Image and Vision Computing 30 (2012) 655–667

Contents lists available at SciVerse ScienceDirect

Image and Vision Computing

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

Multi-operator image retargeting with automatic integration of direct and indirect seam carving ☆

Siqiang Luo [a], Junping Zhang [a,*], Qian Zhang [a], Xiaoru Yuan [b]

[a] Shanghai Key Lab of Intelligent Information Processing, School of Computer Science, Fudan University, 200433, Shanghai, China
[b] Key Laboratory of Machine Perception (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, 100871, Beijing, China

ARTICLE INFO

Article history:
Received 27 June 2011
Received in revised form 16 February 2012
Accepted 17 June 2012

Keywords:
Image resizing
Multi-operator
Seam carving
Cropping method
Image similarity

ABSTRACT

Multi-operator image resizing can preserve important objects and structure in an image by combining multiple image resizing operators. However, traditional multi-operator methods do not take both horizontal and vertical content-aware resizing potential into consideration, which essentially leads to squeeze/stretch effect in the resultant images. In this paper, we propose a new multi-operator scheme that addresses aforementioned issue by integrating direct and indirect seam carving. Compared with previous methods, the proposed scheme remarkably reduces the cost of deciding when to change operators, by employing a newly defined image artifact measure. Furthermore, we propose a novel seam carving enhancement, named ACcumulated Energy Seam Carving (*ACESC*), as a basic operator to improve global structure preservation. By combining horizontal and vertical seam carving, our scheme preserves the shapes of important objects well. We present typical results to demonstrate the effectiveness of our method. User study shows that our method has high user preference.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

In recent years, content-aware image resizing (*a. k. a.*, retargeting) techniques, which can preserve visually important contents and maintain good perceptual invariance in an image when the size and aspect ratio are changed, have evoked a great deal of interests. Such techniques are especially meaningful for transforming images or video clips between multimedia devices with different resolutions.

The techniques can be coarsely divided into five major categories [7]: cropping methods, warping methods, seam carving methods, patch based methods and multi-operator retargeting methods. In cropping methods [8–11], an optimal sub-window of the target size, which contains visually important regions, is searched from the input image. However, a disadvantage is that such methods may discard a large part of essential regions when the important objects are near the image periphery. Warping methods resize images non-homogeneously [3,12–16]. By fixing a mesh on the image, they warp the designed mesh non-uniformly to reach the desired size based on a global optimization function. Patch based methods [17] achieve resizing by optimizing a patch-based similarity measure between the input and target image. One limitation of this method is its high computational cost, as pointed out in [6]. Seam carving (*SC*) methods [4,18] perform image retargeting by iteratively adding or removing the most unimportant curves (*i.e.*, seams) from the image. For images with large homogeneous region going across from its left to right (*e.g.*, the sky region in Fig. 1 going across the image from left to right), a better seam carving scheme, which is called indirect seam carving, first resizes the height by inserting/deleting the horizontal seams in order to achieve the target aspect ratio, and then scales the intermediate image to the target size [14,19]. The advantage of indirect seam carving is that the resizing is mostly done in the homogeneous region, resulting in less visual artifacts. Seam carving methods are especially suitable for images with sparsely distributed objects. It is prone to produce undesirable artifacts when resizing images with many prominent objects in them.

Multi-operator retargeting techniques [1,19] adaptively utilize multiple operators including scaling, seam carving and cropping in turn to achieve content-aware image resizing. Those techniques employ an objective similarity measure to select the best combination from all possible combinations of different operators. A recent survey [6] indicates that compared with most state-of-the-art algorithms, multi-operator frameworks can produce retargeted images highly preferred by people. However, a major quality limitation is that when the resizing task is mono-directional, such techniques neglect the warping potential in the other dimension and are prone to distort the object shape (see Fig. 1 MultiOp, the people are "narrowed"). Another problem is that the computational cost of evaluating the image similarity measures is high and exponentially depends on the number of operators used [20].

Considering the aforementioned pros and cons, we propose a simple yet effective multi-operator resizing scheme to exploit both horizontal and vertical content-aware resizing potential, which has not

☆ This paper has been recommended for acceptance by Michael Goesele.

* Corresponding author. Tel.: +86 21 55664503.
E-mail addresses: sqluo@fudan.edu.cn (S. Luo), jpzhang@fudan.edu.cn (J. Zhang), 10210240044@fudan.edu.cn (Q. Zhang), xiaoru.yuan@pku.edu.cn (X. Yuan).

0262-8856/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2012.06.008

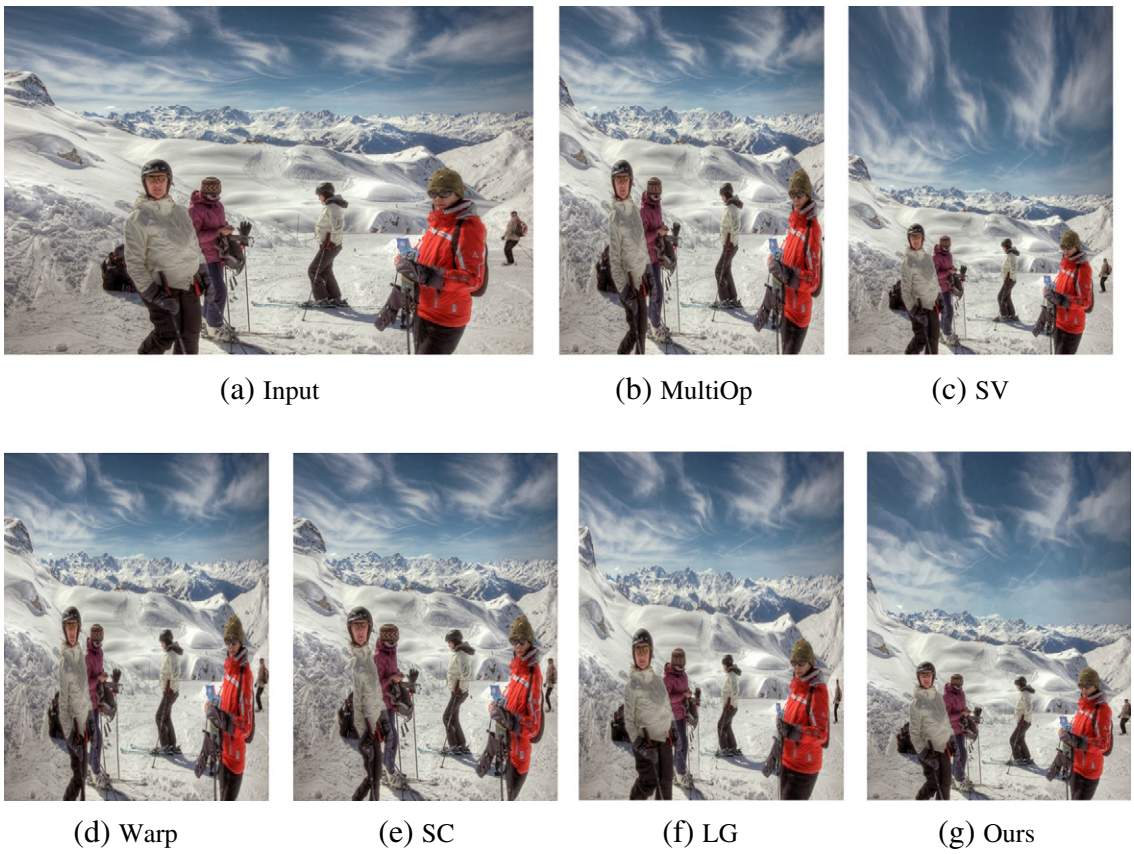

Fig. 1. Examples of retargeting the results of several state-of-the-art methods. (a) Input image; (b) multi-operator [1]; (c) streaming video retargeting [2]; (d) nonhomogeneous warping [3]; (e) seam carving [4]; (f) energy-based deformation [5]; (g) ours. The method notations follow RetargetMe dataset [6]. By exploiting the advantages of integration of direct and indirect *SC* with improved *SC* energy, our method generates better results in comparison with most other methods in terms of keeping the object aspect ratio.

been focused in previous multi-operator techniques [1,19]. The scheme automatically and effectively combines four simple resizing operators. In particular, direct and indirect seam carvings are first performed, and then a similarity transformation is employed to scale the image to a suitable size, followed by cropping the image to the target size. Although we believe that taking the global warping as a basic operator will enhance the ability to distribute distortions in multiple directions, we choose seam carving for its simplicity. Furthermore, we propose a seam artifacts measure to reduce the computational cost of selecting effective changing points among performing different operators. Moreover, we design a novel seam energy scheme for seam carving, named ACcumulated Energy Seam Carving (*ACESC*), to relieve the seam distortion significantly and eliminate the expansion ratio limit of seam carving for enlarging. Compared with several state-of-the-art image retargeting methods, experiments in a benchmark image retargeting dataset indicate that the proposed multi-operator scheme has comparable performance and user preference in preserving the global structure and the shapes of important objects. Its computational cost is remarkably lower than that of the previous multi-operator methods.

