, shape, and symmetry.

The Bayes Nets based method requires training time, and is not a low complexity solution for detecting the main subject on the fly. Also, as this is a Bayes net based approach, the system performance will be poor if the test set is very different from the training examples. With the vast number of possibilities of scene content, scene settings, and user preferences, developing a good set of training examples to guarantee that the neural network would perform well for a varied number of circumstances is difficult.

Wang, Li, Gray, and Wiederhold [6, 7] proposed a wavelet based approach to detect the focused regions in an image from low depth-of-field pictures. Initially, the image is coarsely classified into object-of-interest and background regions by using the average intensity of each image block, and the variance of wavelet coefficients in the high frequency bands. The variance is higher for the focused regions in the image. Blocks are clustered by using k-means algorithm [8] by noting that blocks from a homogeneous image region will have similar average intensities. Each block is further

subdivided into child blocks, and a multiscale context-dependent classification is performed for further refinement. Finally, a post-processing step removes small isolated regions and smoothes the boundaries. The segmentation accuracy of the wavelet-based segmentation algorithm is acceptable with the segmentation error varying between 4 to 7%.

In a spatial-domain approach, Won, Pyan, and Gray [9] developed an iterative algorithm based on variance maps. A local variance map is used to measure the pixel-by-pixel high frequency distribution in the image. This variance map has blob like errors both in the foreground (where the image is relatively smooth) and the background (where the background is highly textured) regions. To eliminate these errors, the authors employ a block-wise maximum a posteriori image segmentation. The block-wise maximum  $\mu$  a posteriori segmentation produces more accurate results compared to the wavelet-based segmentation. However, it requires recursion over image blocks and is computationally demanding. Further refinement of the segmentation by using the watershed algorithm [41, 42] adds to the implementation complexity.

In the problem of main subject detection, the image has to be divided into two separate classes: main subject and background. The main subject region (in focus) has crisp gradients whereas the background (blurred) has low-intensity gradients. Based on the basic assumption that the object in focus has higher gradient components compared to those not in the plane of focus, Serene Banerjee [1] proposed an algorithm which attempts to sharpen these gradients more in contrast to the blurred regions. Edges are detected over this sharpened image and subsequently a continuous smooth contour is defined by using deformable active contour model.

Jianbo Shi and Stella X. Yu [3] considered data clustering problems where partial grouping is known a priori. They formulated such biased grouping problems as a constrained optimization problem, where structural properties of the data define the goodness of a grouping and partial grouping cues define the feasibility of a grouping. They enforced grouping smoothness and fairness on labeled data points so that sparse partial grouping information can be effectively propagated to the unlabeled data. Considering the normalized cuts criterion in particular, their formulation leads to a constrained eigenvalue problem. By generalizing the Rayleigh-Ritz theorem to projected matrices, they found the global optimum in the relaxed continuous

domain by engender composition, from which a near-global optimum to the discrete labeling problem can be obtained effectively.

### 3 Problem Solution

To solve the problem of main subject detection, we combined two methods. The first is the Composition-Guided Image Acquisition [1] which gives us a mask around the main subject of the image. The second is the Normalized Cuts [3] which finds the edges of the discrete regions of an image and draws a red line over these edges.

In particular, we take the original colored image and give it as input to the Composition-Guided Image Acquisition method. The output of this method is a black-white mask like the one which is shown in figure 2. This mask encompasses the main subject of the image but not only this. It has needless areas. Hence, we thought of cropping a rectangle from that mask, not from the black and white image of the mask, but from the original image since we knew the address of each pixel of the mask (the white area of figure 2). Next, with the area that we cropped, we created a new colored image much smaller from the original which is a focus over the main subject of the original image. In other words, the new image is like someone has done zoom in over the main subject of the original image. Finally, we give the new image as input to the Normalized Cuts method. The output of this method is a segmented grayscale image with a red line over the edges of the discrete areas.

We compared the results of our method with the results of the Normalized cuts and we concluded that our method has a great improvement in the matter of the main subject detection of an image. Next we give four original images and the final images of normalized cuts method for the original image as input and for the cropped area as input to it.



