

# Segmentation and background extraction with application to e-catalogue

Jooyoung Hahn<sup>1</sup>, Kiwan Jeon<sup>1</sup>, Chang-Ock Lee<sup>1</sup> and Kilsu Park<sup>1</sup>

1) *Division of Applied Mathematics, KAIST, Daejeon 305-701, KOREA*

Corresponding Author : Chang-Ock Lee, colee@amath.kaist.ac.kr

## ABSTRACT

Background extraction is based on the segmentation by which an image is partitioned into several homogeneous regions. When an image can be divided into an object and the other, so-called the background, the background extraction means removing the background from the image while the object remains unchanged.

This work is motivated by special purpose in e-business to make 3D VR(virtual reality) contents which are one of the basic efficient services to arouse the consumers' interest and make the products be sold on the internet virtual market as well as in the real market. This is because using 3D VR contents of the product people can experience it as in real world though the product is beyond people's reach. Furthermore e-Catalog and e-Manual based on virtual reality have unlimited potential demands in markets even in Japan and China. In order to improve the reliability and performance of 3D VR solutions it needs to establish the system to make the solutions automatic without extensive manual works. In particular, the background extraction of the real photograph has been accomplished by manual labor with the help of commercial programs such as Photoshop or Gimp, which is time-consuming and requires a lot of extra handwork. Therefore to establish the automatic manufacturing system of 3D VR contents it is necessary to develop a software as well as a new and novel algorithms to treat the background processing with simple manipulation. Thus this work is focused on the development of the algorithm which is based on mathematical framework and the development of the software which can implement the algorithm using the commercial program language visual C++ based on MFC.

In this talk we introduce two algorithms which are used in finding the boundary between the object and the background in an image. One is the statistical without-edge algorithm and the other is the modified RAGS (region aided geometric snake). These methods are originated to the segmentation or edge (boundary) detection in an image by minimizing the energy functional. We modify these methods slightly to adjust to the background extraction which requires more accurate detection of the boundary.

One of ways to represent the boundary between regions is the level set method which has been used as an efficient method in many applications in image processing, computer vision and so on. When a curve is given as the boundary in an image, we can represent the curve as the level set of some function of which the values inside the curve are positive and the values outside the curve are negative. The curve is the set of zeros of the function. In this frame the boundary of the image can be detected as the function evolves according to the partial differential equations related to some energy functional. The level set method can treat the topology change problem efficiently.

First, we consider the statistical without-edge algorithm which is reported by Rousson and

Deriche [1] for the segmentation of vector valued images. They suggest an energy functional which is related to the statistical parameters, i.e., means and variance, inside and outside the curve. Using the level set theory the evolving equation is obtained as follows:

$$\frac{\partial \phi}{\partial t} = \delta_{\epsilon}(\phi) \left( \nu \nabla \left( \frac{\nabla \phi}{|\nabla \phi|} \right) - \frac{(u - u_1)^2}{\sigma_1^2} + \frac{(u - u_2)^2}{\sigma_2^2} + \log \frac{\sigma_2^2}{\sigma_1^2} \right) \quad (1)$$

where

- $\phi$  : level set function
- $u = u(x, y)$  : image intensity at  $(x, y)$
- $\mu_1$  : mean inside the curve
- $\sigma_1$  : standard deviation inside the curve
- $\mu_2$  : mean outside the curve
- $\sigma_2$  : standard deviation outside the curve.

When the initial level set is given it is evolved with a gradient descent using the equation (1) while the Gaussian parameters for each region (inside and outside the curve) are updated at each iteration. But in our work we modify the model so that the parameters are analyzed before the iteration to obtain faster convergence to the final level set. Using the histogram distribution of the image, we can infer the mean and variance inside and outside the curve and then use these information as the initial guess to the iteration. If an image is very simple so that it can be regarded as the mixture of two Gaussian distributions of the object and the background, the intersections of two distributions are very useful to represent the boundary between the object and the background. With this analysis we observe the expected segmentation result, determine whether the segmentation will be reasonable or not and then adopt the statistical information to the algorithm. In most cases one iteration is enough to get the final segmentation results, which tells us the efficiency of this algorithm.

