FAST STRUCTURE-PRESERVING IMAGE RETARGETING

Shu-Fan Wang and Shang-Hong Lai

Department of Computer Science
National Tsing Hua University
HsinChu, Taiwan 300 R.O.C

ABSTRACT

Several different methods have been proposed for image/video retargeting while retaining the content. However, they sometimes produce some artifacts, such as ridge or structure twist. In this paper, we present a structure-preserving image resizing technique for the image retargeting applications. Based on the warping-based retargeting technique proposed by Wolf et al.[13], we propose an efficient and adaptive image resizing algorithm that preserves the content and image structure as best as possible. We first downsample the size of the original image by using bilinear interpolation. In order to preserve the content, we introduce the structure constraints derived from the line detection into the large linear system. Then, the mapping matrices are enlarged to the original size by joint-bilateral upsampling and the resized image can be produced to preserve the content and structure as best as possible. Most of the computation is on the low-resolution layer and therefore it can be very efficient. From our experiments, the proposed method can provide resized images with higher image quality and faster speed than that in [13].

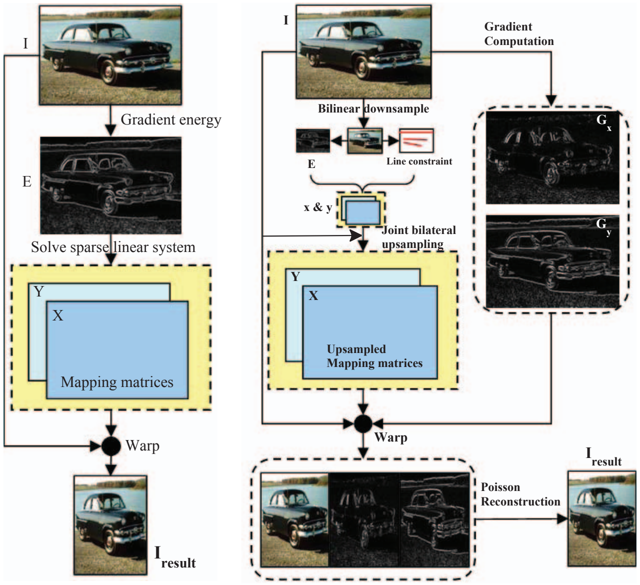
***Index Terms*—** retargeting, structure-preserving, joint-bilateral

1. INTRODUCTION

Content-aware image/video retargeting becomes very important for image display on a variety of display devices of different resolutions or aspect ratios. With the increasing need of retargeting, several methods are proposed to adapt image content to various display settings. In 2003, Chen et al. [3] proposed to automatically detect and transmit the most important region to the mobile device and the system in [4] retargets video to adapt content to small devices. They applied cropping to remove less important regions from the images. Later, Liu and Gleicher [8] tried to find the Region-Of-Interest (ROI) and then applied a linear scaling function to maintain the ROI when the image was warped to fit the specified resolution. They also extended their work to video retargeting [9]. Gal et al. [5] proposed to warp an image to an arbitrary shape while preserving important features. It is achieved by a Laplacian technique to accommodate local constraints followed by a global optimization manner. Recently, Avidan and Shamir [1] proposed an idea of incrementally removing or inserting regions. An 8-connected path of least importance pixels is incrementally removed or inserted to resize an image. However, simply extending seam carving to video retargeting will create jittery artifacts. Therefore, they improved the seam carving algorithm that introduces forward and backward energy to reduce the effect of artifacts [11]. Chen and Sen [2] extended this idea to summarize video into a short segment by removing 2D manifold in space-time volumn. In 2007, Wolf et al. [13] presented a method automatically detect the important region by combining a saliency measure, face detector and motion estimation for video retargeting. They formulated the grid mapping of image resizing as solving a large sparse linear system. This constraint warping scheme is also compared with seam carving and shows better structure-preserving property. In human visual system (HVS), human eyes are very sensitive to some specific shapes. However, most previous works on image/video retargeting did not consider the preservation of global image structure. In this paper, we propose a new image retargeting algorithm that not only improves the resized image quality but also considerably shortens the processing time. The rest of this paper is organized as follows. In section 2, we will describe the proposed algorithm that improves the quality of the resized image as well as the computational efficiency. We validate the effectiveness of the proposed method through qualitative and quantitative experimental results in section 3. The final section concludes this paper.

2. IMPROVED CONTENT-AWARE IMAGE RESIZING

In this section, we will present the proposed retargeting algorithm in details. The main difference between the warping-based retargeting method [13] and the proposed algorithm is shown in Fig.1. In the following subsections, we will briefly review the retargeting algorithm proposed in [13]. Then, we will describe how to preserve the global geometric structure and the quality of resized images. Finally, we will present how to improve the computational efficiency in the proposed algorithm.



