Shafieyan, F., Karimi, N., Mirmahboub, B., Samavi, S., & Shirani, S. (2014, October). Image seam carving using depth assisted saliency map. In *Image Processing (ICIP), 2014 IEEE International Conference on* (pp. 1155-1159). IEEE.

IMAGE SEAM CARVING USING DEPTH ASSISTED SALIENCY MAP

F. Shafieyan[1], *N. Karimi*[1], *B. Mirmahboub*[1], *S. Samavi*[1, 2], *S. Shirani*[2]

Department of Electrical and Computer Engineering
[1]Isfahan University of Technology, Iran
[2]McMaster University, Canada

ABSTRACT

Retargeting algorithms are needed to transfer an image from a device to another with different size and resolution. The goal is to preserve the best visual quality for important objects of the original image. In order to reduce image size, pixels should be removed from less important parts of the image. Therefore, we need an energy function to select less important pixels in seam carving. Various energy functions have been proposed in previous works to minimize the distortion in salient objects. In this paper we combine three different importance maps to form a new energy map. We first use both gradient and depth maps to highlight the values in the saliency map, eventually generates the final energy map. Experimental results using the proposed energy map show better visual appearance in comparison to previous algorithms even at high resizing percentage. The visual artifacts that cause shape deformation in salient objects and deteriorates geometrical consistency of the scene are considerably reduced in our proposed algorithm.

Index Terms— Seam carving, saliency map, depth map, energy map, geometry preservation.

1. INTRODUCTION

An important issue in multimedia technology is to display an image on different devices with different aspect ratios and resolutions. These devices could be cell phones, cameras, and PDAs. Size of image must be changed to fit to the display while the content of image must be preserved with minimum distortion. Basic resizing methods such as scaling and cropping do not take the image content into consideration [1]. Therefore, they produce considerable distortion in the retargeted image. To solve this problem several content-aware methods are proposed for image retargeting. One of these methods is seam carving that is first proposed by Avidan *et al.* [2]. In this method a seam is defined as an 8-connected path of pixels with minimum energy that is calculated based on gradients. In each iteration a vertical or horizontal seam is calculated using dynamic programming and is removed from the input image, until the final desired size of image is reached. Rubinstein *et al.* [3] used graph cut to improve this method. They also combined different operators with optimum sequence of selected seams to produce an image with better visual quality [4].

In spite of all these improvements, the important objects in the resized image still suffer from distortions. Therefore other criteria, rather than gradient, are proposed to be used for image retargeting. Saliency criterion could distinguish important parts of the image. Achanta *et al.* [5] used saliency, instead of gradient, to compute the energy of pixels. It is said that gradient is sensitive to energy change in object edges and causes deformations. In their method the importance of pixels are determined based on color and intensity globally. Cho *et al.* [6] proposed a low complexity method where a removed pixel keeps affecting the energy of its previously neighboring pixels. They introduced an importance diffusion map that preserves visual quality of the entire image and prevents distortion in important objects. Domingues *et al.* [7] used an energy function based on saliency information and considerably reduced image distortion.

Image depth is another criterion that is beneficial in determining the importance of pixels. There are various methods to find the depth map [8-10]. Depth map can be determined manually by the user. Mansfield *et al.* [8] used this method and introduced scene carving. The user divides the input image to several layers based on existing important objects. The depth of each layer is its priority. Retargeting is done using the input image and user-defined depth map. This method produces different results based on different depth maps that are defined by the user. In another algorithm Shen *et al.* [9] introduced depth-aware single image seam carving that used Kinect depth camera and disparity map to estimate the exact depth information. In this method, depths of objects are used to define pixels' energies. They also used just noticeable distortion (JND) model for seam selection that improved seam carving. This method prevented important objects from distortion and had a good performance in producing retargeted image. Basha *et al.* [10] used disparity map from stereo image pairs to estimate the depth and carved the left and right images similarly. They removed piecewise continuous seams and kept the geometric consistency between stereo image pairs.

