Signal Processing: Image Communication 50 (2017) 34–43

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

Image retargeting using depth assisted saliency map

F. Shafieyan[a], N. Karimi[a,*], B. Mirmahboub[b], S. Samavi[a,c,d], S. Shirani[d]

[a] Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran
[b] Department of Pattern Analysis and Computer Vision (PAVIS), Istituto Italiano di Tecnologia (IIT), Genova 16163, Italy
[c] Department of Emergency Medicine, University of Michigan, Ann Arbor 48109 USA
[d] Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada L8S4L8

ARTICLE INFO

Keywords:
Image retargeting
Energy map
Seam carving
Scaling

ABSTRACT

Retargeting algorithms are used to transfer and display images on devices with various sizes and resolutions. All of these algorithms try to preserve the important parts of image against distortions while producing a retargeted image with visual quality comparable with the original one. The main challenge in different algorithms is to find a suitable energy function that properly estimates the importance of each pixel in image. Hence the energy map needs to be improved. In this paper we propose a novel energy function which combines the information from saliency map, depth map and gradient map. We also present an algorithm to adaptively assign proper weights to these three importance maps for each input image. Then we calculate a switching threshold based on energy map that determines when to apply seam carving or scaling. The idea is to use a combination of seam carving and scaling to preserve the structure of important parts of image against distortion when the image size decreases beyond a point. This method reduces shape deformations and visual artifacts in salient regions of images and produces better quality output images. The results of the proposed method show superior visual quality in produced images in comparison to the state-of-the-arts.

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. Naive methods in image resizing, such as homogeneous scaling and cropping, receive the input image and produce an output image, independent of the image contents [1]. In the cropping method, a window with desired target size is selected on the input image and every part of image that resides outside of the window is removed. Some of the deleted parts may include important information. In the scaling method the entire image is deformed uniformly so that new image with desired aspect ratio is produced. In this method some important objects in the scene may be deformed to unwanted shapes. 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.

In spite of many 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 [4–6]. Achanta et al. [4] 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 is determined based on color and intensity globally. Cho et al. [7] proposed a low complexity method where a removed pixel keeps affecting the energy of its previous 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. [8] introduced stream carving algorithm that uses saliency information in addition to gradient to determine the importance of pixels. Their algorithm uses edge detector, face detector and line detector to improve the energy function.

To improve image resizing, some methods combine seam carving with other operators [9,10]. Rubinstein et al. [10] combined seam carving with cropping and scaling. They found an optimum sequence for their combination that produces better results in comparison to each of these operators alone. Decision about optimum switching time

* Corresponding author.
E-mail address: nader.karimi@cc.iut.ac.ir (N. Karimi).

http://dx.doi.org/10.1016/j.image.2016.10.006
Received 3 April 2016; Received in revised form 30 October 2016; Accepted 31 October 2016
Available online 05 November 2016
0923-5965/

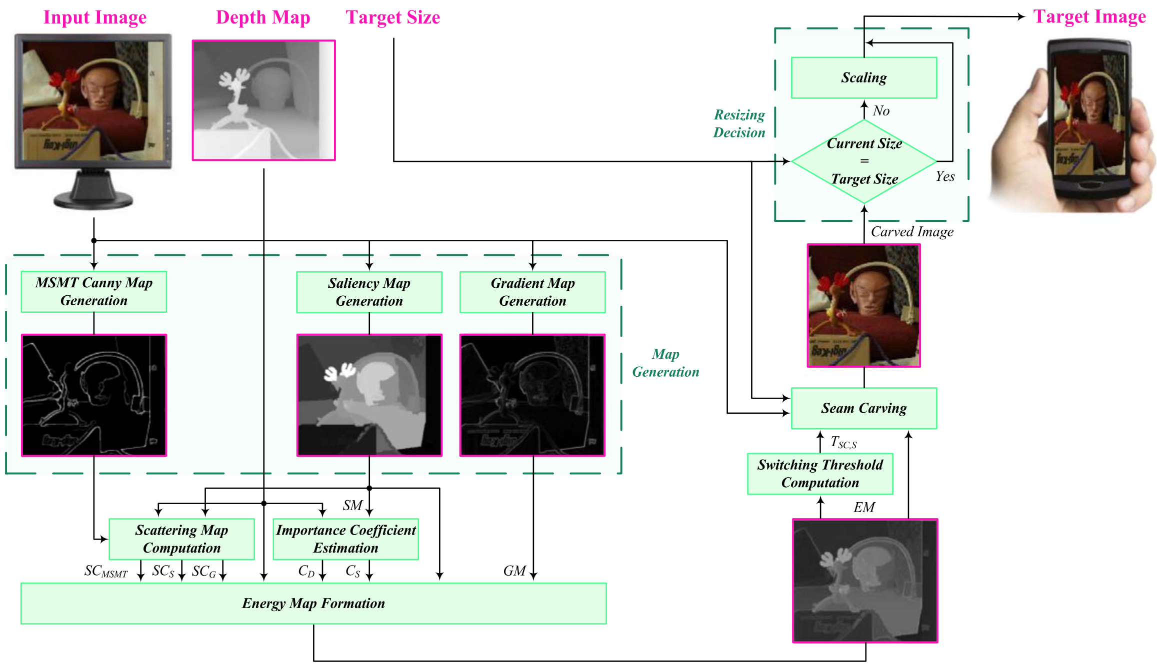

Fig. 1. Framework of the proposed method.

between seam carving and scaling operators are discussed in [11–13]. Dong et al. [13] introduced an image distance function to combine seam carving and scaling. This algorithm removes seams from input image one by one. After each removal, the carved image is scaled to target size and its distance to the original image is calculated. The output image with minimum distance determines the optimum number of seams that can be removed from the original image. This algorithm is computationally intensive and time consuming. Dong et al. [14] added cropping operator to seam carving and scaling to expedite image resizing. Although this algorithm is faster than the previous one; but may crop some important parts of the image and reduce the visual quality of the final image.