Another method to extract the background from an image is the RAGS(region aided geometric snake) model which is reported by Xie and Mirmehdi [2]. This model is very similar to the geodesic snake which finds a geodesic under the new metric given by the image itself. But RAGS improves the geodesic snake by taking account of the region information, so-called the region map. From the region map we can induce the vector field to drive the snake closer to the boundary than the geodesic snake. The equation becomes

$$\frac{\partial \phi}{\partial t} = g(I)(\kappa + c)|\nabla \phi| + \nabla g \cdot \nabla \phi - \nabla R \cdot \nabla \phi \quad (2)$$

where

- $\phi$  : level set function
- $I(x, y)$  : image intensity at  $(x, y)$
- $g$  : edge stop function;  $g(I) = \frac{1}{1+|\nabla I|^2}$
- $\kappa$  : curvature of the curve
- $c$  : constant speed
- $R$  : region map of the image.

The edge stop function is designed to make the curve stop near the edge owing that it approaches 0 as the gradient  $|\nabla I|$  has large value near the edge. As the deformable level set is getting closer to the edge the last term plays an important role and make the curve stop at the correct edge though the edge is not strong due to the diffusive reflection of the light. This algorithm,

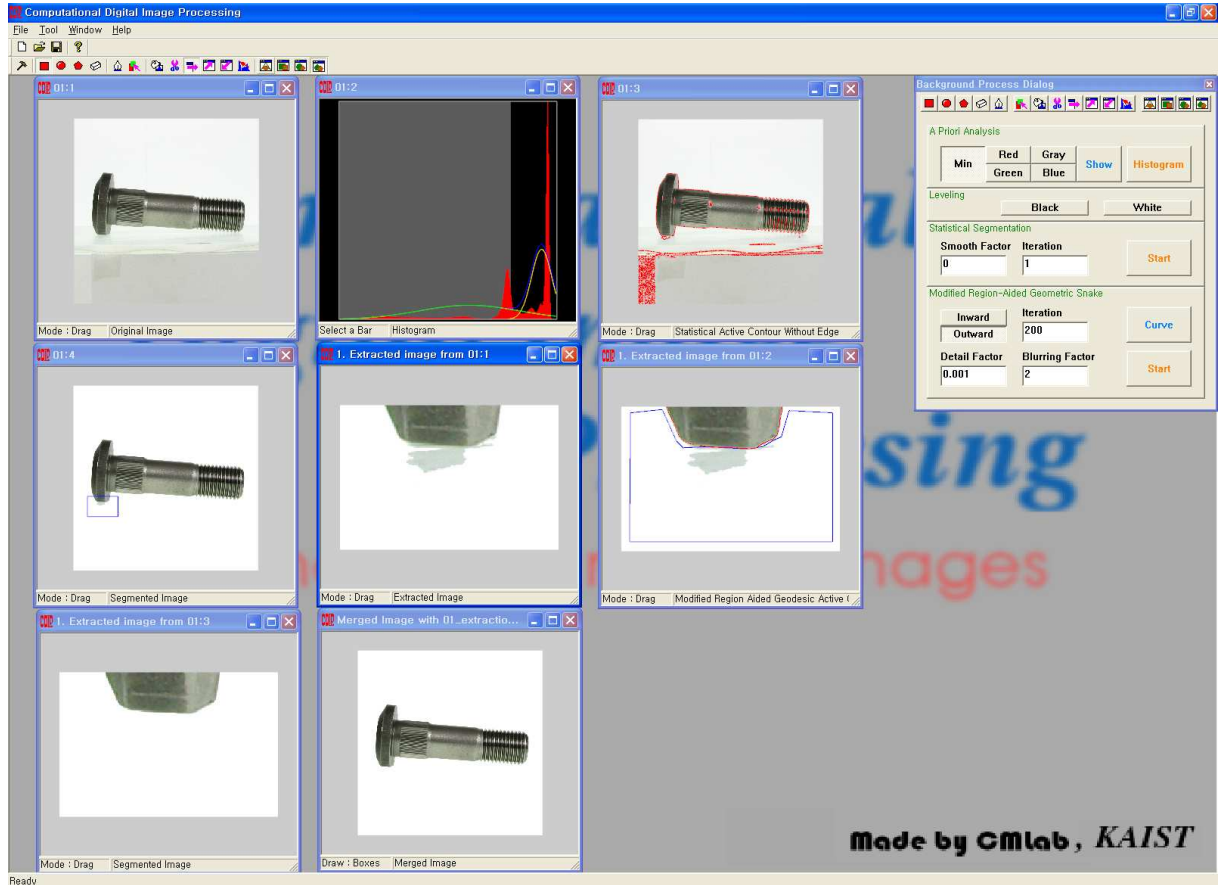


Figure 1. The CDIP software and the background process of the bolt image

however, has defects to determine the constant speed  $c$  and to design the edge stop function  $g$  more accurately, that depends on the image itself. Now we consider the modified RAGS which just uses the intensity difference  $|\nabla I|$  for the region map. Then last two terms in (2) play the same role and hence can be reduced to the term  $\nabla R \cdot \nabla \phi$ . If the moving of the curve by the curvature is ignored then the equation becomes

$$\frac{\partial \phi}{\partial t} = g(I) c |\nabla \phi| - \nabla R \cdot \nabla \phi$$

In this formulation we use three R, G, B components to compute  $|\nabla I|$  in order to get the sensitive edge information. Moreover we take the normalized vector field  $\nabla R$  and let the speed  $c$  vary but keep the magnitude of the ratio of the change fixed. That is, when the level set evolves outward (or inward), we choose the speed  $c$  so that the amount

