A New Snake Algorithm for Object Segmentation in Stereo Images

Shin-Hyoung Kim*, Jae-Ho Choi**, Hyun-Bin Kim***, Jong-Whan Jang*

*Department of Information and Communication Engineering
PaiChai University, Daejeon, Korea
**Chonbuk National University, Jeonju, Korea
***Electronics and Telecommunications Research Institute, Daejeon, Korea
* {zeros, jangjw}@mail.pcu.ac.kr **wave@moak.chonbuk.ac.kr ***hbkim@etri.re.kr

ABSTRACT

In this paper, we propose a novel snake algorithm for object segmentation that can overcome the limitations found in cluttered background, overlapped object placements, and the problems of initial snake points. Our newly designed algorithm is developed by using disparity information taken from several sets of stereo images. The experiment results have exhibited a better performance over the well-known snake algorithm in terms of segmentation accuracy. It is also shown that our proposed new segmentation algorithm can be extended to the tracking of object visible boundary and occlusion detection in the three dimensional disparity space.

1. INTRODUCTION

Object segmentation and tracking is an important research topic for many areas in computer vision such as compression, communication and interactivity, video surveillance, and robotics. DTV, games, VOD are typical examples that can exploit the usefulness of content-based multimedia interactivity [1]. Moreover, a lot attention is drawn on stereoscopic system that can render a realistic 3D display for applications in 3D TV [2]. In all cases, object segmentation is a basic starting point. Recently, several snake algorithms have emerged a powerful tool for obtaining a clear contour of objects[3,4] and tracking[5] in still and image sequence, respectively. They are effective when the image background is relatively uniform and the objects are displaced one another. For these snake algorithms to be successful, the initial snake points need to be placed very close to the object boundary.

In this paper, we propose a novel snake algorithm for object segmentation and tracking the visible boundary of an object that can overcome the limitations found in cluttered background, overlapped object placements, and problems of initial snake points. By exploiting the essential disparity information taken from the stereo images and designing a better energy function, the proposed algorithm can successfully extract clear boundaries of objects in image that are complex and occluded.

This paper is organized as follows. Section 2 covers the fundamentals in 2D snake algorithms. Our novel snake algorithm is developed by using disparity information taken from a set of stereo image is presented in Section 3. Our proposed new segmentation algorithm can be extended to the tracking of object visible boundary and occlusion detection in the three dimensional disparity space. In Section 4, the simulation for evaluating the performance of our method is included, and the conclusions are given in Section 5.

2. FUNDAMENTALS OF SNAKE ALGORITHM

In the discrete formulation of active contour models the contour is represented as a set of snake point \( v_j = (x_j, y_j) \) for \( i = 0, \ldots, N-1 \), \( j = 0, \ldots, M-1 \) where \( x_j \) and \( y_j \) are the x and y coordinates of the snake point respectively and \( v_j \) is \( i^{th} \) snake point at \( j^{th} \) iteration; \( N \) is a total number of snake points; \( M \) is a total number of iterations. Then the typical form of energy function for the snake algorithm can be written as follows:

\[
E_{sna}(v_j) = \min \left\{ \sum_k \left[ aE_{cur}(v_j) + bE_{ext}(v_j) + cE_{int}(v_j) \right] \right\}
\]

In this type of energy function applicable for 2D object segmentation given a single image, the energy function is composed of three energy elements that are computed at \( v_j \).

Those three energy elements are continuity force, curvature force, and external force, respectively [4]. During the iterative energy minimization process, the location gets renewed to a location \( v_j' \). However, these methods were applied to a single image under the assumption that the background is simple, objects are not overlapped one another and the initial snake points need to be placed close to the object boundaries.

3. ENERGY FUNCTION IN DISPARITY SPACE
In this section, our unique energy function defined in a 3D disparity space is described. In order to successfully segment objects in such a complexity, it is essential to obtain a better descriptive method expressing one object from another even when objects are overlapped. The main goal in designing a new energy function is to enhance the object segmentation performance of a snake algorithm in a complex image environment where objects to be segmented are overlapped one another.