In this paper we propose an algorithm that indents to preserve salient objects, at different depths, against retargeting distortion. We introduce a new energy function that consists of three importance maps of gradient, depth and saliency. The gradient map detects strong edges. The depth map distinguishes background from foreground objects. The saliency map indicates visually important objects of the image. In order to emphasize on salient objects of the image, we use both gradient and depth maps in a preprocessing stage to enhance the saliency map. Then the energy function is built by addition of the three importance maps. This energy function assigns high energy to important parts of the image that causes less important pixels to be removed in seam carving. Therefore visual artifacts such as shape deformation and geometry distortion are reduced considerably in salient objects. The rest of the paper is organized in the following manner. Section 2 the proposed method is detailed. In Section 3 experimental results are presented. Section 4 concludes the paper.

2. PROPOSED METHOD

Seam carving method iteratively selects a vertical (horizontal) seam with minimum energy and removes it to reduce the width (height) of the image by one pixel. More precise energy function for pixels leads to more successful image resizing that preserves the important objects. A suitable energy function can decrease the visual artifacts in image. Original seam carving method uses the gradient to obtain energy map [2]:

$$GM(i,j) = \left|\frac{\partial}{\partial x} I(i,j)\right| + \left|\frac{\partial}{\partial y} I(i,j)\right| \tag{1}$$

where $I(i,j)$ is the intensity of a pixel in the input image and GM is the gradient map. The main idea of the proposed method is to use the gradient map, depth map and saliency map altogether. We want to use the important parts of all maps to preserve the salient objects in the image. The structure of the proposed algorithm is illustrated in Fig.1. Gradient map is used to detect the edges of important objects. But if an important object has low contrast in comparison to its surroundings, gradient map alone cannot prevent distortions. In Fig. 2(a) and Fig. 2(b) a sample input image and its corresponding gradient map are shown where the map is produced using Eq. 1.

Depth map is used to find the distance of objects in image from the camera. Near objects are usually more important than background objects. Therefore they are detected with depth map and are prevented from distortion. But sometimes important objects are scattered at different depths of image and some of them may be in background. In such cases, the depth map alone cannot detect the important objects and we need to use saliency map to better preserve them during seam carving. Figures 2(c) and 2(d) respectively illustrate the input depth map, DM, using [10], and the saliency map, SM, using [11], for the image of Fig. 2(a). In order to find a good energy function we try to combine three aforementioned energy maps in a way that common important regions in all of them have high values. This energy function should decrease visual artifacts in the retargeted image.

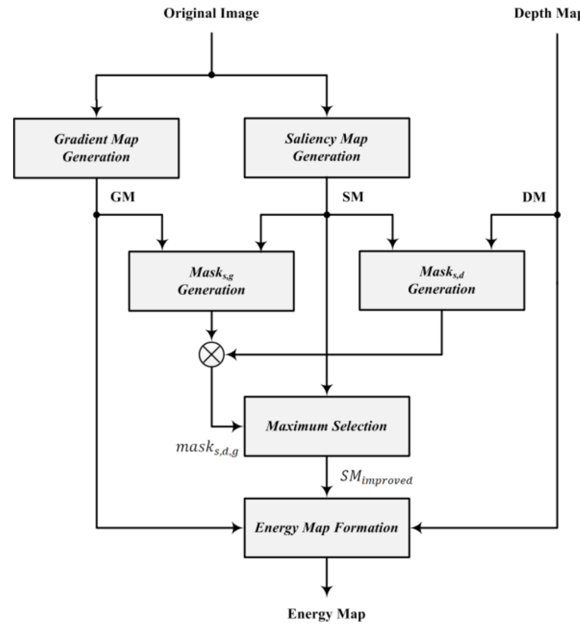

Fig. 1. Framework of proposed mask formation scheme.