$$c - \frac{\nabla R \cdot \nabla \phi}{|\nabla \phi|} = \pm 1$$

away from the edge in order to keep the direction of the evolution of the curve unchanged or keep the curve not disappeared during the evolution. Note that the function  $g$  is almost 1 away from the edge. This choice also enables the curve to evolve uniformly everywhere. On the other hand, since the function  $g$  is very small near the edge the curve evolves under the effect of the last term  $\nabla R \cdot \nabla \phi$ .

These two algorithms are implemented in CDIP program which is coded by means of C++

language based on MFC framework. The program is also equipped with basic tools for drawing the geometric objects (rectangles, circles, polygons) and modifying them. In order to extract the background from the image we suggest to follow a proper procedure. The first step is to expect the segmentation by histogram analysis before starting the statistical without-edge algorithm. If the statistical without-edge algorithm fails to detect the correct boundary, we locally extract the failed part of the image and then apply one of two algorithms to the extracted image and merge the results to the original segmented image to get the final background extraction result. Figure 1 shows the frame of the developed software and the process of the background extraction for the real photograph. From the left top there are original image and histogram analysis and the result of statistical without-edge algorithm. In the middle the first is background extraction result. Other two images are the locally extracted image and the result of the modified RAGS algorithm for the extracted image, where the blue curve represents the initial level set and the red one is the final level set. With the image whose background is processed the final result image is shown in the bottom after the merge process. The dialog box, called Background Process Dialog, is shown in the right, which controls the whole process of the work.

## REFERENCES

1. M. Rousson and R. Deriche, A Variational Framework for Active and Adaptive Segmentation for Vector Valued Image, *Research Report*, N. 4515, INRIA, July 2002.
2. X. Xie and M. Mirmehdi, RAGS: Region-Aided Geometric Snake, *IEEE Transactions on Image Processing*, V. 13 (2004), pp. 640-652.