(a) Wolf et al.[13] (b) The proposed system

Fig. 1. The flow chart of the method in [13] and our system.

2.1. Warping-based image retargeting

The formulation for the image retargeting problem is cast as solving a constrained linear system in [13]. The task is to recover the new coordinate $(X_{i,j}, Y_{i,j})$ of pixels (i, j) under three types of constraints in this application. Take the calculation of $X_{i,j}$ for instance. First, each pixel is supposed to be at a fixed distance from its left and right neighbors: $X_{i,j} - X_{i-1,j} = 1$ and $X_{i+1,j} - X_{i,j} = 1$. The second step is to map each pixel to a location similar to the one of its upper and lower neighbors: $X_{i,j} - X_{i,j+1} = 0$. The third constraint fits the warped image to the dimensions of the target image size: $X_{1,j} = 1$ and $X_{W,j} = W_{target}$.

For content-aware image resizing, an important pixel should be warped to a distance approximately one pixel from its neighbors while less important ones can be blended with their neighbors or be carved off. Therefore, the first two constraints mentioned above should be weighted by the corresponding energy value. Here, the $E_{i,j}$ denote the local gradient energy of the input image at location (i, j):

$$\begin{aligned} &E_{i,j}(X_{i,j} - X_{i-1,j}) = E_{i,j} \\ &E_{i,j}(X_{i+1,j} - X_{i,j}) = E_{i,j} \\ &E_{i,j}(X_{i,j} - X_{i,j+1}) = 0 \end{aligned} \tag{1}$$

All the equations form an over-determined constrained sparse linear system. The optimized new coordinates of the pixels can be obtained by minimizing the sum of squared errors of the above equations, which is equivalent to finding the least-squares solution of the sparse linear system:

$$\begin{aligned} &A\mathbf{x} \approx b \\ &\Rightarrow \mathbf{x} = (A^T A)^{-1} A^T b \end{aligned} \tag{2}$$

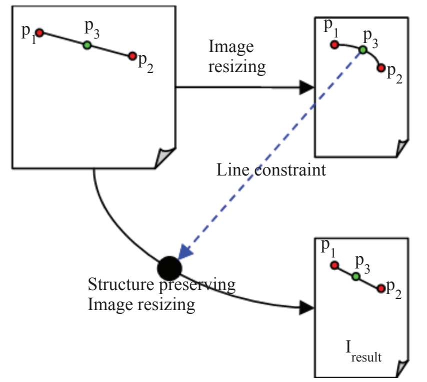


Fig. 2. The inclusion of a line constraint.

Similarly, the coordinate variables $Y_{i,j}$ of pixels (i, j) can also be obtained from the least-square solution of following equations:

$$\begin{aligned} &E_{i,j}(Y_{i,j} - Y_{i,j-1}) = E_{i,j} \\ &E_{i,j}(Y_{i,j+1} - Y_{i,j}) = E_{i,j} \\ &E_{i,j}(Y_{i,j} - Y_{i+1,j}) = 0 \\ &Y_{i,1} = 1, Y_{i,H} = H_{target} \end{aligned} \tag{3}$$

2.2. Line constraint for structure preserving

The retargeting algorithm described in the previous subsection does not consider the preservation of global geometric structure in the original image. However, human vision system is actually sensitive to global image structure, like a line. To preserve the line structure in the resized image, we first apply Hough transform to detect the straight lines in the image and add the associated constraints into the constrained linear system. Before adding the line constraints, we first apply the original retargeting process (i.e. Wolf et al. [13]) to obtain the new locations of the end points of the line segments. Afterwards, the points on the line segments should be constrained on a straight line which is defined by the end points in the resized image. Therefore, we can obtained the new locations of the line pixels and add them into the constrained sparse linear system. Please see Fig.2 for an example, p_1 and p_2 are the end points of a detected line segment and p_3 is the pixel lie on this line segment. After image resizing, the resized line segment can be determined by the new positions of p_1 and p_2. Therefore, the new coordinate of p_3 is determined to be a constraint for line structure preservation. Here, we adaptively determine all the sampled pixels in all line segments as the structure constraints. By this way, we can effectively preserve the line structure in the resized image.