Salient objects are usually near the camera. Therefore saliency map and depth map contain overlap in important regions. We define a binary mask $mask_{s,d}$ to indicate these important regions. Also salient objects usually have large gradients on their edges. Hence, we define a binary mask $mask_{s,g}$ to indicate the overlap between the saliency map and gradient map. To construct the proposed masks, we first calculate addition and subtraction between two relevant importance maps. Therefore to form $mask_{s,d}$ we define two intermediate maps of $A_{s,d}$ and $B_{s,d}$ as defined in Eq. 2 and Eq. 3.

$$A_{s,d}(i,j) = SM(i,j) + DM(i,j) \tag{2}$$

$$B_{s,d}(i,j) = |SM(i,j) - DM(i,j)| \tag{3}$$

The idea is that if a pixel has a large value in $A_{s,d}$, then it has a large value in either SM and/or in DM. If that pixel has small value in $B_{s,d}$, then its value should be large in both SM and DM and that pixel should be considered as an important pixel. Otsu [12] method is used to determine a threshold to separate large and small values in both $A_{s,d}$ and $B_{s,d}$. The value of $mask_{s,d}$ is obtained as below.

$$mask_{s,d} = \begin{cases} 1 & A_{s,d} > T^A_{s,d} \ and \ B_{s,d} < T^B_{s,d} \\ 0 & A_{s,d} < T^A_{s,d} \ or \ B_{s,d} > T^B_{s,d} \end{cases} \tag{4}$$

where $T^A_{s,d}$ and $T^B_{s,d}$ are the thresholds on addition and subtraction of saliency map and depth map respectively.

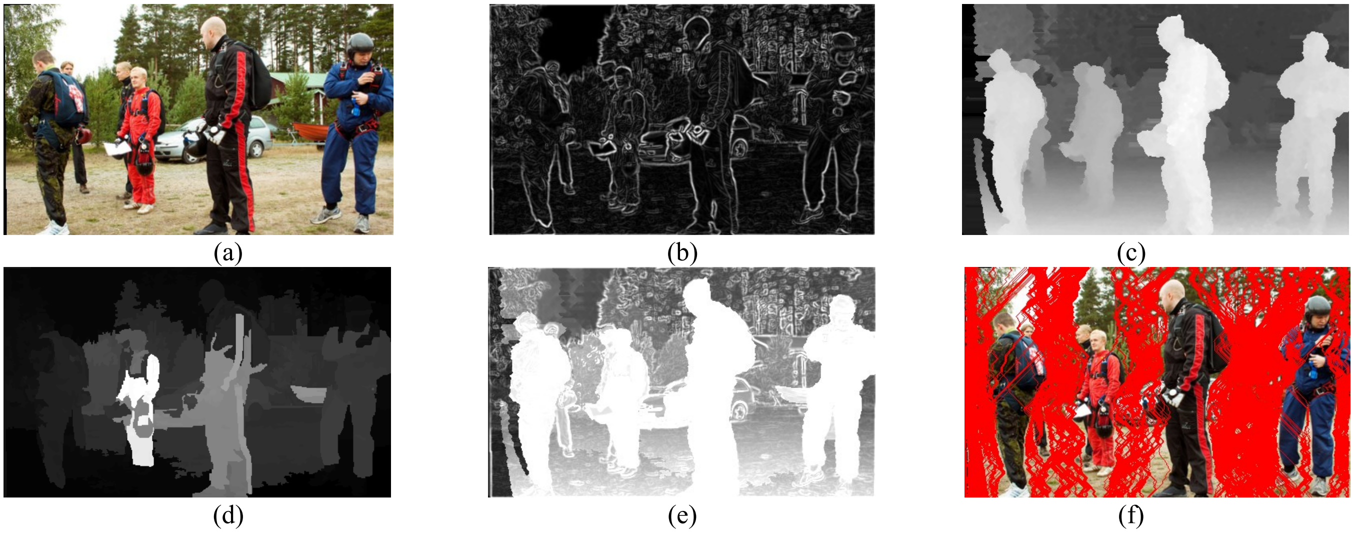

Fig. 2. Illustration of importance maps: (a) Original image, (b) gradient map using Eq. 1, (c) input depth map using [10], (d) saliency map using [11], (e) proposed combined energy map, (f) selected seams for 50 percent width reduction.