Fig.7. The original image

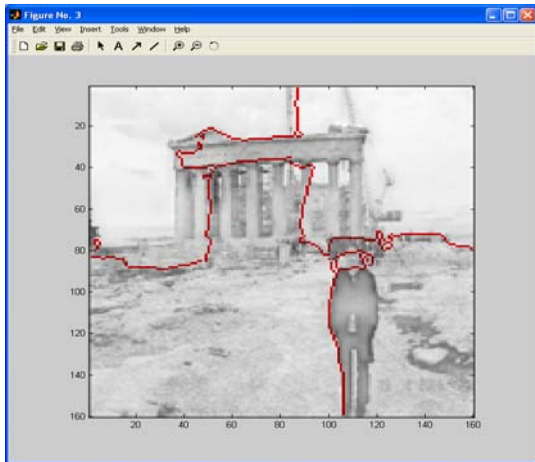


Fig.8. Final image when input was the original

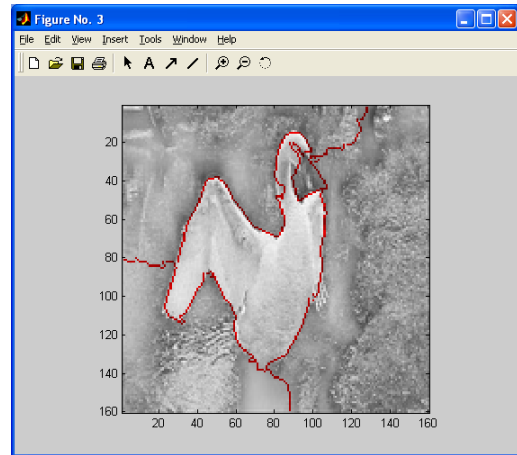


Fig.12. Final image when input was the cropped

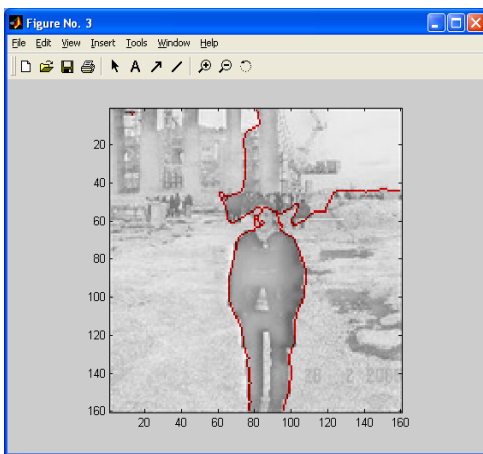


Fig.9. Final image when input was the cropped



Fig.13. The original image



Fig.10. The original image

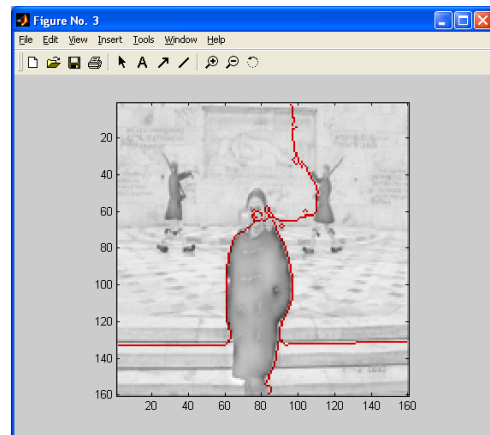


Fig.14. Final image when input was the original

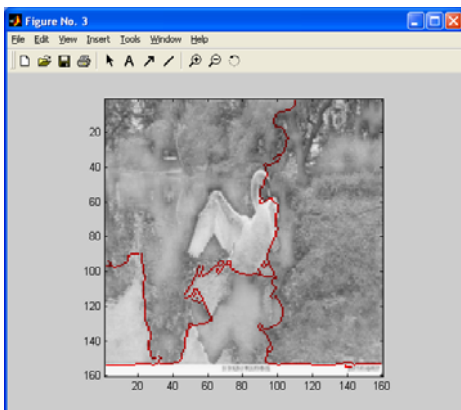


Fig.11. Final image when input was the original

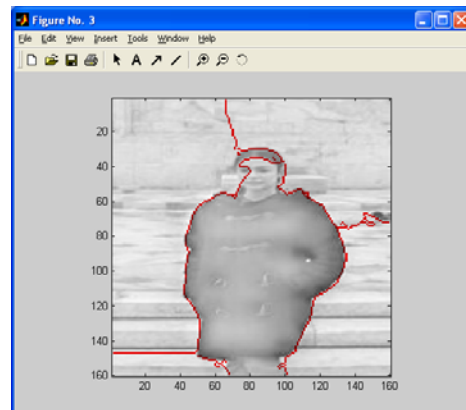


Fig.15. Final image when input was the cropped



Fig.16. The original image

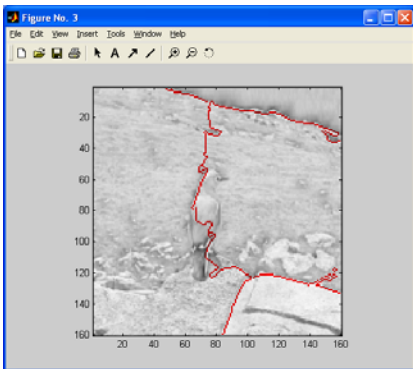


Fig.17. Final image when input was the original

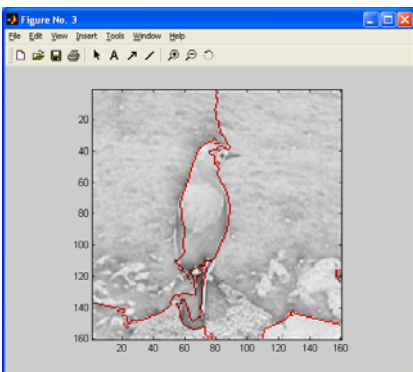


Fig.18. Final image when input was the cropped

#### 4 Conclusion

From the above figures, it is clear that our combination attained the goal of main subject detection. In comparison with the use of Normalized Cuts on the original image, the improvement depends on the specific image and the contrast of light and shadow that the image has. As we can see from our experiment, the best improvement was on the image of figure 10 (the swan in the lake) and the less improvement on the image of figure 16 (the bird on the rock). Further investigation of methods that process the images before they get as input to Normalized Cuts may lead to more interesting results as to the problem of main subject detection.

#### References:

- [1] Serene Banerjee, Composition – Guided Image Acquisition, *Ph. DEE, The University of Texas at Austin*, May 2004.
- [2] M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active Contour Models, *Int. Journal of Computer Vision*, vol. 1, pp. 321-331, 1987.
