Depth Map-Assisted Active Contour Segmentation Algorithm Postprint

Authors: Liu Ji, Zeng Wenliang, Liang Xiaosheng

Date: 2018-11-29T00:00:00+00:00

Abstract

To address the challenge of achieving precise image segmentation when foreground and background exhibit similar appearances, this paper proposes a novel depth-map-assisted region-based active contour image segmentation method, building upon the foundation of active contour segmentation and incorporating depth maps as auxiliary information. Firstly, a filtering algorithm is applied to refine the depth map, yielding a more complete depth representation. Subsequently, a Gaussian mixture model is employed to compute confidence maps for both the color image and the depth map. Finally, these confidence maps are utilized to calculate the weights of color and depth features within a given region, thereby guiding the segmentation process. The proposed algorithm effectively leverages both color and depth information to guide segmentation, enabling more accurate foreground-background separation. Experimental results demonstrate that the segmentation results obtained by this method closely approximate the ground truth, thereby improving the accuracy of image segmentation.

Full Text

Preamble

Vol. 37 No. 1
Application Research of Computers
ChinaXiv Cooperative Journal

Depth-Assisted Active Contour Segmentation Algorithm

Liu Ji, Zeng Wenliang†, Liang Xiaosheng
(College of Computer Science, Chongqing University, Chongqing 400044, China)

Abstract: The similarity between foreground and background presents a significant challenge in precise image segmentation. To address this issue, this paper proposes a novel depth-assisted region-based active contour segmentation method that integrates depth maps as auxiliary information into the active contour framework. First, a filtering algorithm is employed to repair the depth map, yielding a more complete depth representation. Then, Gaussian Mixture Models are used to compute confidence maps for both the color image and the depth map. Finally, these confidence maps are utilized to calculate the weights of color and depth information within a given region, thereby guiding the segmentation process. The proposed algorithm effectively leverages both color and depth information to achieve more accurate foreground-background separation. Experimental results demonstrate that the segmentation results obtained by this method closely approximate ground truth and improve the accuracy of image segmentation.

Keywords: depth image; hybrid filtering; active contours; confidence map

---

0 Introduction

Image segmentation refers to the process of partitioning an image into distinct regions to extract specific targets. Segmentation algorithms can be categorized into several types: region-based methods using color information, edge-based methods, threshold-based methods, and approaches derived from specific theoretical frameworks. Typical traditional methods include Graph Cut [1], watershed algorithms [2], and active contour methods [3]. While these methods can cleverly utilize relevant regional information to separate specific targets, they often fail to achieve satisfactory results due to limited available information, particularly when foreground and background regions exhibit similar characteristics.

To address scenarios with similar foreground and background, researchers have begun introducing depth maps to assist image segmentation [4]. Since depth image grayscale values represent the distance between objects and the camera (with larger values indicating greater distance), depth information can be exploited to separate regions where foreground and background are similar in appearance but differ in depth. For instance, Li Shutao et al. [5] proposed a method using level sets and depth maps for image segmentation, which primarily involves two steps: first segmenting the depth map using level sets, then refining the segmentation boundaries. Ge Ling et al. [6] introduced depth maps into the classical Graph Cut algorithm for interactive image segmentation, with the main idea of using depth information to modify edge weights during segmentation, thereby altering the segmentation results.

However, most methods that utilize depth images for segmentation rely on relatively complete depth maps. In practice, due to equipment limitations (such as Kinect) and environmental interference, acquired depth images suffer from various issues, including missing depth information leading to holes, low depth value accuracy, and image noise. These problems significantly degrade the performance of many depth-based segmentation approaches.

This paper proposes a method that uses depth maps captured by depth cameras to assist image segmentation. The main contributions are twofold: first, an improved filtering algorithm is developed to repair depth maps; second, a confidence map mechanism is introduced to propose a depth-assisted region-based active contour algorithm. The use of confidence maps enables reasonable utilization of both color and depth information, allowing the algorithm to rely primarily on depth information when the repaired depth map is of high quality, while still leveraging color information when depth map repair is suboptimal. Experimental results demonstrate that the proposed method can effectively restore depth images and produce superior segmentation results.

1 Algorithm Description

To address the problem of similar foreground and background in image segmentation, this paper proposes a depth-assisted region-based active contour segmentation method. The overall workflow is illustrated in [Figure 1: see original paper]. The method consists of two key components: depth map inpainting and image segmentation using the repaired depth map. The specific steps are as follows: (a) apply the proposed filtering method to fill holes in the depth map; (b) compute confidence maps for both the color image and the depth map; (c) derive color and depth weight components for a given region from the confidence maps to guide the subsequent segmentation process.