Using saliency map and gradient map we obtain a similar relation for $mask_{s,g}$.

$$mask_{s,g} = \begin{cases} 1 & A_{s,g} > T_{s,g}^{A} \text{ and } B_{s,g} < T_{s,g}^{B} \\ 0 & A_{s,g} < T_{s,g}^{A} \text{ or } B_{s,g} > T_{s,g}^{B} \end{cases} \tag{5}$$

We combine the above two masks using logical AND to obtain an importance map $mask_{s,d,g}$ based on saliency, depth, and gradient maps.

$$mask_{s,d,g}(i,j) = mask_{s,d}(i,j) \times mask_{s,g}(i,j) \tag{6}$$

Now we calculate the maximum between the importance map of Eq. 6 and the normalized version of original saliency map, SM, using Eq. 7. In this way we obtain an improved saliency map that emphasizes important pixels of the image.

$$SM_{improved}(i,j) = \begin{cases} 1 & mask_{s,d,g}(i,j) = 1 \\ SM(i,j) & mask_{s,d,g}(i,j) = 0 \end{cases} \tag{7}$$

Finally the improved saliency map, depth map, and gradient map are added in Eq. 8 to obtain the desired energy map. An example of EM is shown in Fig. 2(e).

$$EM(i,j) = SM_{improved}(i,j) + DM(i,j) + GM(i,j) \tag{8}$$

This energy map gives us a good estimate of pixels' importance. Therefore we can select and remove least important seams to reduce image size. Seam selection is done using dynamic programming. In each iteration memorization table entry $M(i,j)$ is updated by (9).

$$\begin{aligned} M(i,j) &= EM(i,j) \\ &+ \min(M(i-1,j-1), M(i-1,j), M(i-1,j+1)) \end{aligned} \tag{9}$$

At the end of the algorithm, the minimum entry in the last row of the table indicates the end of the least-energy seam which is the candidate for removal. The least-energy seam is found by backtracking from minimum entry in last row to the first row of the memorization table M [2]. Fig. 2(f) demonstrates the seams that are selected in our proposed energy map using this algorithm.

3. EXPERIMENTAL RESULTS

We tested our proposed method on different images [10]. As can be seen in Fig. 3(a) in image of *Diana*, the salient object occupies the main part of the scene. The *People* image contains important objects at different depths. But in *Snowman* image the foreground is more important than the background.

We obtain the importance map and energy map for each image and perform seam carving algorithm on them. Saliency map is produced using hierarchical saliency detection algorithm [11]. Various methods exist to obtain depth map [8-10]. Stereo matching methods are used for stereoscopic images. We use SGM stereo matching algorithm [13] to estimate the depth map. Then we apply hole-filling on some parts of the image that the algorithm did not find a match for them [10].

We compare our proposed method with several existing methods. We tested our method for different resize percentages and compared the outcomes with those of other methods. Fig. 3 shows some sample results of our algorithm, original seam carving method [2], stereo seam carving method [10], and depth-aware seam carving method [9]. These results are obtained by reducing the width of the input image to 50% of its original. We see distortions in the results of the original seam carving in Fig. 3(b), where, for example, humans in the image of *People* are drastically distorted and none of the faces is recognizable. We see similar situations in other images. Also, the stereo seam carving method of [10] shows some visual artifacts in important objects. As shown in the third column of Fig. 3,

