IMPROVED SEAM CARVING FOR IMAGE RETARGETING WITH SIFT FEATURE PRESERVATION

Ke Li, Qianxu Zeng, Bahetiyaer Bare, Bo Yan

School of Computer Science
Shanghai Key Laboratory of Intelligent Information Processing
Fudan University, Shanghai, China

Hamid Gharavi

National Institute of Standards and Technology
100 Bureau Drive, Gaithersburg, MD USA 20899

ABSTRACT

This paper presents an effective and simple image resizing method. Our method is an improved version of seam carving that changes the backtracking basis of seam carving. We use a Scale Invariant Feature Transform (SIFT) feature in our method. SIFT key points are mainly located on high-contrast regions of an image. By using saliency considering SIFT as our backtracking basis, a resized image can preserve the SIFT features of the original image. Therefore, the resized image can be more visually acceptable than the image resized by traditional seam carving.

Index Terms— image resizing, seam carving, SIFT.

1. INTRODUCTION

With the rapid development of diversity and versatility of display devices, the technique of image resizing has played an increasingly important role. There exist many categories of content-aware image resizing methods[1], which have advantages as well as limitations. For instance, a method aimed at energy preservation will remove the pixels with the lowest energy in ascending order. However, it will destroy the rectangular shape of the image as we may remove a different number of pixels from each row. Also, there is a method that removes an equal number of low energy pixels from each row. This preserves the rectangular shape of the image, but will create a zigzag effect which results in destruction of the image.

Image retargeting aims at changing the resolution of the image whilst maintaining its important content. Due to its important application in consumer electronics, it has recently been receiving considerable attention from researchers. Avidan proposed an image retargeting method named seam carving [2][3], which removes pixels in a judicious manner for content-aware resizing. They demonstrated that effective resizing of images should not only use geometric constraints, but also consider the image content. Seam carving is a simple operator that supports content-aware image resizing for both reduction and expansion. A seam is an optimal 8-connected path of pixels on a single image from top to bottom, or left to right, where optimality is defined by an image energy function. By continuously carving-out or inserting pixels in different parts of the image, we can resize the image into target resolution. For image reduction, we remove more of the low energy pixels and fewer of the high energy so that the information of the image can be preserved gracefully.

SIFT keypoints are mainly located on high-contrast regions of an image, such as texture[4]. One of seam carving limitations is it cannot preserve the SIFT features of the image, which may result in many problems, such as visually unacceptability of the resized image and destruction of the image's important regions.

To improve the performance of seam carving, we focus on the SIFT feature of an image. The SIFT feature is invariant to image translation, rotation, scaling, and change in illumination. Therefore, if a resized image is able to keep the SIFT features of the original image, the resized image will be visually acceptable. Although traditional seam carving can preserve the image content better than many other strategies, it cannot preserve the SIFT features.

In this paper, we propose to use the saliency information as seam carving energy function to find all seams in an image, and combine SIFT information and saliency information to select the least important one. By doing this, the resized image can be much more visually acceptable than that produced by the original seam carving algorithm.

The remainder of this paper is organized as follows. In section 2, we will review the basic methods of seam carving and SIFT feature preservation, including some related works. Then we will introduce our improved seam carving approach in section 3. In section 4, we will compare the results of seam carving and our method. We will show the objective assessment scores of the results. Finally, we will draw conclusions.

This work is supported by NSFC (Grant No.: 61370158).

2. REVIEW OF RELATED WORK

2.1. Seam Carving

Seam carving [2] is a useful and effective content-aware resizing method to remove pixels in a judicious manner. The most important thing is how to choose the pixels to be removed. This leads to the following energy function [2]:

$$e_1(I) = |\frac{\partial}{\partial x}I| + |\frac{\partial}{\partial y}I| \quad (1)$$

Let I be an $n \times m$ image and define a vertical seam to be:

$$s^x = \{s_i^x\}_{i=1}^n = \{(x(i), i)\}_{i=1}^n, s.t. \forall i, |x(i) - x(i-1)| \leq 1 \quad (2)$$

where x is a mapping x: $[1, \cdots, n] \rightarrow [1, \cdots, m]$, which means a vertical seam is an 8-connected path of pixels in the image from top to bottom. Each row of the image only contains one pixel [2].