Recently some retargeting methods tried to use depth information of the objects in the scene for seam carving purposes. These information that could help better estimation of importance of the image pixels may be obtained by different means [15–17]. Depth map can be determined manually by the user. Mansfield et al. [15] 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 the 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. [16] introduced depth-aware image seam carving that used Kinect depth camera or 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 difference (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. Shafieyan et al. [17] proposed a new combined energy function using gradient, saliency and depth maps. This method introduced an algorithm that highlights the values in saliency map with gradient and depth information to better emphasize on salient objects of the image. Depth information also can be used in stereo image retargeting methods [18–23]. For example, Basha et al. [18] 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 another algorithm Li et al. [19] introduced warping-based stereo image retargeting method that tried to simultaneously achieve both shape and depth preservation in a retargeted image pair.

A proper energy function plays an essential role in successful retargeting. Therefore, in this paper, we propose a novel combinatorial energy function for better estimation of importance of pixels in order to better preserve salient objects of an image. Our idea is to exploit the advantages of saliency map, depth map and gradient map to recognize the important parts of the image. The proposed algorithm determines the role of each of these importance maps in construction of the final energy map in order to improve seam carving. On the other hand, when the size of image decreases too much, some high energy pixels will be consequently eliminated and visual artifacts will appear in the image. Our method considers this problem and presents an algorithm to find a switching threshold. This switching threshold is used to properly decide when to switch from seam carving to scaling when the image size is being reduced. Hence visual appearance of salient objects in the final image will be preserved by suitable tradeoff between visual artifacts that are caused from seam carving and shape deformations that are caused by scaling.

The rest of this paper is organized as follows. Section 2 describes our proposed image retargeting method using depth assisted saliency map in detail. Experimental results are presented and discussed in Sections 3 and 4 concludes the paper.

2. Proposed method

The main goal in image retargeting is to preserve the most important pixels of the image. In order to properly estimate the importance of image pixels, we combine gradient, saliency, and depth information to form a new energy map. Image contents are dissimilar in different images. Hence, the contribution of each of these three importance maps is different for different images. For example, an image needs more information from saliency map if it has an important object in the foreground. While depth map is more effective in preserving salient regions in an image with important objects scattered in different depths. Therefore we need to determine the contribution of each of the gradient, saliency, and depth maps in construction of the final energy map. Furthermore, in order to decrease visual artifacts that are caused by removing a lot of seams when size of image is reduced too much, we propose an algorithm to determine a switching threshold to apply scaling instead of seam carving. This method prevents drastic distortions from happening in an image by distributing the resizing procedure between seam carving and scaling. After switching to scaling, seams are not removed from high-energy regions of image anymore and the visual quality of the final retargeted image is preserved.

The framework of our proposed method is depicted in Fig. 1. Inputs to the algorithm are the color image, its corresponding depth map, and dimensions of the desired output image. The first part of the algorithm is the Map Generation stage that calculates some preliminary required information from the color image. The Scattering Map Computation and Importance Coefficient Estimation stages are done afterward. They provide required information for the Energy Map Formation stage. This last stage combines importance maps with their corresponding weights to produce the final energy function. A switching threshold is obtained based on the produced energy map that decides when to stop the seam carving process. The last section is the Resizing Decision stage that performs the required image scaling to reach the desired target size. In the following sub-sections we explain the details of each part of the algorithm.

2.1. Map generation

The first part of the proposed algorithm generates a gradient map (GM) and a saliency map (SM). A depth map (DM) is supposed to be fed as an input to the algorithm alongside the original image, as is shown in Fig. 1. The depth map separates far objects from the near ones. This map may be obtained by different means such as Kinect sensors [16], stereo matching algorithms for stereo images [24–26], and predetermined depth maps provided by the user [15]. SM separates visually important parts of the image from other parts. This information can be calculated using different algorithms that exploit several characteristics of the image such as color, intensity and orientation [27]. In our method, we use hierarchical saliency detection method as presented in [28]. This algorithm extracts three image layers of different scales from the input image, and then computes saliency cues from each of these layers. Then, a graphical model is used to fuse these cues into one single map, and produces the final saliency map. GM is calculated using information from image edges, based on (1), where $I(i, j)$ is the intensity of a pixel in the input image [2]. That is,

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

One of the most important concerns in image retargeting is prevention of drastic fractures in an object in the scene, which usually appears as a discontinuity in important image edges. Therefore GM plays an important role when there are many objects in the scene. It should be mentioned that higher values of gradient map do not necessarily mean the more importance of this map. Therefore the contribution of GM in construction of the final energy map should be calculated only based on important edges. We use multi-scale multi-threshold canny (MSMT) algorithm [29] for edge extraction. This algorithm applies Canny edge detection with different thresholds in each scale and averages them. Final edges are the maximum of all detected edges in different scales. Then we use thresholding to remove less important edges and prepare the map of only important edges for the next step of the algorithm that calculates scattering of the gradient map. We refere to this map as *MSMT*.

2.2. Scattering map computation

One criterion that we used in our proposed method is scattering of spatial information of scene objects. This criterion helps us to determine the contributions of GM, SM and DM in building the energy map. The idea of scattering is different for each importance map. SM determines important regions more exactly if salient objects are more concentrated in the scene and are less scattered [31]. Therefore, in such cases the energy map can benefit more from SM. Otherwise, if salient objects are scattered in various depths of the image, then saliency map does not contain sufficient information to estimate a proper energy map. In this case, depth information (DM) would help more to estimate a proper energy function. On the other hand, important edges in the image usually belong to object boundaries. Therefore, existence of various objects in the scene leads to more scattering in image edges and GM plays more important role in the estimation of image pixels importance.