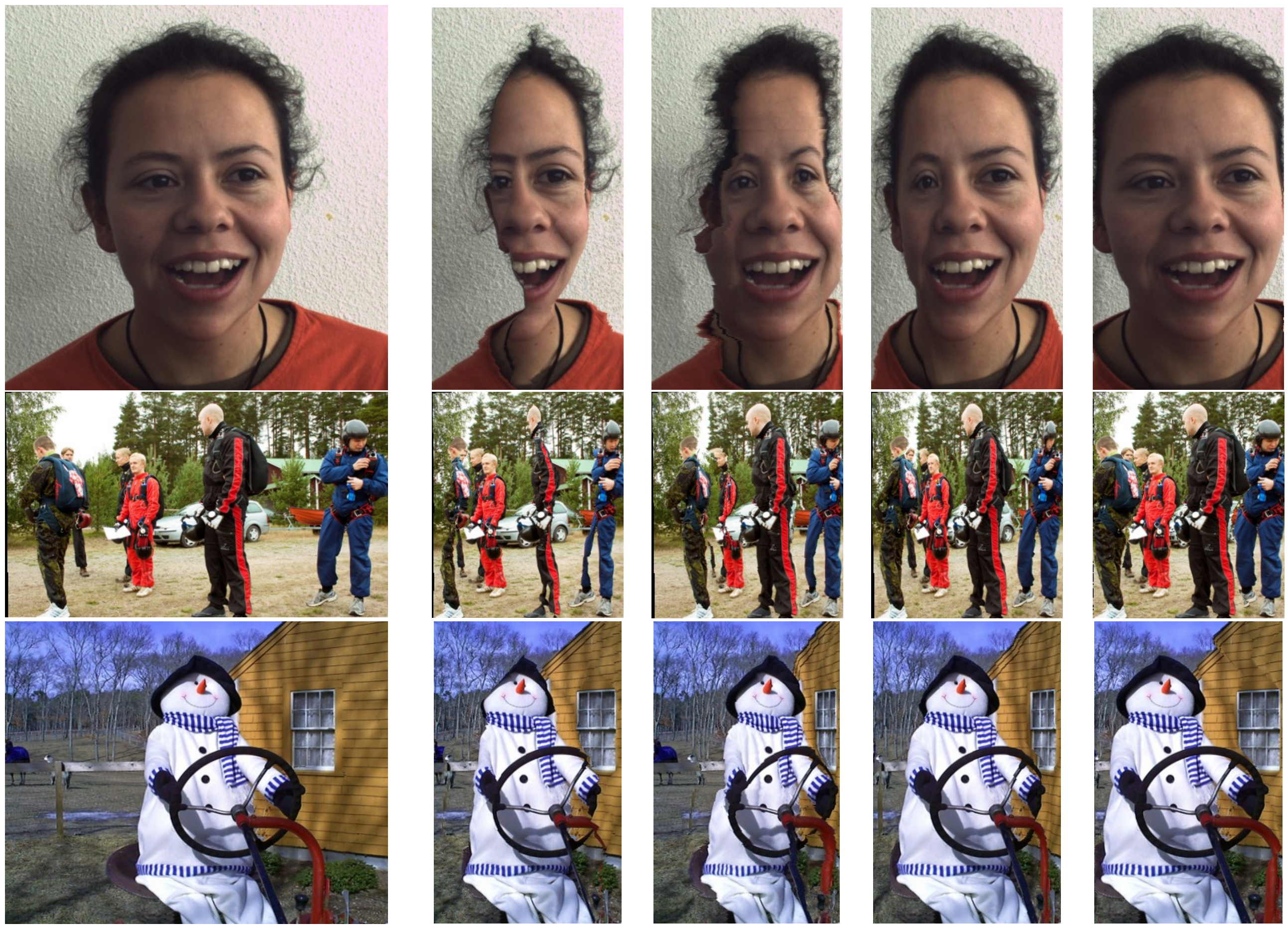

Fig. 3. Comparison of results for 50 percent reduction: (a) Original image, (b) result by original seam carving [2], (c) result by stereo seam carving [10], (d) result by depth-aware seam carving [9], (e) our result.

stereo seam carving [10] produces considerable distortions in images of *Diana* and *Snowman*. The steering wheel and its supporting bar in *Snowman*, which are part of the foreground, are damaged. In *People* the person that is nearest to the camera is well preserved, but others are distorted. The stereo seam carving algorithm has more restrictions in seam selection, because it tries to preserve geometric consistency between left and right images and tries to avoid seam removal from occluded regions. The method in [9] produces better resized images in Fig. 3(d) in comparison with methods of [2] and [10]. But it also causes distortion in important objects when shrinkage increases.