The remainder of the paper is organized as follows. In Section 2, we survey related works. In Section 3, we introduce some preliminary knowledge on image retargeting. Section 4 details the proposed multi-operator scheme. Experiment results are presented and discussed in Section 5. We conclude the paper in Section 6.

2. Related work

In this section, we give a brief survey on the content-aware image and video retargeting algorithms.

Among cropping methods, [21–24] constructed an optimal sub-window to preserve the prominent regions of the original image, which usually involves attention model extraction and object detection. With four psychological attributes (region of interest, attention value, minimal perceptible size, and minimal perceptible time), Liu et al. [8] employed an image attention model to extract main objects within an image. Santella et al. [10] used interactive eye tracking technique to locate important image contents to obtain optimal cropping window. Ciocca et al. [11] utilized different cropping methods to deal with different image types that are classified through a well-trained classifier.

The continuous methods [3,12–16,25] are to construct and optimize a warping function to retain the global structure of the images. Liu and Gleicher [5] employed a fisheye warping strategy to retain both details at regions of interest and context information in the resized image. With Laplacian editing techniques, Gal et al. [12] warped an image into various shapes to preserve the user-specified features. Wolf et al. [3] proposed an image and video retargeting method by utilizing local saliency, face detection [26] and motion detection (for video retargeting) to preserve the visually important pixels. Zhang et al. [13] incorporated a per-pixel cumulative shrinkability map with the random walk model to accelerate the computation and save storage space. Wang et al. [14] employed a significance map [27], which characterizes the visual attractiveness of each pixel, to guide the degree of image deformation. They constructed a warping mesh to distort homogeneous regions while preserving visually prominent regions. Guo et al. [16] combined the triangular mesh and important map as well as some structural constraints to preserve the shapes of important objects and minimize visual distortion. Karni et al. [25] presented a general approach to energy-based image retargeting and 2D shape deformation.

The patch based methods [17] achieve image retargeting based on image patches. Simakov et al. [17] defined a bi-directional similarity,

which considers both completeness and coherence of the resultant image, to measure the similarity between the input and target image. With the measure, they performed resizing step by step, and at each step they iteratively updated the intermediate image in order to minimize the similarity error.

Pritch et al. [28] employed shift-map to record the relative shift of every pixel in the output. They used a global cost, which is considerate of image boundary mapping, saliency information and structural smoothness, to guide the pixel-shifting from the input image to the corresponding resized image.

With the aid of energy functions, seam carving methods [4,18] protect the prominent regions in an image and shrink the image by greedily carving out seams. To enlarge an image, the methods first carve out seams and then duplicate the carved out seams to the original image. Backward energy seam carving [18] performs well by "moving around" important objects when the image has a large area of unimportant inter-object regions. Forward energy seam carving [4] prefers deleting a seam that produces the least contrast in the new image in each iteration. A major disadvantage is that the above seam energy methods neglect the cumulative effect of the deleted seams, resulting in unsatisfactory artifacts. Furthermore, current methods [18] cannot automatically deal with large ratio enlarging. An alternative way to perform seam carving is to use indirect seam carving [14,19] by utilizing the homogeneous regions in the direction orthogonal to the resizing direction to protect the important content.

The multi-operator methods [1,19] perform retargeting using the combination of multiple resizing operators. Rubinstein et al. [1] utilized multi-scale Bi-Directional Warping (*BDW*), which is an image similarity measure, to achieve an effective combination of cropping, seam carving and scaling. They designed a "resizing space" and used the *BDW* to select the best path in the "resizing space". The final result is generated according to the best path. Dong et al. [19] suggested performing seam carving followed by scaling the image to the target size. By fusing the image Euclidean distance and the dominant color descriptors into a bidirectional similarity function, they found an optimal ratio of performing seam carving and scaling.

Recently, several researchers pay attention to video retargeting [1–4,20,29,30]. Unlike still image retargeting, video retargeting requires both spatial and temporal coherence of the resized video. Most video retargeting work added temporal constraints into per-frame image-based techniques. For example, [3] used motion detection to extract the temporal constraints for the visually important pixels preservation in each frame. By extending 1D seam to 2D seam (surface carve), Rubinstein et al. [4] treated video data as graphs and utilized graph cut to perform 2D seam carving. Grundmann et al. [29] relaxed the continuous seam constraints by allowing seams to vary by several pixels between adjacent rows/columns. They designed gradient-variation based measures to estimate spatial and temporal coherence as well as spatio-temporal saliency to guide the retargeting. Nevertheless, it is worth noting that improving the quality of image retargeting is still important as it is the basis of improving the quality of video retargeting.

3. Preliminary

This section briefly summarizes the concepts and some details of original seam carving [18], and an image similarity measure, Bi-Directional Warping (*BDW*) [1], which are referred in the development of the proposed framework in the next section.

3.1. Original seam carving for shrinking and enlarging

Given an $m\times n$ ($height\times width$) image I, a vertical seam is a connected pixel path from the top of the image going downward to the bottom, which is formally defined as follows [18]:

$$\left\{I_{i,x(i)}\middle|1\le i\le m\right\},\ \text{s.t.}\forall i,\ |x(i)-x(i-1)|\le 1$$

where x is a mapping from $[1,\cdots,m]$ to $[1,\cdots,n]$, $I_{i,x(i)}$ denotes a pixel located in the $x(i)$-th column of the i-th row of image I. Similarly, a horizontal seam is defined as:

$$\left\{I_{y(j),j}\middle|1\le j\le n\right\},\ \text{s.t.}\forall j,\ |y(j)-y(j-1)|\le 1$$

where y is a mapping from $[1,\cdots,n]$ to $[1,\cdots,m]$.

Original seam carving assumes the importance (or energy) of each pixel is measured by pixel gradient as follows [18]:

$$e\left(I_{i,j}\right)=\left|\mathbb{I}_{i,j+1}-\mathbb{I}_{i,j}\right|+\left|\mathbb{I}_{i+1,j}-\mathbb{I}_{i,j}\right| \tag{1}$$

where $I_{i,j}$ is the intensity value of pixel $I_{i,j}$. Also, it assumes that the energy of a seam is the sum of each pixel's energy in the seam. Then, original seam carving searches the seam with lowest energy (min-seam) using dynamic programming equation [18] (taking vertical seam as an example):

$$M(i,j)=e\left(I_{i,j}\right)+min\begin{cases}M(i-1,j-1)\\ M(i-1,j)\\ M(i-1,j+1)\end{cases}\ \forall 1\le j\le n,2\le i\le m \tag{2}$$

where $e(I_{i,j})$ is the gradient value (energy defined by Eq. (1)) of the pixel $I_{i,j}$ in the i-th row and the j-th column. Meanwhile, $M(i,j)$ records the minimum pixel-energy-sum of an 8-connected path, which starts from any top row pixel and stops at $I_{i,j}$, with exactly one pixel chosen in each row. With Eq. (2), the vertical seam with the lowest pixel-energy-sum (min-seam) is searched (terminated when $i=n$). By iteratively deleting vertical min-seams, we can reduce the width. Note that height resizing follows the same procedure as width resizing.

If the goal is to enlarge an image, there are two strategies in the original seam carving techniques. One is to iteratively search the min-seam and duplicate it immediately. However, such searched min-seams will be closely located, generating visual artifacts if to duplicate them for enlarging. An alternative way is to separate the "seam-search" step and "duplicate" step. It first searches a group of seams as reduction does, and then duplicates all these seams simultaneously [18]. This strategy is limited when to expand the image significantly, as in this case the selected seams to be duplicated may not be enough (*e.g.*, when the enlarged width is larger than the original width).