In order to compute the scattering inside of an importance map, we use the concept of compactness as proposed in [31]. Each importance map is divided into $M'\times N'$ non-overlapped patches. Scattering of each patch depends on its dissimilarity and distance to other patches in the image. Dissimilarity between two patches $p_{s,t}$ and $p_{i,j}$ is calculated as (2):

$$S(p_{i,j}, p_{s,t})=\|p_{i,j}-p_{s,t}\|_2^2, \tag{2}$$

where $0 < i \le M'$ and $0 < j \le N'$. Also the distance between those two patches is obtained as follows:

$$D(p_{i,j}, p_{s,t})=\max\left(\frac{|i-s|}{M'-1}, \frac{|j-t|}{N'-1}\right). \tag{3}$$

Based on (2) and (3) we estimate the scattering of patch $p_{i,j}$ by:

$$SC(p_{i,j})=\frac{1}{Q(i, j)}\left(\sum_{s=1}^{M'}\sum_{t=1}^{N'}S(p_{i,j}, p_{s,t})\times D(p_{i,j}, p_{s,t})\right), \tag{4}$$

where $Q(i, j)=\sum_{s=1}^{M'}\sum_{t=1}^{N'}s(p_{i,j}, p_{s,t})$. Finally the total scattering of an importance map is calculated as follows:

$$SC_{Map}=\frac{1}{M'\times N'}\left(\sum_{i=1}^{M'}\sum_{j=1}^{N'}SC(p_{i,j})\right). \tag{5}$$

We calculate the scatterings of importance maps of $MSMT$, SM and DM based on (5) and call them SC_{MSMT}, SC_S and SC_D respectively.

2.3. Importance coefficient estimation

In order to better preserve the salient regions of the image, the algorithm of [17] finds an importance mask from SM, GM and DM. We extend their work and propose an algorithm to calculate importance coefficients for SM and DM that determine their contribution in construction of the final energy map as it is shown in Fig. 2.

Salient objects are usually near the camera. Therefore we expect that saliency map and depth map are related together and some important regions are shown in both of them. We propose an algorithm to consider this overlap between SM and DM. We define a binary mask $Mask_{S,D}$ to indicate important regions of the image as mentioned in [17]. On the other hand, SM and DM independently determine some important regions. Such details are represented by binary masks $Mask_S$ and $Mask_D$ which are produced by applying thresholding on SM and DM respectively. Finally in the Map Contribution Estimation stage of Fig. 2, we calculate a coefficient for each of SM and DM based on their contributions to important regions according to:

$$C_S=\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}Mask_S(i, j)\times Mask_{S,D}(i, j)}{\sum_{i=1}^{M}\sum_{j=1}^{N}Mask_{S,D}(i, j)}, \quad C_D=\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}Mask_D(i, j)\times Mask_{S,D}(i, j)}{\sum_{i=1}^{M}\sum_{j=1}^{N}Mask_{S,D}(i, j)}. \tag{6}$$

For an image of size $M\times N$, the coefficients C_S and C_D estimate the validity of salient regions in SM and DM respectively. These coefficients are in interval [0,1]. The closer a coefficient is to 1, the more effective its importance map is in estimation of the final energy map.

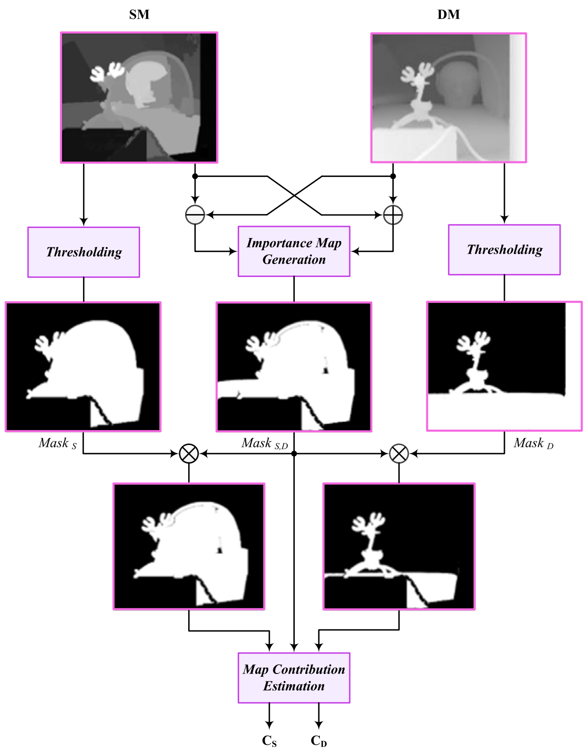

Fig. 2. Framework of the proposed importance coefficient estimation scheme. Symbols of ⊕, ⊖ and ⊗ are pixel-wise addition, subtraction and multiplication operators respectively.

2.4. Energy map formation

In this section we put together the results from previous sections to build a final energy function (EM) by combining SM, DM and GM with proper weights. Saliency map is more important when salient objects are concentrated and less scattered in the scene. Hence, scattering of SM, that is estimated by coefficientSC_S, has inverse relation with its importance in construction of EM. Therefore, in the final energy function, we model the effect of scattering of saliency map as $(1 - SC_S)$. Considering the positive effect of coefficient C_S and the negative effect of coefficient C_D on SM, we introduce coefficient α for SM as below:

$$\alpha=(1-SC_S) \times exp(C_S-C_D). \tag{7}$$

In contrast, depth information plays more important role in construction of EM if its spatial information scattering is high. We calculated SC_D to estimate this scattering of DM. Considering the positive effect of coefficient C_D and the negative effect of coefficient C_S on DM, we introduce coefficient β for DM as follows:

$$\beta=SC_D \times exp(C_D-C_S). \tag{8}$$

In a similar manner, the importance of gradient information is proportional to the scattering its spatial information that is estimated by coefficient SC_{MSMT-T}. On the other hand, when scattering in DM is high, it means that many objects are present in the scene and preservation of their edges is important. Therefore, SC_D has an effect on GM and we introduce coefficient γ for GM as below:

$$\gamma=SC_{MSMT-T} \times exp(C_D). \tag{9}$$

The coefficients C_D and C_S are entered in the energy function in a way that they have increasing effect on their corresponding importance map and decreasing effect on other importance maps. If $C_D=C_S$, then we conclude that salient regions are recognized with the same importance by both DM and SM. Therefore C_D and C_S are ineffective in construction of the energy function and only scattering coefficients are remained. Hence, the proposed energy function is obtained as:

$$EM(i, j)=\frac{1}{N}(\alpha \times SM(i, j)+\beta \times DM(i, j)+\gamma \times GM(i, j)), \tag{10}$$

where $N=\alpha+\beta+\gamma$.

