Interest Contour Region Fusion Method based on Gaussian Mixture Model Segmentation

Xinyu He¹, Shu Chen¹ and Yi Li*²

¹ College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou, Zhejiang, 325035, China
²Corresponding author’s e-mail: liyi@wzu.edu.cn

Abstract. This paper presents an edge feature light-shadow fusion algorithm based on image segmentation. A new edge light shadow feature fusion algorithm based on image segmentation, morphological operation and image filtering is presented to extract target objects from image and fuse them with new background, which is cumbersome, difficult to operate, and lack of fast and effective target acquisition and feature fusion methods. The algorithm first labels the input image through an interactive system, then divides the labeled image to get the image target information, then processes the target object, obtains the edge of the target object, expands the edge to form a mask, combines the target object with the background, and filters through the edge mask to achieve the effect of fusing the target object with the background. The experimental results show that this method can process images with different complexity, not only to generate a combination of portrait and monochrome background, but also to generate a combination of other objects and natural scenery, so that the image can produce transitional natural effects.

1. Introduction

Since the beginning of the 21st century, computer science and technology have been greatly developed. Computer technology also has many application scenarios in our life. Computer vision is an essential part of computer technology, of course, the same is true. In the field of computer vision, image segmentation occupies a place. Image segmentation [1] refers to the process of cutting an image into multiple areas according to the internal characteristics of the image. The segmented image has more distinct features, making it easier for us to understand and analyze the image. Because the boundaries of objects in an image have distinct characteristics, image segmentation is usually located there. More precisely, image segmentation is the process of marking each pixel in an image, which makes a pixel with similar characteristics have the same marker[8]. The result of image segmentation is a collection of one or more areas with similar characteristics or a collection of outlines extracted from the image. Each pixel in a feature area has similar results when compared to a feature or when calculated, such as color, shading, texture, and so on. The adjacent areas are very different in a comparison of certain characteristics.

Edge detection is one of the basic problems in image processing and computer vision because edge detection is a very important part of the basic characteristics of an image[9][10][11]. The purpose of edge detection is to mark the points in the image that are very different from each other. By processing these points, the information in the image can be greatly reduced, and the information that we think is irrelevant can be removed so that the most important part of the image can be preserved. With edge detection, we can easily find the data we need.
Light and shadow feature fusion simply means filtering the image to make it appear more natural and less abrupt when combined with the new background[12]. Image filtering, that is, processing some or all of the pixels by some algorithm or standard, can emphasize some features of the image or fade some features of the image. Image processing operations such as smoothing, sharpening and edge enhancement can be achieved by filtering, and the effect of filtering will directly affect the validity and reliability of subsequent image processing and analysis[13][14]. At present, the research on edge detection and filter processing is very mature, and many effective edge detection and filter processing methods have been proposed[15].

This paper presents an edge light shadow feature fusion algorithm based on image area segmentation technology. By extracting the target from the image and combining it with the new background, and extracting its edge for expanding operation, the natural transition effect is further processed. This algorithm can be applied to all kinds of simple and complex scenes through interactive image segmentation algorithm. The segmented image is segmented by edge detection technology to get the edge feature information of the image. The combined image edge area pixels are processed by expansion operation to get the transition natural edges. Finally, the original image is fused to get the target image. Make the front and background of the new image transition naturally.

2. Dynamic Image Augmented Reality Rendering Algorithm

2.1. Grabcut Image Segmentation Method

The GrabCut[2] image segmentation algorithm is a graph-based segmentation method based on the GraphCut[3] algorithm. GraphCut is often used in foreground and background segmentation, matting, and other scenarios. It first converts the image into a grayscale image, then labels the foreground and foreground to indicate a small number of background pixels B and foreground target pixels O. It then maps the image (Graph) and introduces two special nodes, T and S, to represent the background and foreground, respectively. The algorithm joins the pixel points of the image itself to form an edge, and then connects each pixel point with T and S points to form another set of edges. The energy of edges is expressed by the similarity between two pixel points. Then if the foreground and background can be cut correctly, because the similarity between the two pixels is low, the cutting energy should be minimal. Therefore, in GraphCut, based on the maximum flow and minimum secant theory[4] in graph theory, a secant with the least cost of segmentation is found to match the front and background pixels as closely as possible, and the minimum secant can be obtained quickly and effectively by the algorithm. This enables the separation of the foreground and background.