Consider the 3D disparity space is defined by $x$, $y$ and $d$ which represents horizontal, vertical, and disparity $\text{dis}$ in short axes, respectively [7]. The stereo image description in the disparity space is illustrated in Fig. 1. The figure contains the left and the right images of the stereo image in (a), the disparity map image in (b), and the objects in disparity space using a disparity map in (c). As shown in (c) the objects can be detached by $\text{dis}$ in disparity space, hence each object can be effectively segmented even when given a complex image.

Extending the basic concept of 2D snake algorithms that are effective for object segmentation in a single image, we designed a new energy function that is suitable for stereo images as follows:

$$E_{\text{stereo-snake}}(v^j) = \min \left\{ \sum_{i,j} \left[ E_{\text{int}}(v^j_i) + \beta E_{\text{ext}}(v^j_i) + \gamma E_{\text{cur}}(v^j_i) \right] \right\}$$

where $v^j_i = (x_{i,j}, y_{i,j}, d_{i,j})$ is a snake point in the disparity space.

First, the continuity energy is one of the internal energy in the disparity space, and it represents a distance between the current snake point $v^j_i$ and previous snake point $v^{j-1} _{i+1,j}$ as follows:

$$E_{\text{int}}(v^j_i) = \left\| \overline{d^j} - \overline{v^j_i - v^{j-1}_{i+1,j}} \right\|^2$$

where $\overline{d^j}$ is the average distance between snake points at $j$th iteration.

Second, the curvature energy is defined as follows:

$$E_{\text{cur}}(v^j_i) = \left\| v^j_{i,j} - 2v^j_{i,j} + v^j_{i+1,j} \right\|^2$$

The curvature energy gets a larger where the snake contour is bending.

Third, the external energy involved in this stereo-snake algorithm has a meaningful distinction. As the external energy definition used in 2D snake algorithm relies on edge information taken form a 2D image one can hardly expect satisfactory results when the objects are overlapped and the background is occluded. In order to cope with this complexity, we need a better external energy function.

Here, the edges in the stereo image set map onto a disparity space before being applied to external energy computation. Consider the following description for designing a new external energy function for our stereo-snake algorithm.

![Fig. 1] Image description in disparity space: (a) Stereo images (left image $f(x,y)$, right image $f(x,y)$); where objects $A$, $B$ and $C$ have different $\text{dis}$ (b) Disparity map $f_{\text{dis}}(x,y)$ (c) Objects in disparity space.

![Fig. 2] Proposed external energy in disparity space: (a) Watershed transformation image $f_{\text{WS}}(x,y)$ (b) Mapping image in disparity space $f_{\text{dis}}(x,y)$, $d$ region of object $A$ (d) Edge eliminated image $f_{\text{ed}}(x,y,d)$.

**Step 1.** Obtain a watershed image $f_{\text{WS}}(x,y)$ by applying watershed transformation [8] to one of the stereo images, $f(x,y)$ or $f_{\text{ed}}(x,y)$ as illustrated in Fig. 2. Since the watershed transformation has a drawback of over-segmentation, the region merging should be followed to suppress the over-sampled area.

**Step 2.** Transform the watershed image $f_{\text{WS}}(x,y)$ into $f_{\text{dis}}(x,y,d)$ in disparity space using the disparity map $f_{\text{dis}}(x,y)$ given in Fig. 1 (b). The transformation process is defined as follows:

$$f_{\text{dis}}(x,y,d) = \begin{cases} f_{\text{ed}}(x,y) & d = f_{\text{dis}}(x,y) \\ 0 & d \neq f_{\text{dis}}(x,y) \end{cases}$$

Then, the objects $A$, $B$, and $C$ contained in $f_{\text{dis}}(x,y,d)$ are detached as shown in Fig. 2 (b). However, due to stereo matching window resolution, the contour of an object in $f_{\text{dis}}(x,y,d)$ has fattened area around the object edge as shown in Fig. 2 (c). Since the fattened edge does not form a closed contour it needs to be eliminated. To eliminate the fattened area around edge, we applied a well-known scheme described in [9], and we finally obtained the object contour image $f_{\text{stereo}}(x,y,d)$ in the disparity space for the object $A$ as shown in Fig. 2 (d).
For a performance measure is determined to balance the relative value, the more the snake points converge occupies a bulk in the interested region (IR) linking snake points converges to the boundary of points in disparity space is shown in Fig. 3. The contour coordinate of each snake point (motion model for estimating the motion. In this model, the object using disparity information. We employ an affine motion model for estimating the motion. In this model, the coordinate of each snake point (x_{i,j}, y_{i,j}) on a boundary of object in the current frame \( t \) is related to its corresponding position (\( x_{i,j+1}, y_{i,j+1} \)) in the previous frame. We don’t consider \( d \) axis, because we supposed that \( d \) information of single object is equal,