2 Depth Map Inpainting Filter Algorithm

2.1 Depth Map Hole Generation

Using Kinect as an example, we explain the causes of hole generation and demonstrate the resulting artifacts. The Kinect device comprises three main components: a sensor, a camera, and a microphone. The depth sensor consists of an infrared emitter and an infrared receiver, while the color camera captures color images of the same scene. When acquiring depth images with Kinect, the effective range is 0.8 to 4.0 meters in default mode. Objects outside this range cannot obtain depth information, resulting in data loss and image holes. [Figure 2: see original paper] shows a color image (right) and its corresponding depth image (left) captured by Kinect. In the depth image, black regions represent pixels with a value of 0, indicating hole pixels.

2.2 Filter Algorithm

Building upon the Directional Joint Bilateral Filter (DJBF) algorithm [9], this paper proposes a novel filtering algorithm for depth map inpainting. DJBF, introduced by Le et al., is formulated as follows:

The algorithm first classifies pixels as either hole pixels or non-hole pixels based on whether they reside in hole regions. Priorities are assigned to all pixels within holes, and inpainting proceeds according to this priority order. After completing the hole filling, noise in non-hole regions is removed to obtain the final repaired image. Before calculating the depth value for each target hole pixel, priorities are assigned to all unknown hole pixels, determining the order of inpainting. All hole pixels within the hole region are sorted by priority, and depth values are computed for each pixel in descending priority order. The priority $m(p,q)$ of a target pixel is determined by a linear combination of two factors: the support degree $d(p,q)$ and the confidence degree $c(p,q)$, as expressed in equation (7):

$$m(p,q) = \lambda_0 \cdot d(p,q) + \lambda_1 \cdot c(p,q), \quad \lambda_0, \lambda_1 > 0$$

The support degree $d(p,q)$ refers to the number of known neighboring pixels around the target center pixel. In our experiments, a filter window size of 7×7 is used, resulting in 48 neighboring pixels for the center pixel. The confidence degree $c(p,q)$ represents the similarity between neighboring pixels and the center pixel, including spatial proximity and grayscale similarity. Its calculation method follows the same approach as the neighborhood pixel depth weight calculation in joint bilateral filtering:

The priority calculation proceeds as follows: (a) First, an initial threshold T is set for priority m, which is experimentally set to 1.5. (b) The priority m for all pixels within the hole region is computed using equation (7). All hole pixels with priority greater than T are inserted into a priority queue $TSet$, and their depth values are calculated. (c) After completing the inpainting of all pixels in the current queue, new priorities m are recalculated for all remaining hole pixels by computing their support and confidence degrees. The threshold T is then updated based on the priorities of the remaining hole pixels, and hole pixels are rearranged in the queue according to the new threshold. Here, T is adaptively adjusted: if no pixels remain in the queue, the threshold is decreased by 0.1; otherwise, it is increased by 0.1. This adjustment ensures that pixels with the highest priority are always processed first. In each iteration, new priorities for all remaining hole pixels are computed, the threshold is updated, and new highest-priority hole pixels are continuously added to the queue for depth value calculation. (d) The loop terminates when depth values for all hole pixels have been computed, indicating completion of hole filling.

3 Depth-Assisted Region-Based Active Contour Segmentation Algorithm

3.1 Region-Based Active Contour Algorithm

Active contour algorithms can be broadly classified into two categories: region-based and edge-based [10]. Edge-based methods primarily utilize image gradient information to stop curves at object boundaries, while region-based methods seek to minimize an energy function under boundary conditions during evolution. This paper adopts the region-based active contour segmentation method proposed by Lankton et al. [11]. Unlike previous approaches, this region-based method focuses on local information rather than global image statistics, using local cues to guide the contour evolution process.

When performing image segmentation using a region-based active contour algorithm, for a given image I, let Ω be defined as the two-dimensional space of I. The energy function during evolution can be expressed as:

$$E(\Omega) = \int_{\Omega} \delta_{\phi}(x) \cdot B(x,y) \cdot F(I(y))\, dy\, dx + \lambda \int_{\Omega} \mathrm{div}\left(\frac{\nabla \phi}{|\nabla \phi|}\right) \delta_{\phi}(x)\, dx$$

where E is the evolution energy function, Ω represents the image region, δ_{ϕ} denotes the Dirac function, $B(x,y)$ indicates the distance relationship between two points (equal to 1 when the distance is less than a given radius, otherwise 0), $I(y)$ is the combined value of color and depth components, and λ is a penalty coefficient used in the regularization term to ensure curve smoothness (set as in [11]).

F is an energy measurement function that ensures points in local regions along the contour follow a given model, expressed as:

$$F = \lambda_c \cdot (u - v_c)^2 + \lambda_d \cdot (u - v_d)^2$$

where λ_c and λ_d are determined by the confidence map, v_c represents the mean of interior region pixels, and v_d represents the mean of exterior region pixels.