3.2. BDW for multi-operator retargeting

The core of multi-operator retargeting for content-aware image resizing is to determine the best point of transferring from different retargeting operators. A commonly used technique in the multi-operator schemes is to exploit image similarity measure, such as *BDW* (Bi-Directional Warping) [1]. From the results produced by the possible combinations of operators, *BDW* is utilized to select the best result that is the most similar to the original image. *BDW* measures the similarity between two images with the same height based on Asymmetric-DTW (*A-DTW*) [1], which is a variant of Dynamic Time Warping (*DTW*) [31]. *A-DTW* (s,t) is an asymmetric similarity measure from $1\times|s|$ image's row s to $1\times|t|$ image's row t, where $|\cdot|$ denotes the number of pixels. It finds a mapping g from s to t satisfying that: (1) $g(a)\le g(b)$ if $a\le b$; (2) $\sum_{1\le i\le|s|}d(i,g(i))$ is minimized, where $d(p,q)$ denotes the square difference between the intensity values of pixel p and q. *A-DTW*(s,t) returns $max_{1\le i\le|s|}$ $d(i,g(i))$ as the similarity degree between s and t. We present how to calculate *A-DTW* in Algorithm 1, which is modified from the work [1]. Based on *A-DTW*, a symmetric image similarity measure *BDW* is defined as follows:

$$BDW(S,T)=\max_i\{A-DTW(S_i,T_i)\}+\max_i\{A-DTW(T_i,S_i)\} \tag{3}$$

where S_i, T_i are the i-th row of image S and T, respectively. Note that the computation of *A-DTW* and *BDW* can be performed in both pixel level and image patch level.

Algorithm 1 $maxd = A-DTW_{max}(s[1..|s|], t[1..|t|])$

```
   /* calculating pixel mapping */
 1: Allocate  M[|s|+1][|t|+1], P[|s|+1][|t|+1]
 2:   M[0,0]:=0
 3: for i:=1 to |s| do
 4:   M[i,0]:=∞
 5: for j:=1 to |t| do
 6:   M[0,j]:=0
 7: for i:=1 to |s| do
 8:   for j:=1 to |t| do
 9:     M[i,j]=min{M[i,j-1], M[i-1,j]+d(s[i],t[j])}
10:     if M[i,j]=M[i,j-1]
11:       P[i,j]:=0
12:     else
13:       P[i,j]:=1

   /* backtrack for maximal difference between the match pixel pairs */
14: i:=|s|+1, j:=|t|+1, maxd:=0
15: while(i>0)
16:   temp:=d(s[i],t[j])
17:   if P[i,j]=1
18:     if maxd<temp
19:       maxd:=temp
20:     i:=i-1
21:   else
22:     j:=j-1
23: return maxd
```

4. Our multi-operator framework

The frequently used notations in later discussions are shown in Table 1. We will re-clarify the notations when they are used.

In this section, we first give an overview of the framework in Section 1. Then we discuss the automatic procedure and efficiency issue in Section 2. Finally, we propose an improved seam carving in Section 3.

4.1. Framework overview

In our multi-operator framework, the main concern is how to integrate direct and indirect seam carvings. To simplify, we mainly discuss the typical problem in content-aware image resizing, *i.e.*, shrinking the width. To retarget an image with size $m \times n$ to a smaller size $m \times n'$ ($n' < n$) (aspect ratio changes from $\frac{n}{m}$ to $\frac{n'}{m}$), our proposed framework consists of the following four steps:

- Step 1 (Direct Seam Carving): shrink the width by u ($u \le n - n'$).
- Step 2 (Indirect Seam Carving): enlarge the height by u' (u' should be small enough to let the Step 2 result have a larger aspect ratio than the target aspect ratio $\frac{n'}{m}$).
- Step 3: scale down the image using similarity transformation until the height is equal to m.
- Step 4: crop the image to its target size.

To better understand the proposed multi-operator retargeting framework, we illustrate an example of reducing the width of an image in Fig. 2. Our framework first reduces the width by performing direct seam carving (Step 1) until it begins to generate unsatisfactory artifacts. The generation of such artifacts will be automatically detected and be regarded as a critical point, indicating the switching to indirect seam carving (Step 2), which inserts horizontal seams into the image. When indirect seam carving also begins to cause visual artifacts, the intermediate image will be rescaled to its original height (Step 3), during which the aspect ratio does not change. Finally, a saliency based cropping (Step 4) method crops the image to the desired size if necessary. The final step is suitable as a bit loss of the image content is usually more acceptable than distortions caused by seam carving [6]. To the best of our knowledge, the proposed framework is the first attempt to automatically integrate the direct seam carving and indirect seam carving, which extends the seam carving potential from one dimension to two dimensions.

4.2. Selection of u and u′

Similar to the previous multi-operator framework [1], we can employ the similarity measure *BDW* to select suitable parameter u in the proposed framework (in Section 1). In particular, we can examine every possible ratio between performing seam carving and scaling to reduce a width of $n - n'$. For each retargeted image we compute the similarity (using *BDW*) between it and the original image, and select the most similar one. u is set to be the number of seams deleted corresponding to the selected image. We can formalize this parameter tuning method as:

$$u = \underset{0 \le u_0 \le n-n'}{arg\,min} \left\{ BDW\left(I, \mathcal{F}_{u_0, n-n'-u_0}(I)\right) \right\} \tag{4}$$

where n and n' are the original width and the desired width, respectively. $\mathcal{F}_{u_0, n-n'-u_0}(I)$ is the image by first removing u_0 vertical seams from I and then scaling it to the width n'. Eq. (4) computes the best change point u such that to carve out u vertical seams from I and then scale it to the target width n', the resultant image will be the most similar to the original image I.

Though effective, the above selection is inefficient since it will take $n - n'$ times of computing *BDW* measure to select u. To address this issue, we suggest that during the iterative seam carving step, only a more obvious seam distortion compared with that of the previous step will trigger the calculation of *BDW*. The rationale is that the occurrence of more obvious artifacts caused by seam carving is a reliable signal that seam carving may be over-performed. Therefore, we define a quantity measure, *ATF*, to estimate the seam artifacts. We note that seam artifacts come from misalignment between the original image and the seam carved image. Therefore, we define $ATF(t)$ as the maximal patch misalignment between images I and I_t (the image after removing t vertical seams from I). Specifically, for each $d \times d$ (d is odd) patch in I_t whose central pixel is p, there must be a corresponding patch in I of the same size whose central pixel is also p. For each patch pair we calculate the sum of square difference between them as the patch misalignment. *ATF* is the maximal patch misalignment. One may concern that if the image has repeating textures, the artifacts measure may drop to 0. Since we actually focus on the *increase* of *ATF* rather than the absolute value of it, however, such a phenomenon will not be a problem.

With respect to image I and the width to be reduced w, formally, we define *Candidate Set* (*CS* for short) recording when *ATF* increases

Table 1
Notations.

Notations	Description
$I_{i,j}$	The pixel in i-th row and j-th column of image I
I^b	A blurred image of I
I_t / I_{-t}	Image after deleting/adding t vertical seams from/into I
$\mathcal{F}_{k_1,k_2}(I)$	Result by first deleting k_1 vertical seams from I and then scale to further reduce k_2 width
$ATF(t)$	Artifacts measure for I_t
$CS(I,w)$	Candidate set for shrinking the width of an image I by w

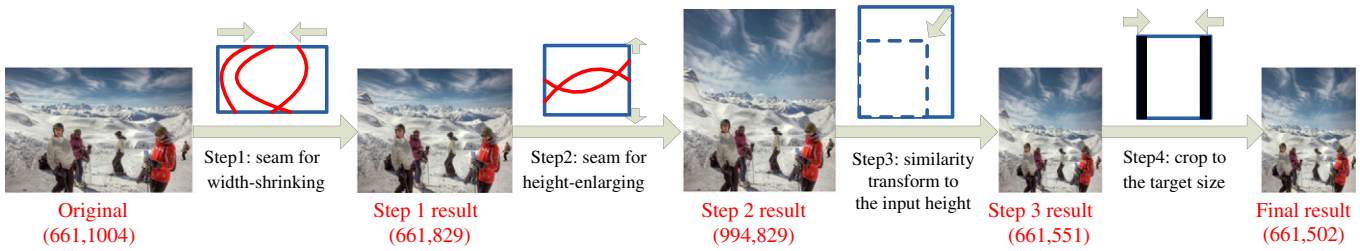

Fig. 2. The general procedure of our framework to shrink the width of an image by half. Original: input image of size 661×1004; step 1 result: shrink the width by 175 pixels (661×829); step 2 result: enlarge the height by 333 pixels (994×829); step 3 result: scale to 661×551 to achieve the same height as original; step 4 result: crop the width by 49 pixels to achieve the desired size.