2.5. Switching threshold computation

Image pixels have different importance. Removing low-energy pixels has little effect on the visual quality. But some pixels belong to salient objects and their removal makes considerable distortion in the image. But sometimes many important objects are scattered all over the image and most parts of the image have high-energy. In this case, seam carving method may select and remove some high-energy parts of the image to reach the desired target size, causing considerable distortions.

On the other hand, using only scaling method, for high reduction percentages, will cause serious shape deformations in salient objects. If we detect the low-energy parts of the image and carve them before scaling, we can reduce the distortion that is caused by image scaling. In another words, combination of seam carving and scaling methods in resizing process can alleviate the mentioned two problems of geometrically deformation and distortion in salient objects. For this purpose we propose an algorithm to find a switching threshold between seam carving and scaling methods based on energy distribution of EM. This threshold is different for each image. Our algorithm consists of three steps of seam energy estimation, seam energy grouping, and threshold estimation.

2.5.1. Seam energy estimation

Our goal from determining the proper amount for carving is to remove only the lowest energy seams. We first assume that the entire image is going to be carved and obtain the energy of all the image seams by applying dynamic programming [2] on EM. In each step of the algorithm the memorization table entry $M(i, j)$ is updated as:

$$M(i, j)=EM(i, j)+\min(M(i-1,j-1), M(i-1,j), M(i-1,j+1)). \tag{11}$$

At the end of algorithm the last row of the table contains the energy of all image seams in a vector that we call E_{Seam}. In this step we just find preliminary estimates for priorities of seams to be removed but we actually do not remove any of them. In real seam carving after removing each seam, gradient information and consequently the energy map will change. But these preliminary estimates help us to decide about proper time to switch from seam carving to scaling.

2.5.2. Seam energy grouping

Seams can be categorized based on the energy estimation that was presented in the previous section. For categorization, we sort the seams in ascending order based on their energy and apply thresholding on them. Thresholding is done in two steps as depicted in Fig. 3. In the first step, the threshold $T^{[1]}$ is calculated that divides the seams into two groups of low energy and high energy. In the second step, thresholding is done in each of the two groups to find two thresholds of $T_L^{[2]}$ and $T_H^{[2]}$. These two last thresholds determine the lower and upper bounds of the seams with medium energy. The seams are now categorized as low-energy, medium-energy, and high-energy. This categorization can be summarized as follows:

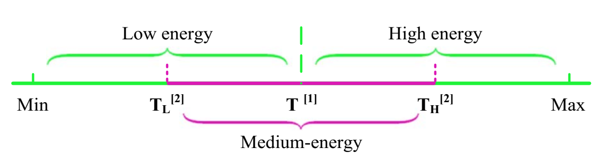

Fig. 3. Seam energy grouping.

Fig. 4. Evaluating the effectiveness of our combined 3-criteria energy map in comparison with each individual criterion for *People* image. (a) Original image, (b) top to down: saliency [28], gradient [2], depth, and proposed combined maps as the energy map, (c) selected seams, and (d) corresponding seam carving results for 20% width reduction, respectively.

$$E_{Seam,i} \in \begin{cases} E_l & E_{Seam,i} < T_L^{[2]} \\ E_m & T_L^{[2]} \le E_{Seam,i} \le T_H^{[2]}, \\ E_h & E_{Seam,i} > T_H^{[2]} \end{cases} \tag{12}$$

where $E_{Seam,i}$ is the energy of i-th seam and $1 \le i \le n$. E_l, E_m and E_h are the intervals of energies for low-energy, medium-energy and high-energy seams respectively. The interval E_l is the best candidate for seam carving; because removing the seams in this group causes least fractures in the image. Scaling is suitable operator to apply to interval E_h; because it prevents serious distortions in salient objects by evenly decreasing the size of image.

2.5.3. Threshold estimation

E_m is the interval of medium-energy seams with k sorted elements. We use this interval for finding a switching threshold. First we apply a Gaussian smoothing filter to reduce small fluctuations in the medium energy band. Then we look for the first element where energy difference is bigger than the average of energy difference in the entire interval. Then we select that element as the proper threshold $T_{sc,s}$. That is given by:

$$J = \left\{ j 0 < j < k \quad , \quad E_{m,(j+1)} - E_{m,j} > \frac{E_{m,k} - E_{m,1}}{k-1} \right\}, \quad T_{sc,s} = E_{m,\min\,(J)} \tag{13}$$

where $E_{m,j}$ is the j-th element in interval E_m and $T_{sc,s}$ is the global switching threshold between [25] seam carving and scaling that we observe a jump in energy level. In fact this threshold indicates the pixels that their carvings have minimum effects on the image visual appearance.

Fig. 5. Comparison of results for *Reindeer* which consists of salient regions both in foreground and background: (a) Original image, (b) proposed energy map, (c) results from our method (without scaling) and (d) results of [16].

2.6. Seam carving

In seam carving method, the selection and removal of the seams are done by applying dynamic programming on the proposed *EM* according to (11). 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. This process is continued until the desired target size or threshold $T_{sc,s}$ is reached.