3.2 Confidence Map Approach

When using depth-assisted active contour algorithms for image segmentation, confidence maps are required to specify λ_c and λ_d. The discrimination method is as follows:

$$\lambda_c, \lambda_d = \begin{cases} \alpha, 1-\alpha & \text{if } S_c > S_d \text{ and } C_c > C_d \\ \beta, 1-\beta & \text{if } S_c > S_d \text{ and } C_c \le C_d \\ 1-\beta, \beta & \text{if } S_c \le S_d \text{ and } C_c > C_d \\ 1-\alpha, \alpha & \text{if } S_c \le S_d \text{ and } C_c \le C_d \end{cases}$$

where S_c and S_d represent the number of values belonging to the foreground in the corresponding regions of the color and depth confidence maps, respectively, C is the area of the given region, and α and β are constants experimentally set to $\alpha = 8$ and $\beta = 4$.

This paper employs Expectation-Maximization Gaussian Mixture Models (EM-GMM) to compute confidence maps. Several alternative approaches exist, including Gaussian models, histograms, Gaussian mixture models, and expectation-maximization algorithms. After comparing these methods, EM-GMM proved most effective. The EM-GMM computation involves two key steps: initialization and solution using the EM algorithm. The calculation proceeds as follows: (a) Initialize the Gaussian models π_k, μ_k, and Σ_k; (b) For each pixel p in image I, assign $z_{ik} = 1$ if $k = \arg\max_j \pi_j N(p|\mu_j, \Sigma_j)$; (c) Update parameters: $\pi_k = \frac{N_k}{N}$, $\mu_k = \frac{1}{N_k}\sum_{i=1}^{N} z_{ik}p_i$, $\Sigma_k = \frac{1}{N_k}\sum_{i=1}^{N} z_{ik}(p_i-\mu_k)(p_i-\mu_k)^T$; (d) Repeat steps 2-5 until maximum iteration conditions are met. By obtaining confidence maps for both color and depth, the weights can be adjusted during segmentation to modify function F, enabling depth-assisted region-based active contour segmentation.

4 Experimental Results

The experimental environment and parameters were properly configured before testing. Our method is based on NTU2000 [14]. The depth map inpainting, confidence map computation, and final active contour segmentation are executed sequentially, where optimal results from each stage contribute to the final optimal solution. During depth map inpainting, the priority threshold T is initially set to 1.5, with other parameters configured as in [9]. During active contour segmentation, λ_c and λ_d range from 1 to 10, with experiments showing that $\lambda_c = 2$ and $\lambda_d = 0.5$ yield the most suitable results, while other parameters follow [11].

Following the experimental procedure in [12], we validate our algorithm from two aspects: first, verifying the effectiveness of the filtering algorithm for depth map repair; second, validating the effectiveness of the depth-assisted region-based segmentation algorithm.

4.1 Depth Map Inpainting Results of the Filter Algorithm

The effectiveness of the depth map inpainting algorithm is demonstrated in [Figure 3: see original paper] (bell) and [Figure 4: see original paper] (tower). In these figures, from left to right are: ground truth depth map, artificially holed depth map, inpainted depth map (with errors marked in red), and final inpainted depth map. For objective evaluation, we measure Error (filling error), Ratio (hole inpainting proportion), and PSNR.

Let $Error = |HoleFilling - Groundtruth|$, where $Error$ represents the number of incorrectly filled pixels. Let $Radio = 1 - \frac{Error}{SumHole}$, where $SumHole$ is the total number of hole pixels. PSNR is an objective standard for measuring image distortion or noise level, commonly used for image quality assessment:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right)$$

where MAX represents the maximum image color value (255 for 8-bit images). MSE is the mean square error between the $m \times n$ ground truth image I and the inpainted depth image K:

$$\mathrm{MSE} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(K(i,j) - I(i,j))^2$$

Since our filter algorithm utilizes local information, the final result cannot be identical to the ground truth. However, as shown in the rightmost inpainted depth maps of [Figure 3: see original paper] and [Figure 4: see original paper], the filter algorithm effectively repairs imperfect depth maps. The data in also demonstrate that the depth maps are satisfactorily restored after filtering.

4.2 Segmentation Results of Region-Based Active Contour Algorithm

To validate the effectiveness of the depth-assisted region-based segmentation algorithm, we conduct two types of comparative experiments. The first compares with segmentation results using only color information, selecting contour-guided color segmentation [7] (ICCV 2015) and superpixel segmentation [8] (CVPR 2015). The second compares with the RGBD-based Graph Cut method proposed by Ge Ling et al. [6]. Additionally, we include FCN (Fully Convolutional Network) for semantic segmentation, though as a deep learning approach requiring extensive training data, it cannot produce satisfactory results for untrained images.