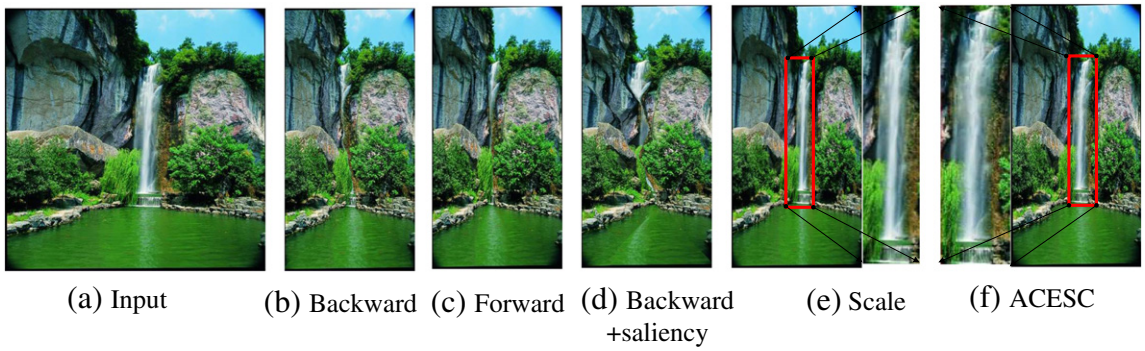

Fig. 3. (a) Input; retargeting results of (b) backward *SC* [18]; (c) forward *SC* [4]; (d) backward *SC* with saliency; (e) bicubic scaling (f) *ACESC*. There are noticeable artifacts caused by pixel over-deletion in backward/forward *SC* methods. Compared with (e), our result is more content-aware (*i.e.*, the larger width of the fall).

as in Eq. (5). According to this strategy, the range of u in Eq. (4) (*i.e.*, $0 \le u_0 \le n - n'$) is restricted to $CS\ (I, n-n')$.

$$CS(I,w) = \{t | 0 \le t \le w \wedge ATF(t+1) > ATF(t)\} \tag{5}$$

We employ a similar technique to speed up the computation of u'. Specifically, we iteratively insert horizontal seams (seam carving for enlarging) into image I_u. When seam artifacts increase by inserting a seam, we further enlarge the height of the intermediate image by scaling to reach the desired aspect ratio ($\frac{n'}{m}$). For each generated image we select the most similar one to the original image, and u' is the number of horizontal seams inserted for producing the selected image.

4.2.1. Time complexity analysis

ATF $(t+1)$ computation can conveniently utilize $ATF(t)$, as only very limited patch pairs are affected after removing a seam from I_t. Such an incremental computation is very efficient and can be neglected compared with that of computing *BDW*. As a result, employing *Candidate Set* reduces the computing cost.

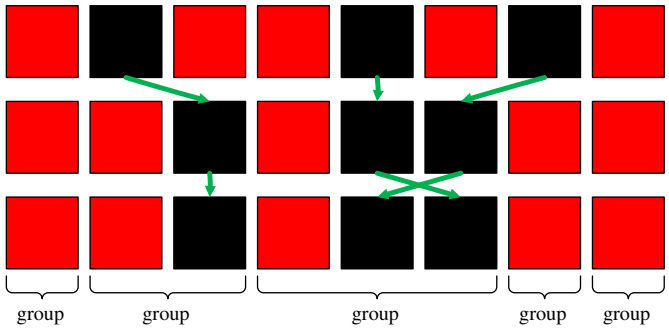

Fig. 4. Illustration of the seam map and pixel groups for SWD_{shr}. A 3×8 image I is carved by 3 seams (black), resulting the seam map $SM(I,I_3)$.

4.3. Improving the operator: ACcumulated Energy Seam Carving (ACESC)

In our proposed multi-operator scheme, direct and indirect seam carving techniques are two basic operators. However, the original seam carving is prone to delete too many pixels in certain region of the image in many cases, generating undesirable artifacts, as shown in Fig. 3(b) and (c). Even with the help of saliency maps [27,32–34], the result can still be unsatisfactory, *e.g.*, Fig. 3(d). Furthermore, as we pointed out in Section 1, seam carving is less effective due to its weakness in large ratio enlarging.

To solve the above issues, we propose *ACESC* to refine the seam carving operator considering the accumulated effect of the sequence of removed/inserted seams. The main idea is to restrict carving too many seams in the same image region. In this model, we define Seam Warping Distance (*SWD*) as a global guidance of seam carving. We will detailedly introduce how to employ the *ACESC* model to deal with both shrinking and enlarging.

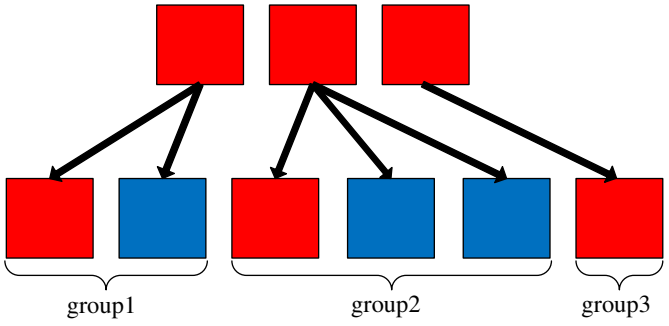

Fig. 5. Illustration of pixel groups for SWD_{enl}. A 1×3 image (top) is enlarged by 3 (bottom). Blue pixels are duplicated (the arrows connect the identical pixels). Any pixel and its duplicates are put into the same group, as shown by under braces in the figure.

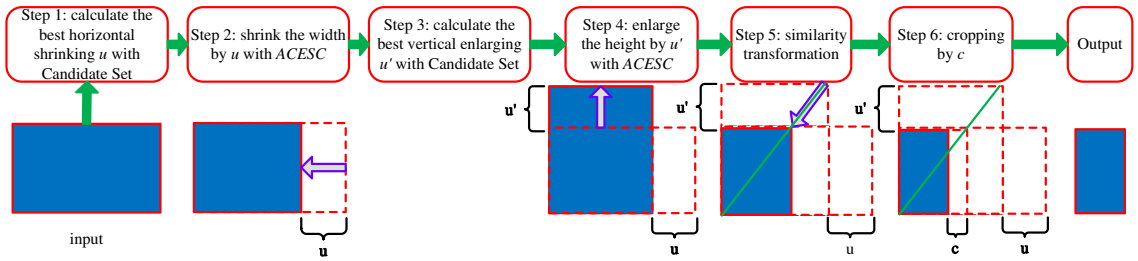

Fig. 6. The flowchart of our multi-operator scheme. The blue boxes refer to input image (under Step 1), intermediate results (under Step 2 to 5) and output image (under Step 6).

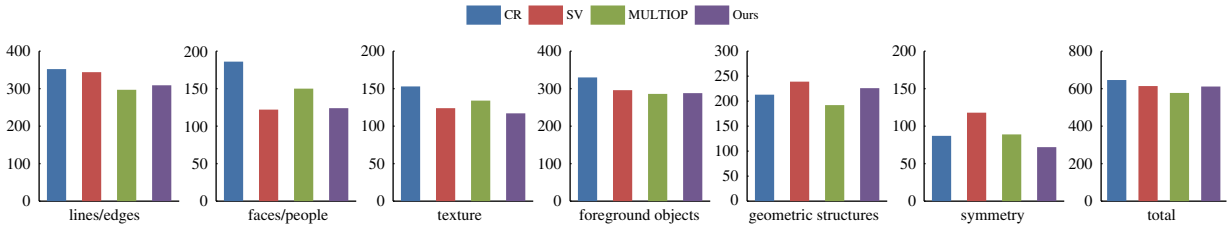

Fig. 7. The number of votes in 6 image classes and total results of four methods, when the reference (original) image was shown.

4.3.1. Shrinking the width

Basically, *ACESC* follows the procedure of original seam carving [18] (*i.e.*, repeatedly deleting min-seams from the image) except how it extracts the min-seams. If currently t seams have been deleted from image I, resulting image I_t, then the key of seam carving is to search the min-seam of I_t. Technically, it is a problem about how to define the energy (importance) of each pixel in I_t. Our model introduces the following steps to characterize pixel energy:

- (Preprocessing) Assign each pixel in I a newly defined pixel value.
- Construct a seam map SM for each image I_t to consider the accumulated effect of seam carving.
- Design a global cost criterion, Seam Warping Distance (SWD), based on SM to estimate the seam distortion in I_t.
- Calculate the seam energy based on SWD to prevent pixel over-deletion in the procedure of seam carving.

To begin with, we assign each pixel value of I:

$$PV = \mathbb{I}\left(I \times \beta + I^b \times (1-\beta)\right) \tag{6}$$

where I^b is obtained with a Gaussian low-pass filter with size $b \times b$ and variance $\sigma^2 = 9$ on I. The sum $I \times \beta + I^b \times (1-\beta)$ with a weight of β is a mixed image of original image and its blurred version. Notation $\mathbb{I}(\cdot)$ retrieves an intensity map. The reason of using I^b is that the smoothness resulting from the Gaussian low-pass filter allows that the pixel values contain more surrounding information, which is closely related to the structure of an image. Meanwhile, the tuning of b has very little influence on the performance as long as b ranges between [10,30]. We empirically set $b = \sqrt{\frac{height \times width}{2000}}$ to adapt to different image sizes. Parameter β is set to be 0.2 in our experiment.