The results of our proposed algorithm are shown in Fig. 3(e). With this method, scene consistency is well preserved at 50% shrinkage and important objects have acceptable visual appearances. In the challenging image of *People*, our method preserves shape and geometry of all persons in different depths. For example two persons on the left and right of the image suffer from less shrinkage in comparison to other methods. Our method is the only one which has preserved the backpack of the front person. Also, the background has acceptable visual quality. Our proposed method preserves the whole face of *Diana* from distortion. Also, the foreground of *Snowman* has better geometrical consistency in comparison to other methods.

4. CONCLUSIONS

In this paper we proposed an energy function that combines the benefits of three different importance maps to avoid distortion in important parts of an image. This energy map assigns high values to important pixels using saliency, depth, and gradient maps in order to improve visual appearance of the resized image. Experimental results for high percentage image shrinkage show that our method produces less distortion in comparison with the state of the art methods. Also salient objects in different depths of image are well preserved.

5. ACKNOWLEDGEMENTS

We would like to thank J. Shen, D. Wang, and X. Li for providing us the code for their depth aware SC algorithm.

6. REFERENCES

[1] L. Chen, X. Xie, X. Fan, W. Ma, H. Zhang, and H. Zhou, "A Visual Attention Model for Adapting Images on Small Displays," *Multimedia Syst*., vol. 9, no. 4, pp. 353-364, 2003.

[2] S. Avidan and A. Shamir, "Seam Carving for Content-aware Image Resizing," *ACM Trans. on Graphics (TOG)*, vol. 26, no. 3, pp. 10.1-10.9, 2007.

[3] M. Rubinstein, A. Shamir, and S. Avidan, "Improved Seam Carving for Video Retargeting," *ACM Trans. on Graphics (TOG)*, vol. 27, no. 3, pp. 1-9, 2008.

[4] M. Rubinstein, A. Shamir, and S. Avidan, "Multi-operator Media Retargeting," *ACM Trans. on Graphics (TOG)*, vol. 28, no. 3, pp. 1-11, 2009.

[5] R. Achanta, and S. Susstrunk, "Saliency Detection for Content-aware Image Resizing," in *IEEE Inernational Conf. on Image Processing (ICIP)*, pp. 1005-1008, 2009.

[6] S. Cho, H. Choi, Y. Matsushita, and S. Lee, "Image Retargeting Using Importance Diffusion," in *IEEE Inernational Conf. on Image Processing (ICIP)*, pp. 977-980, 2009.

[7] D. Domingues, A. Alahi, and P. Vandergheynst, "Stream Carving: an Adaptive Seam Carving Algorithm," in *IEEE Inernational Conf. on Image Processing (ICIP)*, pp. 901-904, 2010.

[8] A. Mansfield, P. Genler, L. Gool, and C. Rother, "Scene Carving: Scene Consistent Image Retargeting," in *Proc. European Conf. Computer Vision (ECCV),* pp. 143-156, 2010.

[9] J. Shen, D. Wang, and X. Li, "Depth-aware Image Seam Carving," *IEEE Trans. on Information Cybernetics*, vol. 43, no. 5, pp. 1453-1461, 2013.

[10] T. Basha, Y. Moses, and S. Avidan, "Stereo Seam Carving a Geometrically Consistent Aproach," *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 35, no. 10, pp. 1816-1823, 2013.

[11] Q. Yan, L. Xu, J. Shi, and J. Jia, "Hierarchical Saliency Detection," *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 1155-1162, 2013.

[12] N. Otsu, "A Threshold Selection Method from Gray-level Histograms," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.

[13] H. Hirschmuller, "Stereo Processing by Semiglobal Matching and Mutual Information," *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 30, no. 2, pp. 328-341, 2008.