GrabCut has made many improvements based on GraphCut. First, GrabCut uses the Gaussian Mixing Model (GMM) to statistical priori information, so color maps are naturally supported. In terms of labeling, GrabCut only needs to label the background in the way of a picture frame, which is more convenient than GraphCut. Then GrabCut uses an iterative algorithm that can be developed in the process of estimation and parameter learning to complete multiple energy minimizations instead of one. It can repeat iterations to ensure that the energy of f-segment is less and less, ultimately, to achieve image segmentation.

2.2. Canny Image Segmentation Method

Canny[5] edge detection algorithm is an edge-based segmentation method. It first uses a Gaussian filter to suppress the noise of the image and smooth the image. Then the map intensity and gradient direction of each pixel in the image are calculated, each pixel point in the image is traversed, and two pixel points in the target pixel and the positive and negative gradient direction are compared. If the gradient intensity is the largest, it is marked as an edge pixel. After non-maximum suppression, the remaining pixels are closer to the true edges of the image. However, there are still some non-edge pixels that are marked as edge pixels due to noise or color changes. Next, a double threshold detection is performed, with one high threshold and one low threshold selected. Generally, a high threshold is three times the low threshold, such as 5 and 15. Retaining points with gradient intensity higher than
the threshold in the image and deleting points below the threshold will greatly reduce the appearance of false edges. Finally, the pixels with gradient intensity between the high and low thresholds around the retained pixel points are retained together so that a more complete edge can be obtained.

2.3. Expansion and Bilateral Filtering Method

Expansion [6] is a basic operation in mathematical morphology. Its target is the highlighted part of the image, that is, the part with a large pixel value. After processing, the image will have a larger highlighted area than before. Mathematically, the dilation operation is to use the convolution kernel to process the highlighted pixels in the image, calculate the maximum value of the pixel points in the area covered by the convolution kernel, and assign the maximum value to the pixels in the area covered by the convolution core, so that the highlighted areas in the image will grow gradually.

Bilateral filter [7] is a non-linear filter. The traditional Gauss filter calculates the weights of pixel points based on distance, only considering the spatial location relationship between pixels, but not the relationship between pixel values. Therefore, the Gauss filter treats all the pixel points equally, even the pixels on the edge are weighted average using the nearby pixels. The direct result is that the edge of the image becomes blurred, that is, the edge information is lost. On the basis of Gauss filter, bilateral filter is further considered and weighted according to the relationship between the values of pixels, which has a good effect in preserving image edges.

3. The Main Algorithm Processing Flow

The main algorithm processing flow designed of edge light shadow feature fusion algorithm based on image area segmentation technology which include image segmentation, edge detection, image combination, and filter processing. The algorithm processing structure is shown in Figure 1.

4. Target Image Segmentation Processing

The purpose of image segmentation is to separate the foreground and background of the image to obtain accurate target with clear edges. Image segmentation is divided into two steps. The first step is to load a picture with the target object, drag the mouse to pull a rectangle to frame the target object, that is, to mark the foreground and background, and then use the algorithm in this paper to process the image.

After tagging the front and background pixels, the algorithm will generally use five Gaussian models to model the target and background, and the five Gaussian models will be superimposed into a Gaussian mixture model. For pixel X, there are only two possibilities, either from a Gaussian component of the target GMM or from a Gaussian component of the background GMM.
The algorithm matches the RGB color value of the pixel with the foreground and background to find the minimum region item energy. Here is a specific representation of the energy of the region item:

$$\Theta = \left\{ \pi(a,k), \mu(a,k), \Sigma(a,k), a = 0,1, k = 1...K \right\}$$

(1)

In RGB space, we use Euclidean distance instead of gray value to measure the similarity of two pixels. Specifically expressed as follows:

$$V(a,z) = \gamma \sum_{(m,n) \in C} \left[ a_n \neq a_m \right] \exp(-\beta z_m - z_n^2)$$

(2)

Where parameters $\beta$ depending on the contrast of the image, if the contrast of the image is low, the difference between the two pixels with the difference is still very low, so a larger factor is required. $\beta$ to magnify the difference between two pixels and, conversely, if the contrast is large, take a smaller factor $\beta$ to reduce the difference between two pixels. This allows the boundary item energy to work properly regardless of contrast. After obtaining the minimum cut method, all you have to do is iterate over and over again to get the minimum cut. By iterating over and over, you will gradually optimize the GMM model and the segmentation effect until you get the desired result, as shown in Figure 2. It is clear that the results obtained through continuous iteration are getting closer to the real goal.