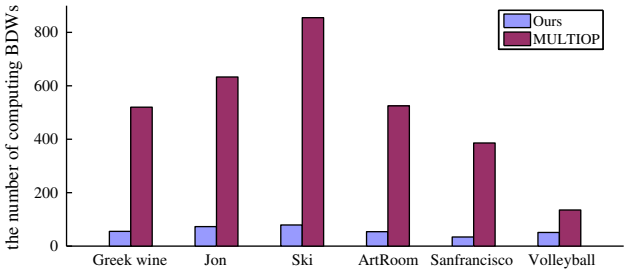

Fig. 8. Comparison between our method and the previous multi-operator method [1] with respect to the number of computing *BDWs*.

We maintain a seam map $SM\ (I, I_t)$, which is the same size as the original image I, to record which pixel of I still remains in I_t. For illustration, each pixel of the seam map is colored to indicate whether the pixel is remained in I_t, as shown in Fig. 4. The red pixel means it remains in I_t while black indicates not. It is clear that $SM\ (I, I_t)$ describes position relationship between the deleted pixels and remaining pixels of I_t.

Once the seam map is constructed, we group its pixels for later calculation of the seam warping distance between I and I_t, denoted as $SWD_{shr}(I, I_t)$. Each red pixel and all contiguous, adjacent black pixels to the right of this red pixel are put into one group. To deal with the remaining black pixels, if any, at the beginning of each row, we pad a left-most column containing all red pixels with the maximal pixel value of this image. Obviously, every group contains only one red pixel (remained in I_t or the auxiliary pixel padded) and several neighboring black pixels (deleted from I). Then we define *group cost* as $cost_{shr}(g) = (a-b) \times s$ for group g. Here a, b and s mean the maximal pixel value, the minimal pixel value and the number of pixels in g, respectively. Intuitively, the value $(a–b)$ estimates the influence of pixel-deletion in this region, and a large value of $(a–b)$ indicates the pixels in this group are dissimilar. Meanwhile, the value s counts the frequency of pixel-deletion and plays a role of penalizing over-deletion. Overall, $cost_{shr}(g)$ depicts the degree of pixel deletion in certain non-homogeneous regions. With the above definitions, SWD is defined as the sum of group costs:

$$SWD_{shr}(I, I_t) = \sum_{g \in Gp(I, I_t)} cost_{shr}(g) \tag{7}$$

where $Gp(I, I_t)$ is the set of groups. $SWD_{shr}(I, I_t)$ measures the degree of the pixel over-deletion in non-homogeneous regions of I_t.

To generate low $SWD_{shr}(I, I_t)$ value for every t, we define the seam energy of I_t as follows:

$$E_t(seam) = SWD_{shr}(I, I_t - seam) - SWD_{shr}(I, I_t) \tag{8}$$

The definition of the group cost guarantees that $E_t(seam)$ is non-negative. Based on this definition, therefore, in every *ACESC* iteration

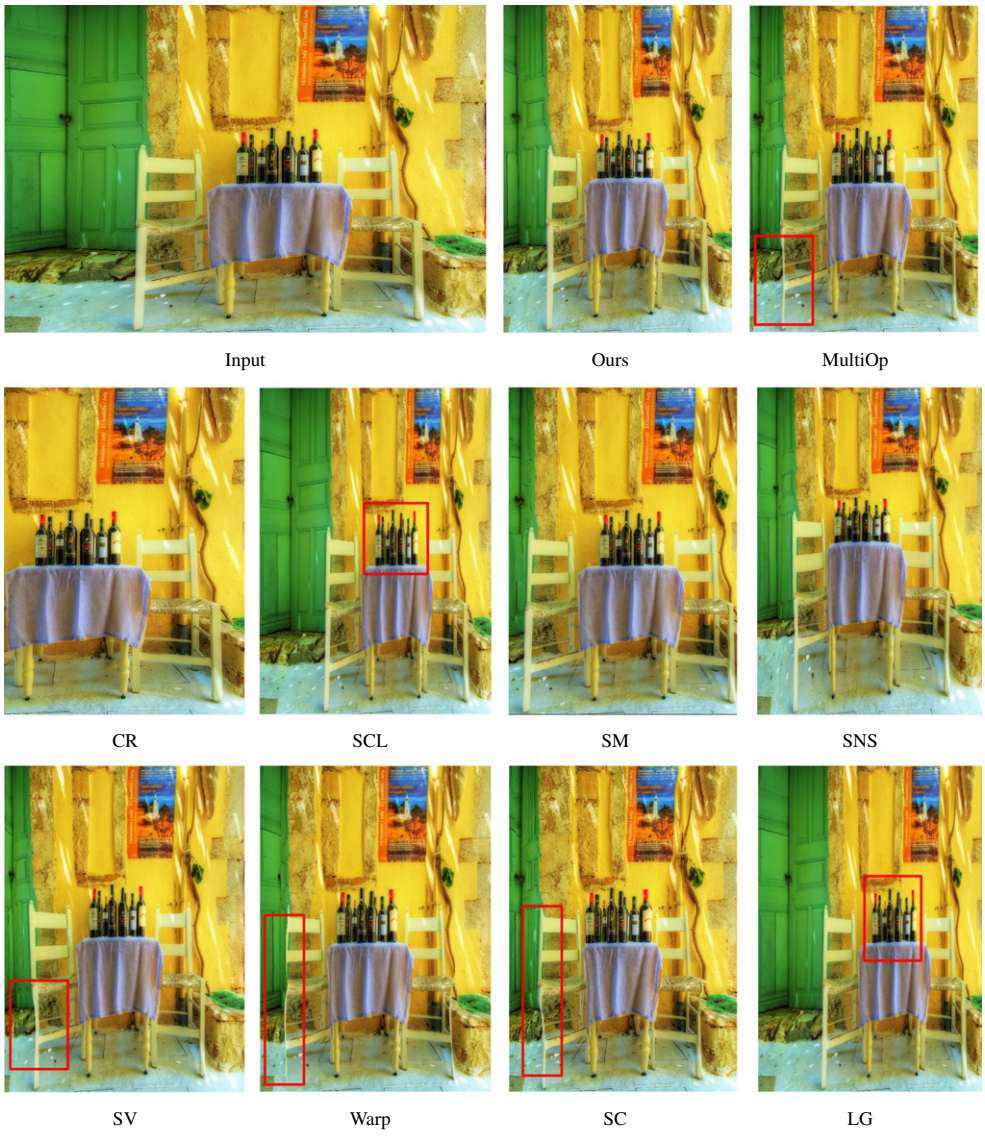

Fig. 9. Image retargeting results generated by our method and SM, SNS are more visually pleasing. The chair legs are severely distorted in results of MultiOp, SV, Warp, and SC. CR incurs an unacceptable content lost. LG and SCL squeeze the bottles.

the removal of the min-seam causes the least increase of SWD_{shr}. Note that although we only take horizontal coherence into account for defining the seam energy, the vertical coherence can also be ensured since we always carve out seams. Implementation details of *ACESC* are shown in 2.

4.3.2. Enlarging the width

To enlarge the image, as mentioned in Section 1, problems occur when the enlarging ratio is large, especially when the added width n_1 is larger than the original width n. The problem is even worse in our multi-operator framework since it does not restrict the expanding ratio with respect to n. We propose SWD_{enl} (SWD for enlarging) to address this image enlarging issue. With this newly defined seam energy, seam carving for enlarging is performed by iteratively inserting seams. It will not insert the same seam as original seam carving does since in our model the position of the min-seam varies as t increases. With this new model, furthermore, the enlarged size is not restricted by the original size.

Similar to SWD_{shr}, we group the pixels of I_{-t}, define the group cost and sum up the costs as SWD_{enl}. Unlike shrinking, here the pixels in I form a subset of the pixels in I_{-t}, where I_{-t} is formed by duplicating seams into image I. Therefore, grouping strategy is different from shrinking. Specifically, the pixels in I_{-t} corresponding to the same pixel in I form a single group, as shown in Fig. 5. To define the group cost, we first calculate the gradient values of I as: $G_{i,j} = |\mathbb{I}_{i,j+1} - \mathbb{I}_{i,j}| + |\mathbb{I}_{i+1,j} - \mathbb{I}_{i,j}|$. Then for each group, we denote the number of pixels in this group as s and the common gradient value as μ. We then define the cost of a group g as $cost_{enl}(g) = \mu \times 2^{s-1}$. The reason for using the exponential form as the group cost is to avoid duplicating the same pixel too many times.