Given an energy function e, seam carving defines the cost of a seam as [2]:

$$E(s) = E(I_s) = \sum_{i=1}^{n} e(I(s_i)) \quad (3)$$

The optimal seam s^* is the seam that minimizes this seam cost [2]:

$$s^* = \min_s E(s) = \min_s \sum_{i=1}^{n} e(I(s_i)) \quad (4)$$

Seam carving uses dynamic programming to find the optimal seam. The first step is to traverse the image from the second row to the last row and compute the cumulative minimum energy N for all possible connected seams for each entry (i, j) [2]:

$$N(i,j) = e(i,j) + \min(N(i-1,j-1), N(i-1,j), N(i-1,j+1)) \quad (5)$$

At the end of the process, the minimum value of the last row in N will indicate the end of the minimal connected vertical seam. In the second step, seam carving backtracks from this minimum entry on N to find the path of the optimal seam. Thus, we know which seam should be removed [2].

2.2. SIFT Feature Preservation

SIFT keypoints [5] are mainly located in important regions or high-contrast regions of an image. SIFT consists of four major stages: scale-space peak selection; keypoints localization; orientation assignment and keypoint descriptor.

The first step is to build the scale space of an image. It is an initialization operation aimed at simulating the image's multi-scale features. By scanning all possible scales and image locations, potential interest points will be identified. Gaussian function will be used in this step, as it is the only possible scale-space kernel. Then a two-dimensional image's scale space can be defined as:

$$L(\mathrm{x}, \mathrm{y}, \sigma) = \mathrm{G}(\mathrm{x}, \mathrm{y}, \sigma) * \mathrm{I}(\mathrm{x}, \mathrm{y}) \quad (6)$$

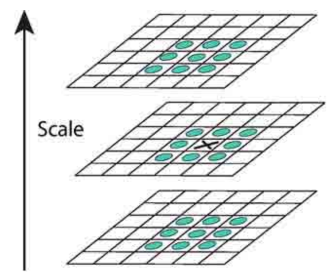

Fig. 1: Maxima and minima of the DoG images are detected by comparing a pixel(X) to its 26 neighbors in 3×3 regions at the current and adjacent scales [5].

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (7)$$

where $I(x, y)$ is a convolution function of an image. $G(x, y, \sigma)$ is a variable-scale Gaussian function; (x, y) is the coordinate and σ is the feature scale which determines the image's smoothness [5].

From here a series of difference-of-Gaussian (DoG) images are established. The DoG function provides a close approximation to the scale-normalized Laplacian of Gaussian, $\sigma^2\nabla^2 G$:

$$\sigma^2\nabla^2 G = \frac{\partial G}{\partial \sigma} = \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma} \quad (8)$$

The maxima and minima of $\sigma^2\nabla^2 G$ produce the most stable image features, which are among a range of image functions [5].

The second step is to localize candidate keypoints to sub-pixel accuracy. If a keypoint is found to be unstable, it will be eliminated. In the third step, the dominant orientation of each keypoint will be identified according to its local image patch. Finally, a local image descriptor for each keypoint will be built based on the image gradients in its local neighborhood. A SIFT descriptor is a vector in 128 dimensions.

Research focusing on SIFT features has caught more and more people's attention. For instance, the SIFT descriptor characterizes a SIFT region invariantly to image translation, scaling, rotation and change in illumination. By using SIFT descriptors, Yue *et al*. [6] proposed an image compression scheme to make use of external image contents to reduce visual redundancy among images. Kaze *et al*. [7] proposed an image resizing method in order to preserve SIFT features, which is based on warping method [8].

SIFT feature preservation can make contributions to image resizing, as it can help preserve the important contents of an image. Thus, we highlight the importance of SIFT features in our method.

(a) original image (b) saliency (c) SIFT saliency

Fig. 2: A example of an image's SIFT saliency.(a) the image; (b) the saliency map calculated by [9]; (c) the saliency considering SIFT regions.

3. SEAM CARVING WITH SIFT FEATURE PRESERVATION

The most important thing of our approach is how to choose the pixels to be removed. We use an approach similar to what used in seam carving. We also use dynamic programming to find seams. However, we will find all the seams that start from pixels in the first row. Then we will choose one seam to be removed. The challenge is which one should be removed.