- [3] Stella X. Yu and Jianbo Shi, Segmentation Given Partial Grouping Constraints, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.26, No.2, February 2004.
- [4] S. Banerjee and B. L. Evans, A Novel Gradient Induced Main Subject Segmentation Algorithm for Digital Still Cameras, in *Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers*, Nov. 2003.
- [5] J. Luo, S. P. Etz, A. Singhal, and R. T. Gray, Performance - Scalable Computational Approach to Main Subject Detection in Photographs, in *Proc. SPIE Conf. on Human Vision and Electronic Imaging*, vol. 4299, pp. 494-505, Jan. 2001.
- [6] J. Li, J. Z. Wang, R. M. Gray, and G. Wiederhold, Multiresolution Object-of-Interest Detection for Images with Low Depth of Field, in *Proc. IEEE Int. Conf. on Image Analysis and Processing*, pp. 32-37, Sept. 1999.
- [7] J. Z. Wang, J. Li, R. M. Gray, and G. Wiederhold, Unsupervised Multiresolution Segmentation for Images with Low Depth of Field, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 85-90, Jan. 2001.
- [8] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, An Efficient k-Means Clustering Algorithm: Analysis and Implementation, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 881-892, July 2002.
- [9] C. S. Won, K. Pyan, and R. M. Gray, Automatic Object Segmentation in Images with Low Depth of Field, in *Proc. IEEE Int. Conf. on Image Proc.*, pp. 805-808, Sept. 2002.
- [10] C. S. Won, A Block-Based MAP Segmentation for Image Compression, *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 8, pp. 592-601, Sept. 1998.
- [11] L. Vincent and P. Soille, Watersheds in Digital Space: An Efficient Algorithm Based on Immersion Simulations, *IEEE Trans. on Pattern Matching and Machine Intelligence*, vol. 13, pp. 583-598, June 1991.