Let $Gp(I, I_{-t})$ be the group set of I_{-t}, SWD_{enl} is calculated using Eq. (9). The seam energy is decided by Eq. (10).

$$SWD_{enl}(I, I_{-t}) = \sum_{g \in Gp(I, I_{-t})} cost_{enl}(g) \tag{9}$$

$$E_{-t}(seam) = SWD_{enl}(I, I_{-t} + seam) - SWD_{enl}(I, I_{-t}) \tag{10}$$

To avoid distractions, we show more corresponding technical details in Eq. 3.

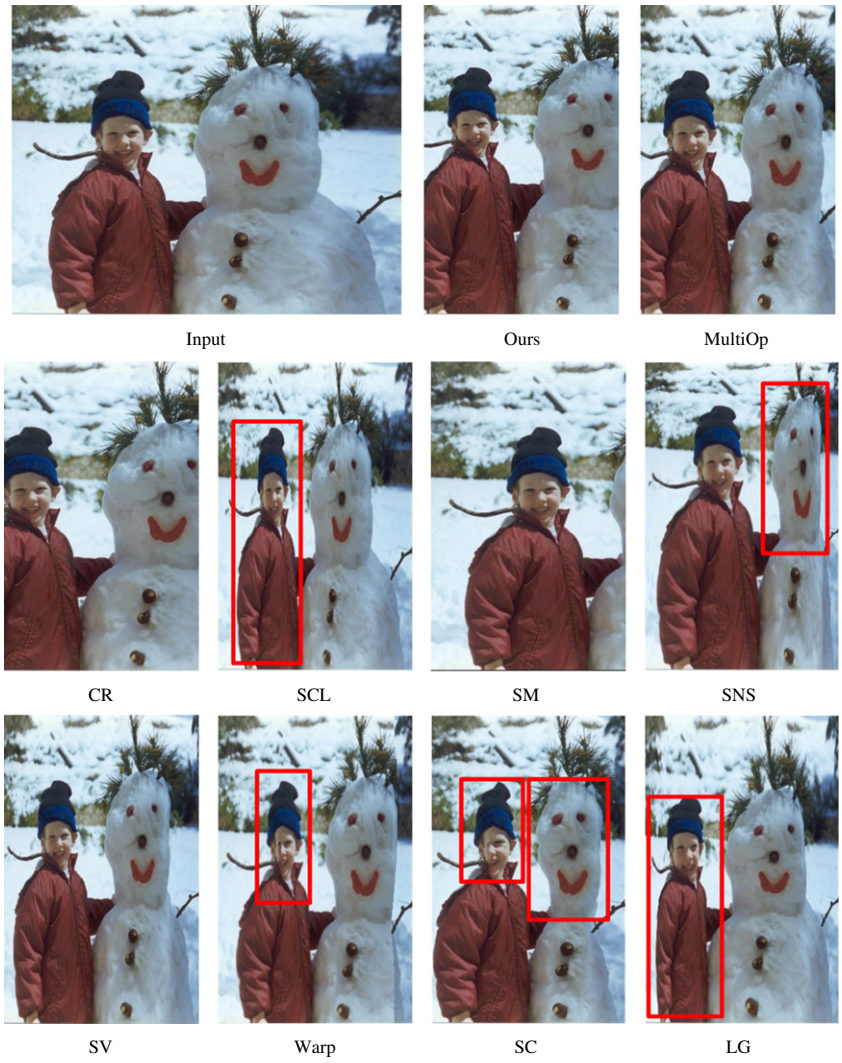

Fig. 10. Image retargeting results generated by our method and other methods. Our method crops less compared with CR and SM, and produce less artifacts than the other methods (shown in red box).

4.4. Scheme summary

Until now, we summarize the whole image retargeting procedure in more detail (compared with scheme overview in Section 1). Taking a task of resizing $m \times n$ image I to $m \times n'$ as an example, we detail the procedure as follows:

- Goal: Resizing an $m \times n$ image I to $m \times n' (n' < n)$.
- Step 1: Find a candidate set $\mathcal{C}_1$ for shrinking the width by $n - n'$ (*i.e.*, calculate CS $(I, n - n')$ with Eq. (5)), and from $\mathcal{C}_1$ select an optimal number of vertical seams to be removed.
- Step 2: Perform *ACESC* to shrink the width of I by u to reduce the image I to the reduced I_u of size $m \times (n - u)$.
- Step 3: Find a candidate set $\mathcal{C}_2$ for enlarging the height of I_u, and from $\mathcal{C}_2$ select an optimal number of horizontal seams to be inserted.
- Step 4: Perform *ACESC* to enlarge the height of I_u by u' to enlarge the image to size $(m + u') \times (n - u)$.
- Step 5: Perform a similarity transformation to scale the intermediate image to the original height m (image of size $m \times ((n - u) \times m / (m + u'))$).
- Step 6: Crop the intermediate image to shrink the width to the target width n' as the final result. We employ a simple cropping in this step to remove the left most or right most column iteratively, depending on which column has less total saliency values [34].

The corresponding flowchart is illustrated in Fig. 6.

5. Results and discussions

In this section, we compare our proposed multi-operator scheme with several recently published algorithms including: 1) multi-operator (MultiOp) [1]; 2) streaming video (SV) [2]; 3) shift-maps (SM) [28]; 4) scale-and-stretch (SNS) [14]; 5) seam carving (*SC*) [4]; 6) nonhomogeneous warping (Warp) [3]; and 7) energy-based deformation (LG) [25]. We also include the results of 8) direct scaling operator (SCL), and 9) manually chosen cropping windows (CR) [6]. The comparison is performed in a benchmark image retargeting data set, *i.e.*, RetargetMe dataset [6]. The dataset consists of 80 different images with 6 different attributes including Lines/Clear Edges, Faces/People, Evident foreground objects, Geometric Structures, Symmetry and Texture elements/repeating pattern. Due to the space limitation, we only demonstrate some representative results in this paper, and show the complete results online.[1] Note that all

[1] http://people.csail.mit.edu/mrub/retargetme/contribute.html.

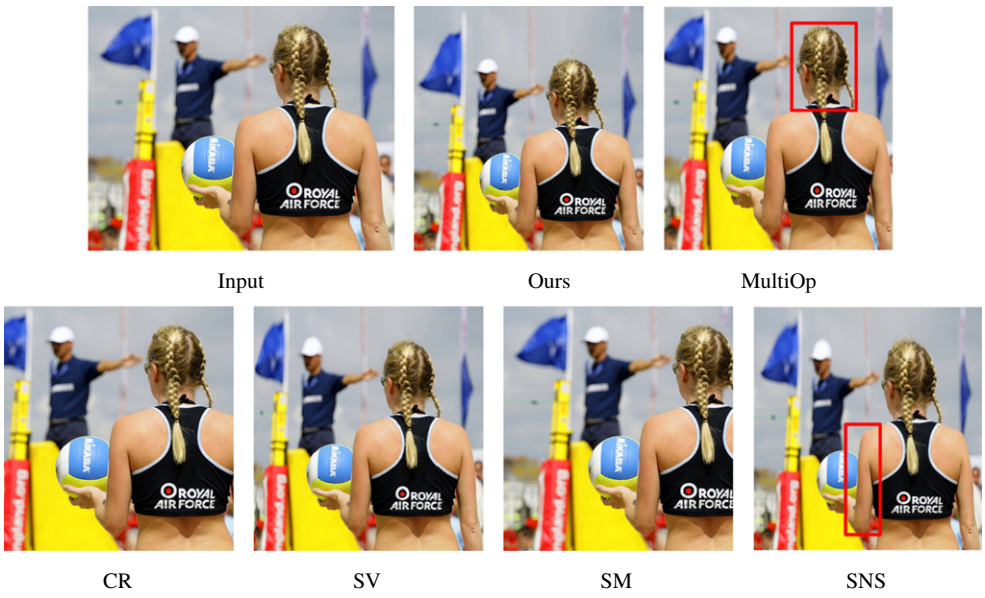

Fig. 11. Image retargeting results generated by our method and other methods. Our method and SV keep the girls' shape relatively better than MultiOp and SNS do (shown in red box). Less crop has been done compared with the results of SM and CR.

the results of other algorithms are obtained from Rubinstein et al.'s work [6]. Based on a web-interface image retargeting survey system provided by Rubinstein et al. [6], we also conducted a user study [6] to analyze the user preference of the results of different methods.