Our method considers the SIFT features of the image. For each keypoint ν, we calculate a radius r_ν using the scale of ν. Then we get a circle for each keypoint with the radius and location. We define the square, which is tangent to the circle of one keypoint, as the "SIFT region". Then the saliency map M, indicating the importance of pixels, would be obtained by using effective method [9].

In order to reflect the concept of "SIFT region", we calculate the average saliency s_ν of each "SIFT region". That's to say, we add up every pixel's saliency then divide it by the number of pixels in this region:

$$s_v = \frac{\sum_{i=\lceil x-r_v \rceil}^{\lceil x+r_v \rceil} \sum_{j=\lceil y-r_v \rceil}^{\lceil y+r_v \rceil} M_{ij}}{(\lceil x+r_v \rceil - \lceil x-r_v \rceil) \times (\lceil y+r_v \rceil - \lceil y-r_v \rceil)} \quad (9)$$

where x and y are the horizontal and vertical coordinates of ν respectively.

In order to avoid large distortion, we use a constraint here. Only for SIFT regions whose corresponding radius r_ν is less than 10, will the saliency be recalculated. Then, the saliency of each pixel in this region will be changed to the average value. If a pixel is included in more than one "SIFT region", we choose the maximum of all the average saliencies of the SIFT regions that cover the pixel as the pixel's saliency. If a pixel is not covered in any "SIFT region", the saliency of the pixel will not be changed. For each "SIFT region", we repeat the above steps and get a new saliency map M'. We call the new saliency of an image "SIFT saliency"(Fig. 2).

As we discussed above, we get all the seams S^{all} by equation (5) and the saliency of each pixel. Now we should choose one seam S^* to be removed. For each seam, we add up saliency of each pixel on this seam. We define a new function E_s to measure the importance of k^{th} seam S^k

$$E(S^k) = \sum M'_{ij} \ , \ .s.t \ pixel \ (i,j) \ \in S^k \quad (10)$$

where k is the index of the seam. (i,j) is the coordinate of the pixel on the k^{th} seam. The S^* could be found by minimizing the seam cost:

$$S^* = \min_{S^k \subset S^{all}} E(S^k) \quad (11)$$

Then we repeat the steps above. Every time we choose one seam with minimum saliency to be removed. Finally, we will get the image on the target resolution, which is much more visually acceptable than the result of seam carving.

Fig. 3: Results of different methods and $W'/W = 0.7$. The images from left to right:original, results of [2] and ours.

	$SeamCarving$	$OurMethod$	$Gain(\%)$
Fish	0.78	0.80	2.56
Surfer	0.81	0.87	7.92
Family	0.85	0.87	2.64

Table 1: Scores when images' width is reduced to 70%

Fig. 4: Results of different methods and $W'/W = 0.6$. The images from left to right:original, results of [2] and ours.

	$SeamCarving$	$OurMethod$	$Gain(\%)$
Fish	0.70	0.73	3.92
Surfer	0.73	0.79	8.14
Family	0.78	0.81	3.70

Table 2: Scores when images' width is reduced to 60%.

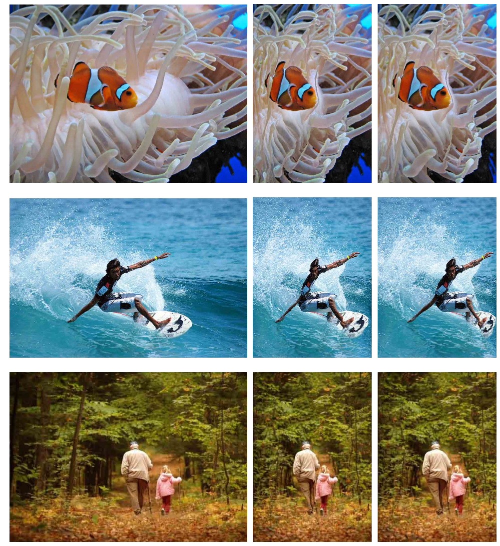

Fig. 5: Results of different methods and $W'/W = 0.5$. The images from left to right:original, results of [2] and ours.