2.3. Efficiency improvement

The Hough transform and two-stage retargeting process involved in the proposed algorithm require more processing time than that in [13]. Generally, the most effective way to

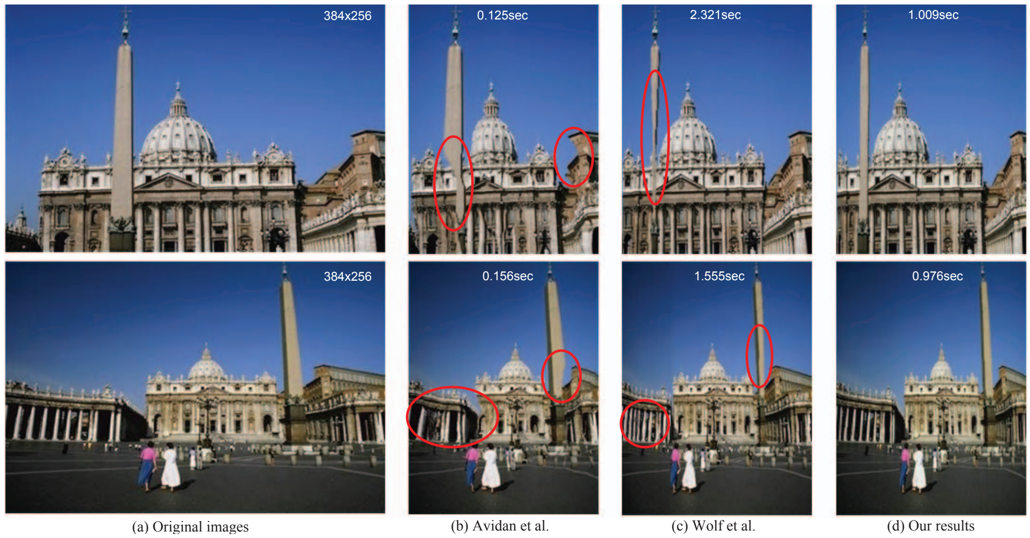


Fig. 3. The quality and processing time comparison among different content-aware image resizing techniques.

reduce computational cost is to process on a subsampled image and try to upsample the resized image back. However, naive subsampling, resizing and upsampling can not obtain a resized image with correct content structure. In this paper, we follow the basic idea of subsampling the original image and find the X and Y mapping matrices by solving the linear system in Eq.2. Instead of upsampling the resized image, we upsample the grid X and Y by a content-driven method; namely, the joint bilateral upsampling technique [6].

The traditional bilateral filtering is an edge-preserving filter which uses both a domain filter kernel and a range filter kernel evaluated on the data values themselves. Recently, a concept of joint bilateral filtering was proposed such that the range filter is applied to another guidance image. Kopf et al. [6] extended the idea to upsampling, which can be used here to upsample the coordinate matrices X and Y. Assume we have a higher-resolution image $\tilde{I}$ and a low-resolution coordinate matrix X, the upsampled $\tilde{X}$ can be computed by:

$$\tilde{X}_p = \frac{1}{k_p} \sum_{q_\downarrow \in \Omega} X_{q_\downarrow} f(\|p_\downarrow - q_\downarrow\|) g(\left\|\tilde{I}_p - \tilde{I}_q\right\|) \quad (4)$$

where p and q denote the coordinates of pixels in $\tilde{I}$ and $p_\downarrow$ and $q_\downarrow$ denote the corresponding coordinates in the low-resolution coordinate matrix X. Similarly, the low resolution coordinate matrix Y can also be upsampled to the target size by joint bilateral upsampling.

In some cases, the retargeting process still creates noticeable visual artifacts in the resized image if the color gradually changes in the original scene. Therefore, similar to [1], our algorithm is performed in the gradient domain to resize the x and y derivatives and then reconstruct the resized color image by using the Poisson reconstruction technique [10].

3. EXPERIMENTS

In this section, we validate the proposed method by conducting several experiments for quality and efficiency verification. In the experiments, we retarget the images from the dataset [7, 12], which contains a wide variety of images. First, we will compare the results from seam carving [1], the warping-based method by Wolf et al. [13] and the proposed method on strongly structured images. The second experiment is conducted to show some cases of image retargeting with visible artifacts in [13], but the proposed algorithm can produce more satisfactory results very efficiently. In our experiment, the seam-carving method [1], is implemented in Visual C++, the warping-based method by Wolf et al. [13] and the proposed method are implemented in Matlab. The following experiments are conducted on a PC with Intel 1.86GHz CPU and 2048MB RAM.