2.7. Resizing decision

In the last stage of our algorithm, we check the size of the carved image after removing the low-energy seams. If the target size has not been achieved by seam carving, we apply image scaling to reach the desired final size.

3. Experimental results

We implemented our proposed algorithm by MATLAB© R2012a and applied it on images with different characteristics [16,18,24,25]. We present our evaluation results in three sub-sections. First we evaluate the effect of the proposed energy function in improving the seam carving process. Then we compare the proposed algorithm with the state-of-the-art algorithms. In the third sub-section, we discuss failure cases of the proposed method. In our proposed method for construction of the energy map we obtained the image depth information from [16] for *Man*, from [18] for *People*, *Car*, *Diana* and *Snowman* images, and from [24,25] for *Cones*, *Teddy*, *Art*, *Aloe*, and *Reindeer* images. These depth information have 8-bit precision and are obtained from Kinect sensor [16], semi-global matching (SGM) stereo method [26] and structured-lighting approach [24] respectively. Saliency maps are also obtained from [28]. For scattering computation we chose window size of 8×8 pixels. MSMT Canny is extracted using 9 different thresholds, evenly distributed in interval [0.1, 0.9], and 5 different scales 1–5. To select a thresholding scheme to be used in the proposed method, we considered three different approaches of Otsu [30], Kapur [32], and Kittler [33]. To establish an optimum threshold, Otsu's method minimizes the weighted sum of within-class variances of the foreground and background pixels; Kapur's scheme maximizes the sum of entropies and Kittler's method minimizes the total misclassification error of the foreground and background pixels. In our proposed method, we selected to use Otsu [30] as the thresholding approach which has achieved better overall results in our experiments.

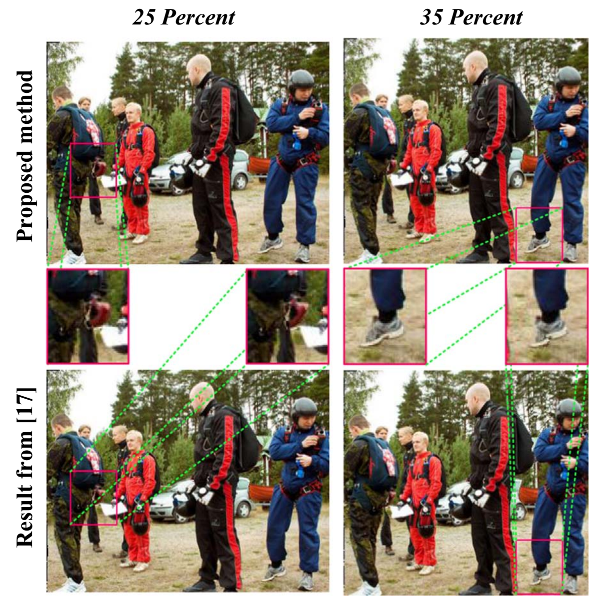

Fig. 6. Comparison of results for 25 and 35% width reduction between proposed method and method of [17]. Two parts of *People* image are zoomed in to show the improvement of our method in comparison with the other method.

3.1. Evaluation of effect of proposed energy function in improving seam carving process

In order to show the effectiveness of the proposed energy map in seam carving process, we evaluated the effectiveness of our combined 3-criteria energy map in comparison with each individual criterion. To do so, we used the gradient map [2], saliency map [28], depth map and our combined map independently as the input energy map. As an example, in Fig. 4 the obtained results for *People* image for 20% image width reduction is shown. In this image (Fig. 4(a)) salient objects are scattered in the entire image and they have different depths. It can be seen when the saliency alone is used as the energy map, salient objects are not well distinguished. Therefore, body shapes of persons, in the left and right of the image, are considerably deformed. Results obtained from gradient as an energy map also contain serious distortions in the structural geometry of the entire image. In such a case, depth information can be more helpful to produce sufficient results. The carved image from depth map shows the effectiveness of depth information in preserving scattered important objects. Here our

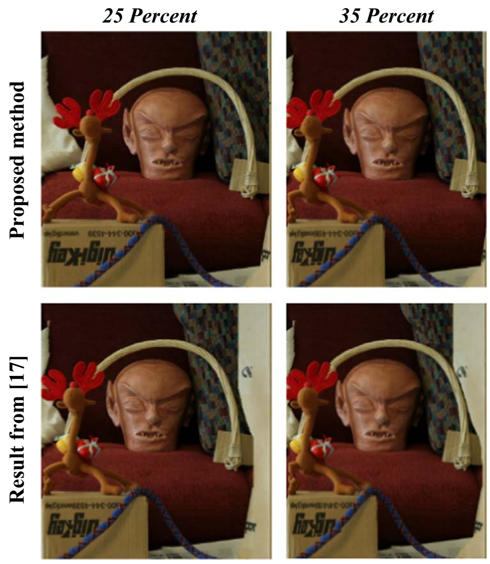

Fig. 7. Comparison of results for 25 and 35% width reductions between the proposed method and method of [17].

proposed energy map realizes that for this image it has to give more weight to depth information while it has to use some of the information of other maps too. Hence, our map produces better global appearance than the use of sole depth map. An example of this superiority is that the distortion that is present between the two persons in the right of the image is lower than the results from the depth map. This is done by removing pixels from the space between two persons in left of the image. This is more apparent by comparing the corresponding selected seams in the third row of Fig. 4.

To compare our method with the other depth based algorithms, we selected method of [16] which is a depth-aware image seam carving scheme. To emphasis the functionality of our energy map in seam carving process, the Resizing Decision stage of the proposed method are bypassed. The obtained results of our method by this consideration are referred to as “proposed method (without scaling)” in the reset of the paper. As an example, Fig. 5 shows the results of our proposed method in comparison to [16] for several different percentages of width reduction of the *Reindeer* image. In comparison to the results of [16] in Fig. 5(d), our proposed energy function performs better in emphasizing important regions of the image. Our method preserves the statue that is located in the image background and produces results with better global appearance. This successful performance is due to simultaneous

Fig. 8. Comparison of results for 50% width reduction. For *People* image the main goal is to preserve the persons in the image as the salient objects. In *Art* and *Cones* images, salient objects are scattered and preserving their geometrical structure is challenging. In *Teddy* image salient objects are located in different depths and occupy the main parts of the image. (a) Original images, (b) original seam carving [2], (c) depth-aware image seam carving, (d) ours method (without scaling), (e) scaling and (f) our method.