	$SeamCarving$	$OurMethod$	$Gain(\%)$
Fish	0.62	0.65	4.01
Surfer	0.66	0.68	4.00
Family	0.72	0.74	2.78

Table 3: Scores when images' width is reduced to 50%.

4. RESULTS AND ASSESSMENT

In order to evaluate the performance of our proposed method, we test it on the RetargetMe image retargeting benchmark provided by [10]. Without loss of generality, we only change the width of the images in our results. Let W and W' represent the width of original image and retargeted image respectively. We have 3 groups of results with different values of W'/W: Figure 3 ($W'/W = 0.7$), Figure 4 ($W'/W = 0.6$) and Figure 5 ($W'/W = 0.5$). The results of seam carving are shown as comparison to our method. We will try to analyze our results from subjective visual perception and objective assessment.

As shown in Figure 3 to Figure 5, most of the results of our method are much more visually acceptable than the results of seam carving. For instance, in the resized image of the image named $Fish$, the fish is the most important content of the images and in our results are less deformed compared with the results of seam carving. As we can observe, when images are resized to half the original width (Figure 5), seam carving removed too much seams through the fish which makes the fish look much shorter than the original, while our result looks much better as our method removed less seams through the fish. Also, our results for the image named $Surfer$ are more natural than the results of seam carving.

In addition, we got objective assessment for the 3 result groups with different targeted widths by using the method in [11]: Table 1($W'/W = 0.7$), Table 2($W'/W = 0.6$) and Table 3($W'/W = 0.5$). These scores indicate that the results of our method are better than seam carving in objective evaluation. The proportional gain can be up to 8.14%.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed discrete image resizing method for preserving SIFT features. We changed the strategy for choosing seams of seam carving, which allows the seam with the lowest SIFT saliency to be removed. We imposed a constraint on the matrix of SIFT saliency in order to avoid large distortion. Most of our experimental results demonstrated that our method is much more visually acceptable than seam carving as the SIFT features of the image were preserved. In addition, our method outperforms seam carving in terms of objective assessment. More importantly, our method could preserve SIFT features, which could be used for image retrieval and image classification.

6. REFERENCES

[1] Ariel Shamir and Olga Sorkine, "Visual media retargeting," in *ACM SIGGRAPH ASIA 2009 Courses*, 2009, pp. 11:1–11:13.

[2] Shai Avidan and Ariel Shamir, "Seam carving for content-aware image resizing," in *ACM SIGGRAPH 2007 Papers*, 2007, SIGGRAPH '07.

[3] Michael Rubinstein, Ariel Shamir, and Shai Avidan, "Improved seam carving for video retargeting," in *ACM SIGGRAPH 2008 Papers*, 2008, pp. 16:1–16:9.

[4] Ran Gal, Olga Sorkine, and Daniel Cohen-Or, "Feature-aware texturing," in *Proceedings of the 17th Eurographics Conference on Rendering Techniques*, 2006, pp. 297–303.

[5] David G Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[6] Huanjing Yue, Xiaoyan Sun, Feng Wu, and Jingyu Yang, "Sift-based image compression," in *2012 IEEE International Conference on Multimedia and Expo (ICME)*, July 2012, pp. 473–478.

[7] Kazu Mishiba and Takeshi Yoshitome, "Image resizing with sift feature preservation.," in *IEEE International Conference on Image Processing (ICIP)*, Sept 2013, pp. 991–995.

[8] Yanwen Guo, Feng Liu, Jian Shi, Zhi-Hua Zhou, and M. Gleicher, "Image retargeting using mesh parametrization," *IEEE Transactions on Multimedia*, pp. 856–867, Aug 2009.

[9] Jonathan Harel, "A saliency implementation in matlab: `http://www.klab.caltech.edu/~harel/share/gbvs.php`," .

[10] Michael Rubinstein, Diego Gutierrez, Olga Sorkine, and Ariel Shamir, "A comparative study of image retargeting," in *ACM SIGGRAPH Asia*, 2010, pp. 160:1–160:10.

[11] Yong-Jin Liu, Xi Luo, Yu-Ming Xuan, Wen-Feng Chen, and Xiao-Lan Fu, "Image retargeting quality assessment," *Computer Graphics Forum*, pp. 583–592, 2011.