3.1. The resizing of structured images

Objects presented in an image can be classified into two categories, natural or artificial. Man-made objects are often structured, such as the building in Fig.3. We resized the images of size 384x256 to 192x256 with the three different methods and compare the quality of the resized images and their efficiency. From the second column in Fig.3, we can see that the seam carving [1] destroyed the original image structure severely although it always removes the seam with minimum energy. The warping-based method by Wolf et al. [13] also produced noticeable artifacts in the image structure. Obviously, both methods can not preserve the line structure well. It is obvious from Fig. 3 that the resized images by using the proposed algorithm outperforms those of the other two methods.

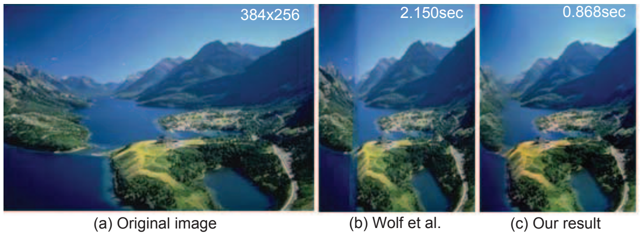


Fig. 4. The image containing smooth color variations usually leads to artifacts after resizing [13].

3.2. The resizing of images with gradual color change

In a natural landscape, the color may change gradually, such as sky, sea or shadow. Fig.4 depicts an example that color changes gradually across the sky and the lake. The image resized by using [13] contains noticeable visual artifacts. Since we resize images in the gradient domain and apply the Poisson reconstruction in the last step, our results are free from this kind of artifacts.

3.3. The issue of computational efficiency

We measure the processing time in all the experiments which is also shown in Fig.3 and Fig.4. Seam carving [1] shows the highest efficiency because of the simple calculations involved in the dynamic programming. However, it is not acceptable that the global image structure is severely destroyed. In overall, the proposed retargeting algorithm produces the best image quality and its computational speed is at least twice faster than the previous warping-based method [13]. When the size of the original image becomes larger (larger than 500×500), the processing speed can be up to five times faster than that in [13] and it can still maintain high image quality and global image structure.

4. CONCLUSION

In this paper, we proposed an efficient image resizing algorithm to retarget images to different sizes with the global image structure preserved. In the proposed algorithm, we first detect straight lines and introduced line constraints by computing the new coordinates of the pixels in the lines to enhance the line structure preserving capability. The main computational cost in the proposed resizing process is to solve a large and sparse linear system. This is reduced by performing the resizing process in a low-resolution space, thus improving the computational efficiency. In addition, we employed the content-driven joint bilateral filter to upsample the mapping matrices and obtain the final resized image. The experimental results demonstrate the improved image quality and efficiency of the proposed image resizing algorithm.

5. REFERENCES

[1] S. Avidan and A. Shamir. Seam carving for content-aware image resizing. *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, 2007.

[2] B. Chen and P. Sen. Video carving. *Eurographics*, 2008.

[3] L. Q. Chen, X. Xie, X. Fan, W. Y. Ma, H. J. Zhang, and H. Q. Zhou. A visual attention model for adapting images on small displays. *ACM Multimedia Systems Journal*, 9(4):353–364, 2003.

[4] X. Fan, X. Xie, H.-Q. Zhou, and W.-Y. Ma. Looking into video frames on small displays. In *MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia*, pages 247–250, New York, NY, USA, 2003. ACM.

[5] R. Gal, O. Sorkine, and D. Cohen-Or. Feature-aware texturing. In *Proceedings of Eurographics Symposium on Rendering*, pages 297–303, 2006.

[6] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele. Joint bilateral upsampling. *ACM Trans. Graph.*, 26(3):96, 2007.

[7] J. Li and J. Z. Wang. Automatic linguistic indexing of pictures by a statistical modeling approach. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25:1075–1088, 2003.

[8] F. Liu and M. Gleicher. Automatic image retargeting with fisheye-view warping. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 153–162, New York, NY, USA, 2005. ACM.

[9] F. Liu and M. Gleicher. Video retargeting: automating pan and scan. In *MULTIMEDIA '06: Proceedings of the 14th annual ACM international conference on Multimedia*, pages 241–250, New York, NY, USA, 2006. ACM Press.

[10] P. Perez, M. Gangnet, and A. Blake. Poisson image editing. *ACM Trans. Graph.*, 22(3):313–318, 2003.

[11] M. Rubinstein, A. Shamir, and S. Avidan. Improved seam carving for video retargeting. *ACM Trans. Graph.*, 27(3):1–9, 2008.

[12] J. Z. Wang and G. Wiederhold. Simplicity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:947–963, 2001.

[13] L. Wolf, M. Guttmann, and D. Cohen-Or. Non-homogeneous content-driven video-retargeting. *Proceedings of the Eleventh IEEE International Conference on Computer Vision (ICCV-07)*, 2007.