Fig. 9. Comparison of results for 50% width reduction. In *Diana* image salient region is the face of the person that is in foreground and occupies the main area of the image. But in *Man* image several objects in foreground in addition to person's face are important. In *Snowman* image, geometric structure of snowman, steering wheel and its connected rod in foreground have more priorities to preserve than visual quality of the background. In *Reindeer* image both foreground and background of the image include salient objects. (a) Original images, (b) original seam carving [2], (c) depth-aware image seam carving [16], (d) ours method (without scaling), (e) scaling and (f) our method.

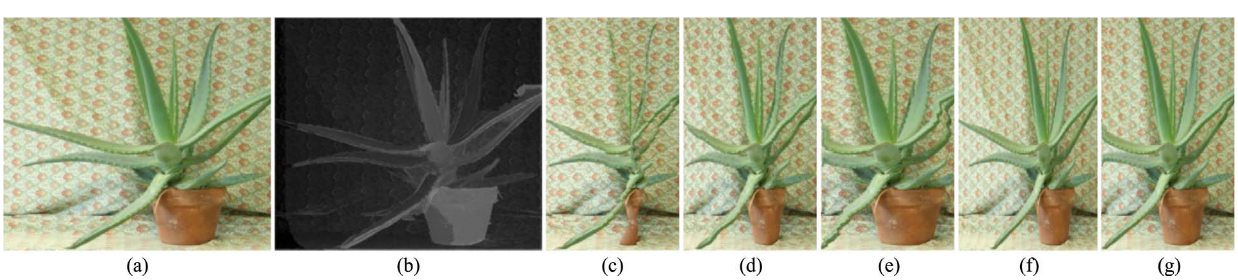

Fig. 10. A failure case. In *Aloe* image salient object is stretched all over the image. (a) Original image, (b) corresponding proposed energy map, (c) original seam carving [2], (d) depth-aware image seam carving [16], (e) our method (without scaling), (f) scaling and (g) our method. Comparison of the results of different methods shows that the pure scaling method produces the best visual quality for this image.

consideration of saliency and depth information to recognize important parts of the image.

On the other hand, proper selection of contribution of each importance map will improve the visual quality of final image. In order to show this issue, two samples are shown in Figs. 6 and 7 that compare the results of proposed method with the results of [17]. In Fig. 6, we see two improvements of our proposed method in comparison with [17]. Our method better preserves the foot of the person on the right and also the leg of the person who is standing deeper in the image are preserved against distortion. Also, our method produces

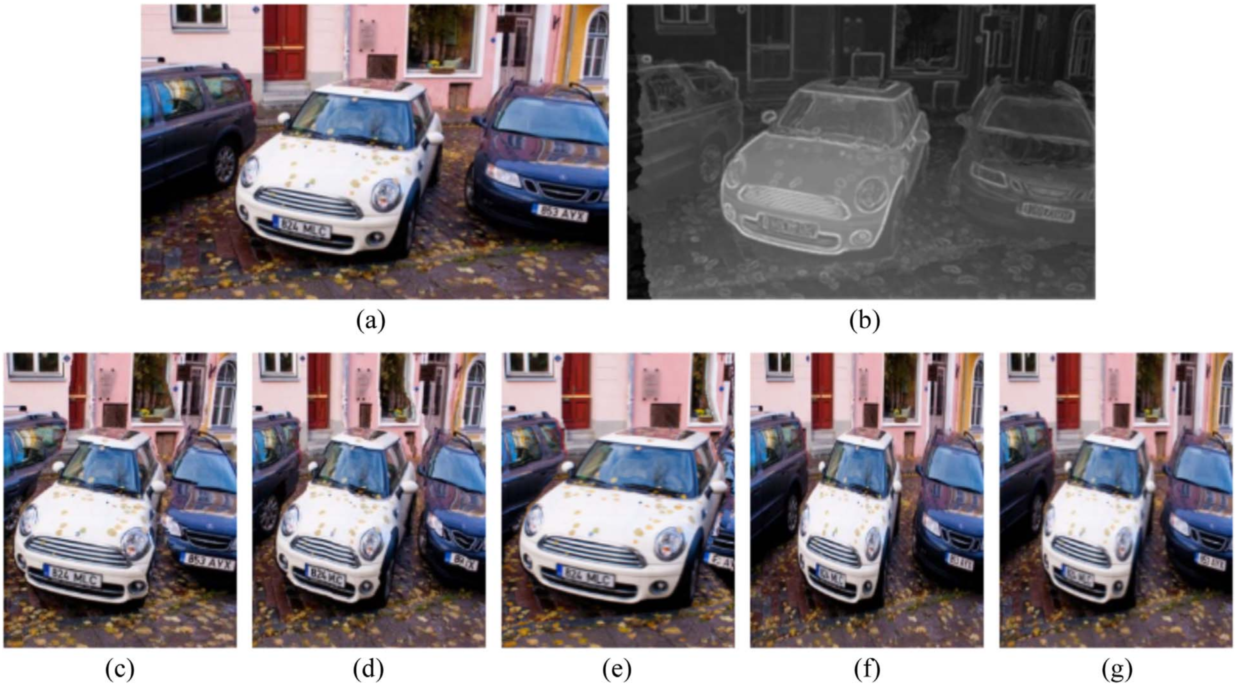

Fig. 11. A failure case. In *Car* image salient objects almost cover the entire image. (a) Original image, (b) corresponding proposed energy map, (c) original seam carving [2], (d) depth-aware image seam carving [16], (e) our method (without scaling), (f) scaling and (g) our method. It seems that pure scaling would provide better global visual appearance for this image.

better visual quality and less distortion in Fig. 7. For example, undesired shrinkage is visible in the right side of the results of [17] that is resolved in our results. In this case, the wall in the right side of the image is not important in proposed energy function and is removed to preserve other salient regions.