5. Edge Detection and Expansion Processing

Generally speaking, edge detection needs to binary the image, and then filter the image to smooth the edge. In the image segmentation process, black and white masks have been obtained, so only filtering is needed. Complex images often have a variety of colors and cluttered details. Edge detection algorithms do not detect edges very well. After the segmentation process, the image becomes very simple and it is easy to obtain good continuous edges. The smoothing effect and edge detection effect are shown in Figure 3.
Now we have the edges of the target object, but it is obviously not enough to only process the edge pixels, so we need to use the expansion algorithm in mathematical morphology to expand the edges, so we can get more edge area pixels, as shown in Figure 4.

6. Image Combination and Filtering
The purpose of this step is to use the mask generated by the image segmentation to extract the target object from the image and combine it with the background image, as shown in Figure 5, for preprocessing. The image is processed by bilateral filtering because it introduces weights related to the pixel values, so the pixel value weights are expressed as $G_r$, and the distance weights are expressed as $G_s$.

$$G_s = \exp \left( -\frac{||p - q||^2}{2\sigma_s^2} \right)$$  \hspace{1cm} (3)$$

$$G_r = \exp \left( -\frac{I_p - I_q^2}{2\sigma_r^2} \right)$$  \hspace{1cm} (4)$$
Then the filter result $BF$ can be expressed as:

$$BF = \frac{1}{W_q} \sum_{p \in S} G_i(p)G_j(p) * I_p$$  \hspace{1cm} (5)

The sum of the weights of all the pixel values $W_q$ used for filtering $\frac{1}{W_q}$ is the normalization of the weights.

In areas far from the edge, the pixel value weights of each pixel point in the filter are similar, and the distance weights play an important role in the filter. In the edge area, pixel points on the same side of the edge have similar pixel value weights and are much larger than those on the other side of the edge. Therefore, it is difficult for non-ipsilateral pixel points to influence the filter result and can also play a role in protecting edge information. The effect of bilateral filtering is shown in Figure 5.

![Figure 5. Image Fusion and Filtering Results.](image)

7. **Image Edge Pixel Fitting Optimization**

The purpose of this step is to fit the pixels near the edges of the filtered image to the unfiltered image, which can preserve the natural transition between the target and background caused by the bilateral filter, and overcome the drawbacks of losing image details and blurring the image after filtering. The filtered edge area pixels and the fitted image are shown in Figure 6.
8. Sections, subsections and subsubsections

To test the universality and validity of the algorithm in this paper, we use the algorithm in this paper to process different images to get the effect of foreground and background fusion. The following experiments are performed on Windows 1064 bit operating system, CPU Intel(R) Core(TM) i7-10710U CPU @ 1.10GHz, 16G memory, Python 3.8.9 programming language, and Visual Studio Code. The experimental results are shown in Figure 7.

From the histogram distribution of Figure 7, it can be seen that the initial images of Figure 8 are similar and belong to the images of simple background. The experimental results of these two graphs show that the algorithm can achieve ideal results for target extraction and target fusion of single and multiple objects under simple background.
From the histogram distribution of Figure 8, it can be seen that the initial image in Figure 7(1) and Figure 7(3) is not of the same type, the initial image in Figure 7(1) is of a simple background, the initial image in Figure 7(3) is of a complex background, and the target image in Figure 7(1) and Figure 7(3) can be seen whether it is a simple background image or a complex background image. The algorithm can achieve good results.

![Figure 8. Histogram Comparison of Simple Background.](image1)

The histogram distribution of Figure 9 shows that the initial images of Figure 7(3) and Figure 7(4) belong to the complex background images, which are distinguished as the background of Figure 7(3) is a monochrome gradient background, and the background of Figure 7(4) is a natural scene. From the result images of Figure 7(3) and 7(4), it can be seen that the algorithm can still achieve good results for images with a large color distribution span.

The experimental results show that the algorithm proposed in this paper has strong robustness and applicability. Target extraction can be achieved for different target segmentation images, and pixels in the edge area can be filtered to make the target and background combined with a natural transition effect.