\[
\begin{align*}
    x_{i,j} &= a_1 x_{i,j-1} + a_2 y_{i,j-1} + a_3 \\
    y_{i,j} &= a_1 x_{i,j-1} + a_2 y_{i,j-1} + a_4
\end{align*}
\]

Where \( \{a_1, \ldots, a_4\} \) are the affine parameters that include moving, scaling, and rotation. Fig. 4 shows boundary prediction in the presence of occlusion area. We pick up visible three snake points in the previous frame \( t-1 \) and obtain a necessary affine parameter set from them. Then, every other snake points in current frame \( t \) are estimated using the affine parameters. The problem that another object occludes the tracked object is solved using disparity information. If snake point is occluded by another object, \( d \) is different at the position. Hence, we can detect occlusion area by comparing the disparity information.

4. EXPERIMENTAL RESULTS

The proposed algorithm is implemented in the disparity space in order to assess the performance of our stereo-snake algorithm using Pentium IV machine running at 2GHz with 512MB of memory and the Window 2000 operating system.

We have performed a set of simulations using stereo image pairs and stereo image sequence. User selects IR including OS. Initial snake points are set evenly and automatically on the boundary of IR. The choices of values for \( \alpha \), \( \beta \) and \( \gamma \) are determined to balance the relative strengths of the terms. DIF for a performance measure is defined as follows:

\[
DIF = \frac{\sum |v_{i,t} - \alpha_{i}|}{N}
\]

where \( N \) is the total number of the snake points; \( v_{i,t} \) is the snake point at the end of iteration; \( \alpha_{i} \) is the snake point correctly positioned in the boundary of the object. The smaller the DIF value, the more the snake points converge well in OS’s boundary.

The simulation process and its results are illustrated in Fig. 5 and 6; and the performance comparison in terms of DIF is summarized in Table 1.

Consider the Experiment I where the complexity of the stereo image pair is relatively high as shown in Fig. 5. Objects are overlapped one another making the background of an object much cluttered. Noticeably, the previously presented algorithm could not segment the contour of OS (Coca Cola can) due to heavy and adverse effects of surrounding edges as shown in Fig. 5 (c). However, our proposed algorithm shows a superb convergence on the
boundary of OS because our algorithm could exploit OS’s dis information along with dis information of surroundings edges. The very successful result of our stereo-snake algorithm for segmenting an object out of the complex image is shown in Fig. 5 (g).

Fig. 6 describes an object tracking result in the stereo image sequence. We obtained snake points on the boundary of OS in the first frame using Eq. (7) and computed the position of snake points in next frame using affine motion model. We can detect snake points in occlusion area when snake point is occluded by another object. The simulation result is very successful as shown in Fig. 6 (161 and 163 frame).

The Experiment I result is summarized in Table 1 in terms of DIF. The capability of our method is well shown with comparably low DIF’s even when the image is complex in comparison to the other method.

5. CONCLUSIONS

In this paper, we presented a new snake algorithm extending conventional snake algorithms by utilizing a pair of stereo images. Our unique energy function newly defined in the disparity space was a major contribution to our snake algorithm enabling successful boundary segmentation of objects even when objects were overlapped one another and the background is cluttered. Moreover, tracking visible boundary of object using affine motion model and disparity information is successful. To further enhance the performance object segmentation in terms of speed and accuracy, it will be necessary to obtain a better stereo matching method. In the future, the results obtained in this study could be extended to a real-time object-based 3D system.

6. REFERENCES


<Table 1> Performance comparison in terms of DIF

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Previous algorithm</th>
<th>Proposed algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object Number of snake points</td>
<td>DIF (pixel)</td>
<td>DIF (pixel)</td>
</tr>
<tr>
<td>Exp. I</td>
<td>41</td>
<td>12.14</td>
</tr>
</tbody>
</table>