Parameter settings

There are several parameters to be predefined in our scheme. In the calculation of *CS*, we set patch size d to be 15×15 and 31×31 to calculate their respective *CS* and then union them as the final *Candidate Set*. We empirically set the patch size for calculating the *BDW* as 32×32. Furthermore, we select a considerable resizing ratios (25% or 50%) [6] for shrinking image.

5.1. Comparisons

Several results are shown in Figs. 9 to 12. In these cases, our method is more capable of preserving the shapes of important objects and structure (*e.g.*, the shape of the girl in Fig. 11 and the heart in Fig. 12) than the previous multi-operator scheme [1] and other methods . The main reason is that our multi-operator scheme exploits the resizing potential orthogonal to the resizing direction with indirect seam carving, which was not considered in the previous schemes. As a result, our two-dimensional seam carving protects the aspect ratio of the main object better for one-dimensional image retargeting.

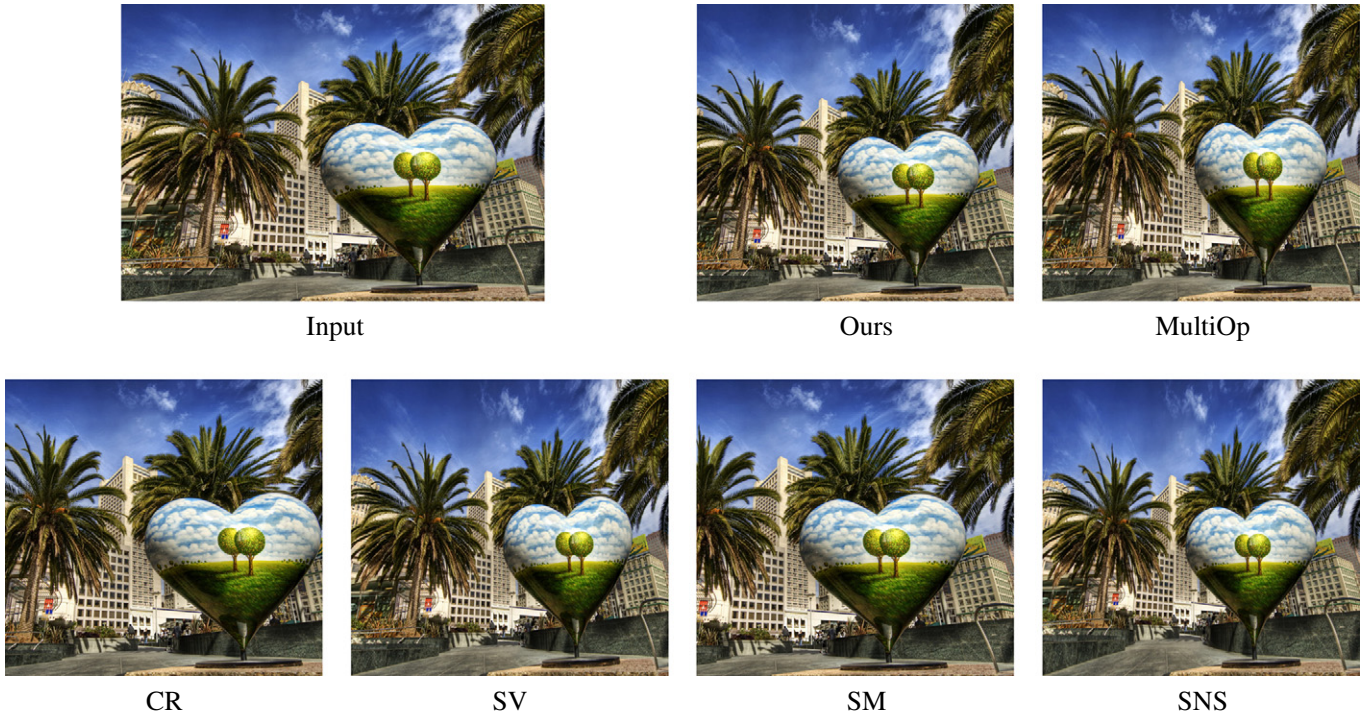

Fig. 12. Image retargeting results generated by our method and other methods. Our method and SNS better preserve the "heart shape" than MultiOp does, while incurring less cropping than SM and CR do. Compared to the result of SV, the "heart structure " keeps better by our method.
SNS result is similar to ours.

Table 2
Running times of computing *BDW* once.

Patch size	8×8	16×16	32×32
Overlapping	2.33 s	3.88 s	6.58 s
Non-overlapping	0.46 s	0.34 s	0.30 s

More specifically, in Fig. 12, our method performs indirect seam carving to avoid squeezing the heart shape, while the previous multi-operator scheme (*i.e.*, MultiOp) squeezes the heart shape because of only using one-directional scaling operator. Compared with one-directional warping [3] (Warp), similar effects of object squeezing can be observed in Figs. 9 and 10. During such warping, each pixel is shrunk to the same direction, unavoidably causing squeezing the image objects to some extent. Moreover, it can be observed from Figs. 9 to 12 that, by carefully combining *SC* and cropping, our method avoids cropping too much contents as the cropping method does, and meanwhile preventing from obvious seam distortion, which the original *SC* (Figs. 9, 10) may bring in.

Some of our results are similar to those of global warping based methods [14] [25] [2] since the flexible warping directions are also considered in those methods. The main difference between our method and theirs is twofold: (i) in our method, seam carving and homogeneous rescale are used to resize the main objects, whereas the saliency based warping are used in their methods; (ii) Our method employs cropping as an effective alternative to maintain the main structure in the image.

Since the output pixels come from the input image, shift-map [28] unavoidably discards a large part of the image content (Figs. 10 and 11). Although our method also employs cropping that may discard border contents, it only crops a little since most resizing job has been done very well by *SC*.

5.2. User study

User study acts as an important role in the evaluation of image retargeting quality [6]. In Rubinstein et al.'s user study [6], three retargeting methods (SV, MultiOp, and CR) have been experimentally proven to have better image retargeting qualities than other methods. In our user study, therefore, we perform a total of $\binom{4}{2} = 6$ pairwise comparisons with only the three methods. To reduce the workload of users, we select 39 images from the RetaretMe dataset. Each time 8 images are randomly drawn and 6 pairwise comparisons are tested (*i.e.*, a total of 48 comparisons) for each user per time.

A total of 52 valid volunteers took part in the test, casting a total of 2496 valid votes. We rank the four methods according to the number of votes received, *i.e.*, the number of times a method was preferred over a different method. According to six different attributes and total images, 7 ranking results are shown in Fig. 7. From the figure we can see that CR is the method the most preferred by users. However, the user preference of CR method highly relies on the *manually* chosen cropping windows. Among the other three automatic methods, our method and SV gains almost the same votes in total while slightly higher than MultiOp. Furthermore, we also perform a paired *t*-test to the four methods. With a statistical significance level of 5%, we find that the four methods are actually statistical indistinguishable, which means that our methods are comparable to other methods.

5.3. Effect of Candidate Set (CS)

According to the analysis of the significant speed-up of employing the proposed *CS* in Section 2, our method can approximately reduce the *BDW* computation cost by a factor $\frac{w}{|CS(I,w)|}$ compared with previous scheme [1], where w is the width to be reduced and $|CS(I,w)|$ is the size of $CS(I,w)$. The actual number of computing *BDW* is shown in Fig. 8, including Ski (661×1003, resizing ratio 50%, Fig. 1), Green-wine (697×1024, 50%, Fig. 9), Jon (397×499, 50%, Fig. 10), Volleyball (400×500, 75%, Fig. 11), San Francisco (718×1024, 75%, Fig. 12) and ArtRoom (813×1024, 75%, Fig. 15). The ratio of computing *BDW*s between previous multi-operator [1] scheme and our method is in an interval of [3,10] or wider, varying according to the contents and sizes of images.

5.4. Efficiency

We report the running time of our method. The software environment is based on Visual Studio 2008. The computer used for the evaluation is with an Intel Xeon Duo E5520 2.27GHZ and 2 GB of DDR3 memory. The running time is dominated by the computing of *BDW* and closely related to the parameters (or strategies) of computing *BDW*. In Table 2, we give the running times of computing *BDW* once for a 600×600 image with respect to different settings (*i.e.*, patch size and whether patches allow overlapping). When implementing patch-based *BDW*, if the patches allow being overlapped, the results are slightly better than that of not allowing patch overlapping, but the running time is several times longer. Taking 8×8 patch-based *BDW* for shrinking half width of a 600×600 image as an example, allowing patch-overlapping typically runs 5–10 min while not allowing patch overlapping takes only 1–2 min. Compromising between running time and quality is using restricted overlapping (with respect to maximal overlapping) which requires certain "distance" between central points of different patches.