Evaluation metrics include PSNR, Jaccard index, Dice coefficient, Precision, and Conformity coefficient. The Jaccard index measures similarity between sample sets, defined as the intersection over union:

$$\kappa_j = \frac{|\Omega_1 \cap \Omega_2|}{|\Omega_1 \cup \Omega_2|} \times 100\%$$

Higher Jaccard values indicate greater similarity and segmentation results closer to ground truth. The Dice coefficient is another similarity measure, where larger values indicate greater similarity:

$$\kappa_d = \frac{2|\Omega_1 \cap \Omega_2|}{|\Omega_1| + |\Omega_2|} \times 100\%$$

Precision represents the proportion of correctly segmented samples, where higher values indicate more accurate segmentation. The Conformity coefficient, proposed in [13], provides a more sensitive and strict evaluation metric than Dice and Jaccard, offering better discrimination for detecting subtle differences in segmentation results:

$$\kappa_{cd} = \frac{|\Omega_1 \cap \Omega_2|}{\max(|\Omega_1|, |\Omega_2|)}$$

Segmentation results are shown in [Figure 5: see original paper] and [Figure 6: see original paper]. It is evident that even with holes in the depth map, our method produces results closer to ground truth because we first repair the depth map using the filter algorithm. The specific evaluation metrics in and demonstrate that the segmentation results obtained by our proposed depth map inpainting algorithm combined with the depth-assisted region-based active contour algorithm achieve superior performance across all metrics, indicating results closest to ground truth.

5 Conclusion

Image segmentation remains a major research focus in computer vision. This paper proposes a depth-assisted region-based active contour segmentation method. To address holes in practically acquired depth maps, we introduce a novel filtering algorithm based on DJBF for depth map inpainting. Our method first repairs the depth map, then computes confidence maps for both color and depth images, derives weight values for color and depth components from these confidence maps, and finally completes the segmentation process. Experimental results demonstrate that the proposed method effectively repairs depth maps and produces segmentation results closer to ground truth.

References

[1] Wen Peizhi, Chen Xiao, Wu Xiaojun, et al. Automatic target segmentation algorithm of GrabCut based on cubic spline interpolation [J]. Application Research of Computers, 2014, 31 (07): 2187-2190.

[2] Wang Lan, Zhang Hui, Zhang Haitao. Watershed image segmentation algorithm which combining with Otsu in new color space [J]. Application Research of Computers, 2017, 34 (12): 3873-3879.

[3] Li Tianqin, Zhang Yi, Liu Zhi, et al. A review of Snake model [J]. Computer Engineering, 2005, 31 (9): 1-3.

[4] Zanuttigh P, Marin G, Mutto C D, et al. Time-of-Flight and Structured Light Depth Camera [M]. [S. l.] : Springer Publishing Company, 2016.

[5] Lu Ting, Li Shutao. Image matting with color and depth information [C]// Proc of International Conference on Pattern Recognition. 2012: 3787-3790.

[6] Ge Ling, Ju Ran, Ren Tongwen, et al. Interactive RGB-D image segmentation using hierarchical graph cut and geodesic distance [C]// Proc of Pacific Rim Conference on Multimedia. Cham: Springer, 2015: 114-124.

[7] Fu Xiang, Wang C Y, Chen Chen, et al. Robust image segmentation using contour-guided color palettes [C]// Proc of International Conference on Computer Vision. 2015: 1618-1625.

[8] Li Zhengqin, Chen Jiansheng. Superpixel segmentation using Linear Spectral Clustering [C]// Proc of Computer Vision and Pattern Recognition. 2015: 1356-1363.

[9] Le A V, Jung S W, Won C S. Directional Joint Bilateral Filter for Depth Images [J]. Sensors, 2014, 14 (7): 11362-11378.

[10] Hu Ping, Bing Shuai, Liu Jun, et al. Deep level sets for salient object detection [C]// Proc of Computer Vision and Pattern Recognition. 2017: 540-549.

[11] Lankton S, Tannenbaum A. Localizing region-based active contours [J]. IEEE Trans on Image Processing, 2008, 17 (11): 2029–2039.

[12] Lei Tao, Jia Xiaohong, Zhang Yanning, et al. Significantly fast and robust fuzzy c-means clustering algorithms based on morphological reconstruction and membership filtering [J]. IEEE Trans on Fuzzy Systems, 2018.

[13] Chang H H, Zhuang A H, Valentino D J, et al. Performance measure characterization for evaluating neuroimage segmentation algorithms [J]. NeuroImage, 2009, 47 (1): 122-135.

[14] Ju Ran, Liu Yang, Ren Tongwen, et al. Depth-aware salient object detection using anisotropic center-around difference [J]. Signal Processing: Image Communication, 2015, 38 (10): 115-126.

Note: Figure translations are in progress. See original paper for figures.