![Figure 9. Histogram Comparison of Complex Background.](image2)
9. Conclusions
Based on the above mathematical models and experimental results above, this paper an edge feature light-shadow fusion algorithm based on image segmentation. In order to extract the target from the image and fuse it with the new background, there are some problems such as cumbersome steps, high operation difficulty, lack of fast and effective target acquisition and feature fusion methods, an edge light shadow feature fusion algorithm based on image area segmentation technology is proposed, which uses image segmentation, edge detection, morphological expansion, and so on. Image processing methods, such as filtering, achieve the effect of extracting target objects and natural fusion through simple operations. The experimental results show that the algorithm can be applied to images of simple objects under simple and complex backgrounds, and it can produce good results. However, for the extraction of target objects, it is the interactive picture frame selection. The number of iterations of the algorithm determines the effect of the extraction of target objects. For images with clear foreground and foreground, the effect is obvious. For images with complex foreground and background and unclear edges, the number of iterations should be selected multiple times to get the best results. In order to solve these problems, it is necessary to automatically select the appropriate number of iterations according to the complexity of the image and the image with unclear foreground and background, so as to make the target extraction more accurate and the synthesis effect more obvious. Since the Grabcut method of interactive feature area extraction is used to obtain the target segmentation area, in the following research, we will make the edge feature fusion algorithm of image segmentation more versatile by exploring the algorithm of image feature detection and semantic segmentation based on neural network, and overcome the difficulties of image segmentation caused by complex environmental factors.

Acknowledgments
This work was supported by The National Key Research and Development Program of China (No. 2018YFB1004904); Foundational Science and Technology Project of Wenzhou (No. G20200025).

References
[1] Nunez J, Otazu X, Fors O, et al. Multiresolution-based image fusion with additive wavelet decomposition[J]. IEEE Transactions on Geoscience and Remote Sensing, 1999, 37(3):1204-1211.
[2] Rother C. GrabCut : Interactive foreground extraction using iterated graph cut[J]. Acm Trans Graph, 2004, 23.
[3] Boykov Y. Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images[J]. Proc of IEEE Int'l Conf an Computer Vision, 2001.
[4] Pajares G, Cruz J. A wavelet-based image fusion tutorial[J]. Pattern Recognition, 2004, 37(9):1855-1872.
[5] Canny J. A computational approach to edge detection[J]. IEEE Trans. Pattern Anal. & Mach. Intell, 1986.
[6] Liu W, Li S. Multi-morphology image super-resolution via sparse representation[J]. Neurocomputing, 2013, 120(Complete):645-654.
[7] Raman S, Chaudhuri S. Bilateral Filter Based Compositing for Variable Exposure Photography. 2009.
[8] Linlong Fan, Yi Li, Xiaojin Zhang. Generative Adversarial Network Based Local Facial Stylized Generative Algorithm[J]. Journal of Graphics, Vol.42 No.1, 2021, ISSN 2095-302X, CN 10-1034/T, 2021
[9] Charbonnier, P, Blanc-Feraud, et al. Deterministic edge-preserving regularization in computed imaging[J]. IEEE Transactions on Image Processing, 1997.
[10] Charbonnier, Blanc-Feraud, Aubert, et al. Deterministic edge-preserving regularization in computed imaging[J]. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society, 1997, 6(2):298-311.
[11] Huang F, Yang G, Song K. The Soft Morphology Applied to Detecting Image Edge. JOURNAL OF IMAGE AND GRAPHICS, 2000.
[12] Zhou W, Li Y. Feature Fusion Detector for Semantic Cognition of Remote Sensing[J]. arXiv e-prints, 2019.
[13] Nunez, Jorge, Otazu, et al. Multiresolution-Based Image Fusion with Additive Wavelet Decomposition.[J]. IEEE Transactions on Geoscience & Remote Sensing, 1999.
[14] Huang Q, Wen G, Cai W. Thresholding technique with adaptive window selection for uneven lighting image[J]. Pattern Recognition Letters, 2005, 26(6):801-808.
[15] Ramprasad R, Rajapriya C, Mca S S. Biometric Iris Recognition System Using Edge Detection and Gabor Filter for High Security[J]. Advances in computational sciences and technology, 2019, 12(1):7-21.