3.2. Comparison of proposed seam carving algorithm with state-of-the-art algorithms

In order to compare our proposed method with three methods of scaling, original seam carving [2] and depth-aware image seam carving [16], we chose the images in Figs. 8 and 9 with different characteristics. We evaluate our method in two ways. In the first evaluation we just use the proposed energy map for seam carving method; but in the second experiment we combine scaling with seam carving according to our algorithm to reduce geometrical distortion and improve final visual appearance. As we see in Figs. 8(b) and 9(b) original seam carving method produces serious distortion in structural geometry of salient parts of the image. Therefore the results of this method have lower visual quality in comparison to the results of other methods. Figs. 8(c) and 9(c) show the results of depth-aware image seam carving [16]. The images in Figs. 8(d) and 9(d) are the results of our proposed seam carving method without scaling. For the *People* image our proposed method better preserves the image appearance rather than the method of [16]. This is particularly true for the persons who are standing near the image left and right margins. For the *Art* image our energy function tries to preserve the statue that is the most salient part of image while considering the geometrical structure of the entire image. But some distortions are inevitable in some regions such as objects in the right side because of high percentage of width reduction. Also our energy function gives desired importance to the mask and cones in the *Cones* image and the green puppet in *Teddy*. These are salient objects and our method produces less artifacts in comparison to Fig. 8(c).

Improvement in seam carving method using proposed energy function is more evident for the images in Fig. 9. The proposed method successfully preserves the face of the person in the *Diana* image (Fig. 9(d)) while the result of [16] in Fig. 9(c) shows shrinkage in the left side of the face that is an undesired visual artifact. Also the person's body and camera tripod in *Man* image are better preserved by our method in Fig. 9(d) in comparison with Fig. 9(c). On the other hand, in *Snowman* image although the method of [16] better preserves the cottage roof in the background (Fig. 9(c)) but it produces more considerable deformations in the steering wheel and connected rod rather than our method in Fig. 9(d). In addition the snowman's hat is a bit distorted in Fig. 9(c). In *Reindeer* image the main goal of the proposed energy function is to consider salient parts in both foreground and background simultaneously to produce proper results. The result can be seen in Fig. 9(d) that our method successfully preserves the statue in background against serious artifact in comparison to method [16] in Fig. 9(c).

Some distortions are intrinsic in the results of seam carving method and are inevitable when high percentage of image resizing is requested. In such cases applying image scaling may help reduction of artifacts. But the results that are obtained by the scaling method alone show considerable shape deformations as they are seen in Fig. 9(e). In order to solve this problem our proposed method uses a combination of seam carving and scaling to efficiently avoid serious distortions in the results. These results are shown in Figs. 8(f) and 9(f) where image resizing is first done using seam carving while removing low-energy pixels does not degrade image quality. Then for higher image width reductions, resizing is continued using scaling. Hence, salient regions are preserved and proper final visual appearance is achieved. Some undesired artifacts that are resolved by this method include deformations in people's bodies in the *People* image, distortions in geometrical structure of objects in *Art*, *Cones* and *Reindeer* images and distortions in rod and roof in the *Snowman* image.

3.3. Failures of proposed method

In this section we study two images of *Aloe* (Fig. 10(a)) and *Car* (Fig. 11(a)) that our proposed method does not perform well for them. Considering the proposed energy map in Fig. 10(b), our algorithm gives priority to removing low-energy seams that are close to image margins. This leads to some fractures in side leaves of *Aloe* in the resized image (Fig. 10(e)). The proposed combinatorial algorithm tries to avoid those distortions by applying scaling. But small artifacts are still visible in the leaves near image margins (Fig. 10(g)). Of course, our results are still better than the state-of-the-art method in Fig. 10(d). Our energy function assigns the highest energy to the middle car in Fig. 11(b) because that car is the most important one in both depth and saliency maps. Consequently our proposed seam carving method gives the highest priority to preserve the middle the car and seriously distorts other two cars in the image as it is shown in Fig. 11(e). The proposed combinatorial method considerably avoids this artifact in Fig. 11(g) by applying scaling for a portion of image resizing and produces results that are better than state-of-the-art method in Fig. 11(d). Nevertheless a little distortion is still visible in the left car.

4. Conclusion

In this paper we proposed a novel method for image resizing. First we introduced an energy function to exactly locate salient regions in the image. This energy function is built based on the information of a gradient map, a saliency map and a depth map. Contribution of each importance map depends on scattering of the content of that map and is calculated for each image separately using an algorithm that assigns an importance coefficient to each map. In the second stage we used the estimation of the proposed energy map to find a threshold on seam energy to switch from seam carving method to scaling method. This threshold is triggered when high width reduction percentages are needed and it helps preservation of geometrical structure of image objects. Our proposed image resizing method intelligently employs seam carving and scaling to produce results with desired visual appearances. Comparison of the proposed method with other state-of-the-art methods proved that our method successfully preserves salient regions of the image and provides better global visual quality.

Acknowledgements

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

References

[1] L.Q. Chen, X. Xie, X. Fan, W.Y. Ma, H.-J. Zhang, H.-Q. Zhou, A visual attention model for adapting images on small displays, Multimed. Syst. 9 (4) (2003) 353–364.
[2] S. Avidan, A. Shamir, Seam carving for content-aware image resizing, ACM Trans. Graph. 26 (3) (2007) 10:1–10:9.
[3] M. Rubinstein, A. Shamir, S. Avidan, Improved seam carving for video retargeting, ACM Trans. Graph. 27 (3) (2008) 16:1–16:9.
[4] R. Achanta, S. Süsstrunk, Saliency detection for content-aware image resizing, in: Proceedings of the IEEE International Conference on Image Proceeding (ICIP), 2009, pp. 1005–1008.
[5] L. Mai, Y. Niu, F. Liu, Saliency aggregation: a data-driven approach, in: Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR), 2013, pp. 1131–1138.
[6] H. Jiang, J. Wang, Z. Yuan, Y. Wu, N. Zheng, S. Li, Salient object detection: a discriminative regional feature integration approach, in: Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR), 2013, pp. 2083–2090.
[7] S. Cho, H. Choi, Y. Matsushita, S. Lee, Image retargeting using importance diffusion, in: Proceedings of the IEEE International Conference on Image Proceeing (ICIP), 2009, pp. 977–980.
[8] D. Domingues, A. Alahi, P. Vandergheynst, Stream carving: an adaptive seam carving algorithm, in: Proceedings of the IEEE International Conference on Image Proceeing (ICIP), 2010, pp. 901–904.
[9] S. Luo, J. Zhang, Q. Zhang, X. Yuan, Multi-operator image retargeting with automatic integration of direct and indirect seam carving, Image Vis. Comput. 30 (9) (2012) 655–667.
[10] M. Rubinstein, A. Shamir, S. Avidan, Multi-operator media retargeting, ACM Trans. Graph. 28 (3) (2009) 23:1–23:11.
[11] Y. Wang, C. Tai, O. Sorkine, T. Lee, Optimized scale-and-stretch for image resizing, ACM Trans. Graph. 27 (5) (2008) 8:1–8:8.
[12] D.S. Hwang, S.Y. Chien, Content-aware image resizing using perceptual seam carving with human attention model, in: Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), 2008, pp. 1029–1032.
[13] W. Dong, N. Zhou, J.-C. Paul, X. Zhang, Optimized image resizing using seam carving and scaling, ACM Trans. Graph. 28 (5) (2009) 125:1–125:10.
[14] W. Dong, G. Bao, X. Zhang, J. Paul, Fast multi-operator image resizing and evaluation, J. Comput. Sci. Technol. 27 (1) (2012) 121–134.
[15] A. Mansfield, P. Genler, L. Gool, C. Rother, Scene carving: scene consistent image retargeting, in: Proceedinsg of the European Conference on Computer Vision (ECCV), 2010, pp. 143–156.
[16] J. Shen, D. Wang, X. Li, Depth-aware image Seam carving, IEEE Trans. Cybern. 43 (5) (2013) 1453–1461.
[17] F. Shafieyan, N. Karimi, B. Mirmahboub, S. Samavi, S. Shirani, Image seam caving using depth assisted saliency map, in: Proceedings of the IEEE International Conference on Image Proceeing (ICIP), 2014, pp. 1155–1159.
[18] T.D. Basha, Y. Moses, S. Avidan, Stereo seam carving a geometrically consistent approach, IEEE Trans. Pattern Anal. Mach. Intell. 35 (10) (2013) 2513–2525.
[19] B. Li, L.-Y. Duam, C.-W. Lin, T. Huang, W. Gao, Depth-preserving warping for stereo image retargeting, IEEE Trans. Image Process. 24 (9) (2015) 2811–2826.
[20] K. Utsugi, T. Shibahara, T. Koike, K. Takahashi, T. Naemura, Seam carving for stereo images, in: Proceedings of the 3DTV-Conf., 2010, pp. 1–4.
[21] C.H. Chang, C.K. Liang, Y.Y. Chuang, Content-aware display adaptation and interactive editing for stereoscopic images, IEEE Trans. Multimed. 13 (4) (2011) 589–601.
[22] K.Y. Lee, C.D. Chung, Y.Y. Chuang, Scene warping: layer-based stereoscopic image resizing, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012, pp. 49–56.
[23] J.W. Yoo, S. Yea, I.K. Park, Content-driven retargeting of stereoscopic images, IEEE Signal Process. Lett. 20 (5) (2013) 519–522.
[24] D. Scharstein, R. Szeliski, High-accuracy stereo depth maps using structured light, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2003, pp. 195–202.
[25] H. Hirschmüller, D. Scharstein, Evaluation of cost function for stereo matching, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2007, pp. 1–8.
[26] H. Hirschmüller, Stereo processing by semiglobal matching and mutual information, IEEE Trans. Pattern Anal. Mach. Intell. 30 (2) (2008) 328–341.
[27] H. Jiang, J. Wang, Z. Yuan, T. Liu, N. Zheng, S. Li, Automatic salient object segmentation based on context and shape prior, in: Proceedings of the British Machine Vision Conference (BMVC), Vol. 6, Issue 7, 2011, pp. 110:1–110:12.
[28] Q. Yan, L. Xu, J. Shi, J. Jia, Hierarchical saliency detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 1155–1162.
[29] A. Hernández-Vela, C. Gatta, S. Escalera, L. Igual, V. Martín-Yuste, M. Sabaté, P. Radeva, Accurate coronary centerline extraction, caliber estimation, and catheter detection in angiographies, IEEE Trans. Inf. Technol. Biomed. 16 (6) (2012) 1332–1340.
[30] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man Cybern. 9 (1979) 62–66.
[31] C. Zhang, W. Wang, Object-level saliency detection based on spatial compactness assumption, in: Proceedings of the IEEE International Conference on Image Proceeing (ICIP), 2013, pp. 2475–2479.
[32] J.N. Kapur, P.K. Sahoo, A.K. Wong, A new method for gray-level picture thresholding using the entropy of the histogram, Comput. Vis. Graph. Image Process. 29 (3) (1985) 273–285.
[33] J. Kittler, J. Illingworth, Minimum error thresholding, Pattern Recognit. 19 (1) (1986) 41–47.