5.5. Limitations

Finally, we would like to discuss the limitations of our method. Generally speaking, there are two main factors which affect the retargeting quality, *i.e.*, inappropriate combination of operators and ineffectiveness of the employed operators. The former is mainly caused by the

Fig. 13. Illustration of inappropriate combination of operators. Butterfly: too many vertical seams are deleted, causing distortion. Eagle: crop too much content.

(a) Input (b) Ours (c) SC

Fig. 14. Illustration of seam distortions carried over to our method. (a) Input image; (b) our result; (c) SC result.

combined effect of the *BDW* measure and the *CS*. For example, if the *CS* fails to include a number close to the optimal, the result can suffer severe seam distortions (too much seam carving) or a large content lost (too much cropping). "Butterfly" and "eagle" are two typical failure cases of this type, as shown in Fig. 13. The distortions in the butterfly image result from removing too many vertical seams, while the content lost in the eagle image is due to performing too less seam carving and too much cropping. A better combination of operators can help improve the retargeting quality. Nevertheless, the result may still suffer the ineffectiveness of the employed operators, which is a limitation sharing with all other multi-operator frameworks. Specifically, limitations from the employed operators (*SC*, cropping) will also be carried over to our solution. An example of this failure type is shown in Fig. 14. Due to the fact that seams are discrete, performing seam carving will unavoidably generate undesirable distortions of the smooth curves in the image (Fig. 14(b) and (c)). Compared to only performing seam carving (Fig. 14(c)), however, the operator combination helps relieve the undesirable distortions caused by seam carving as less seam carving is performed (Fig. 14(b)). Besides, our method is less effective to retarget images that contain very limited homogeneous regions, similar to most of the retargeting methods. For example, in Fig. 15, our method automatically chooses cropping, causing a content lost, and meanwhile the MultiOp method [1] performs somewhat like scaling to avoid content lost in this case. However, either cropping or scaling can be a good choice in this case since it definitely avoids severe seam distortions. Another drawback of all the multi-operator methods is the efficiency. Although we have significantly improved the efficiency with the proposed *Candidate Set* strategy, the running time is still stuck by the costly image similarity measure computation. Since the image similarity measure is not the main concern of this paper and our speed-up scheme is independent of the measure, our valuable contributions are non-negligible and the efficiency problem of the multi-operator has close relationship to the progress of image similarity measure.

6. Conclusion

In this paper, we propose a new multi-operator framework that automatically combines direct seam carving, indirect seam carving, similarity transformation and cropping. Our contributions are three-fold: 1) we put forward a multi-operator retargeting framework towards automatic combination of direct and indirect seam carving; 2) we propose a measure to estimate the artifacts caused by seam carving and select the candidate set to speed up the existing multi-operator retargeting framework; 3) moreover, we put forward a novel seam carving model, *ACESC*, that alleviates the pixel over-deletion effect and thus enhance the potential of preserving the structure. Experiments in a benchmark image retargeting dataset show that our method has comparable performance with other state-of-the-art algorithms in terms of human visual perception and user study.

In the future, we will investigate how to further accelerate the proposed scheme. Furthermore, we will pay more attention to the preservation of structure in an image, which is very important to the development of image retargeting.

Acknowledgment

We would like to thank Dr. Michael Rubinstein for valuable suggestions on the multi-operator retargeting algorithm. We would also like to

Ours MultiOp

Fig. 15. Image retargeting results for the image with little homogeneous regions. Our method resorts to cropping to avoid severe seam distortions when the *SC* method fails, while MultiOp result performs somewhat like scaling, avoiding content lost. Usually resorting to either cropping or scaling is a good choice in such scenario.

thank the reviewers for their various comments, which helped improve this paper. This work was partially supported by the National Natural Science Foundation of China (NNSFC no. 60975044) and the National Fundamental Research Program of China (no. 2010CB327900).

Appendix A. The implementation of *ACESC* for shrinking images

We present pseudo code of *ACESC* in Algorithm. 2. Similar to the classical seam carving, the seam carving process of *ACESC* is to iteratively find and delete a min-seam from the input image. The difference between *ACESC* and the classic seam carving is that the energy of *ACESC* is recalculated after each seam removal. For the purpose of the energy calculation and updating, three maps including maximal map A, minimal map B and number map $\mathcal{N}$, are constructed to record the maximal, minimal values, and the number of pixels of every group in the intermediate seam carved image, respectively. Note that if currently t seams have been carved out, then intermediate latest seam carved image is I_t. Lines 2–3 are the initialization of maps, indicating that A and B are initially with pixel values defined in Eq. (6). Function *SHRINKONE* shrinks the width by one pixel. $PE_{i,j}$ records the pixel energy of $(I_t)_{i,j}$ by Eq. (1).

$$PE^s_{i,j} = \left(max\left\{A_{i,j-1}, A_{i,j}\right\} - min\left\{B_{i,j-1}, B_{i,j}\right\}\right) \times \left(\mathcal{N}_{i,j-1} + \mathcal{N}_{i,j}\right) - \left(A_{i,j-1} - B_{i,j-1}\right) \times \mathcal{N}_{i,j-1} - \left(A_{i,j} - B_{i,j}\right) \times \mathcal{N}_{i,j} \quad (1)$$

The above equation for calculating the pixel energy can be easily induced from the seam energy calculation depicted by Eq. (8). In other words, it is the increment of SWD_{shr} if pixel $(I_t)_{i,j}$ is removed. Dynamic programming is performed (line 9) to calculate the min-seam in the intermediate image, which is the same implementation as original seam carving. The min-seam is backtracked and the maps are updated and shrink the image width by 1 in line 10–11.

The update of maps in line 10 is triggered before removing a seam. If a pixel in i-th row and j-th column in the intermediate image is to be deleted, the group described by $(A_{i,j}, B_{i,j}, \mathcal{N}_{i,j})$ and $(A_{i,j-1}, B_{i,j-1}, \mathcal{N}_{i,j-1})$ will be merged into one group described by $(max\{A_{i,j}, A_{i,j-1}\}, min\{B_{i,j}, B_{i,j-1}\}, \mathcal{N}_{i,j} + \mathcal{N}_{i,j-1})$.

$$\begin{cases} A_{i,j-1} = \max\left\{A_{i,j-1}, A_{i,j}\right\} \\ B_{i,j-1} = \min\left\{B_{i,j-1}, B_{i,j}\right\} \\ \mathcal{N}_{i,j-1} = \mathcal{N}_{i,j-1} + \mathcal{N}_{i,j} \\ A_{i,k} = A_{i,k+1}, \quad k = j, j+1, \ldots, n-1 \\ B_{i,k} = B_{i,k+1}, \quad k = j, j+1, \ldots, n-1 \\ C_{i,k} = C_{i,k+1}, \quad k = j, j+1, \ldots, n-1 \end{cases} \quad (2)$$

Algorithm 2 $I = ACESC_{shr}(I, w)$

```
INPUT: I is the input image; w is the width to be reduced.
OUTPUT: the reduced image I_w.
1: allocate A, B, N, PV, PE with same size of I
2: compute PV using Eq. (6)
3: assign PV to A and B
4: assign N with all 1
5: for i:=1 to m do
6:    SHRINKONE
7: procedure SHRINKONE
8:    compute PE using Eq. (1)
9:    use dynamic programing to search min-seam
10:   backtract the min-seam and update with assignment (.2)
11:   remove min-seam
12: end procedure
```

Appendix B. The implementation of *ACESC* for enlarging images

The pseudo code of enlarging can be referred to Algorithm 4. Initially, map N records the numbers of pixels in each group of the intermediate image (line 5), and map G records gradient values of input image I (line 6). Here the intermediate image refers to currently obtained intermediate results during seam carving process. With map N, G and Eq. (10), pixel energy of the intermediate image then can be formulated with Eq. (3) (line 9).

$$PE^e_{i,j} = G_{i,j} \times 2^{N_{i,j}} - G_{i,j} \times 2^{N_{i,j}-1} \quad (3)$$

Dynamic programming is performed to search the min-seam (line 10). After each min-seam insertion happens, the values correspond to the inserted seam pixels in map G will be duplicated (line 13). N is updated in line 14–16 after each seam insertion to maintain the counting of group pixels for each pixel in the intermediate image.

Algorithm 3 $I = ACESC_{enl}(I, w)$

```
INPUT: the input image I; w is the width to be added.
OUTPUT: the enlarged image I_w.
VARIABLE: N — the numbers of pixels in each group; G — the gradient values;
PE — the pixel energy;
1: (m,n) = size(I)
2: allocate N[m][n+w], G[m][n+w], PE[m][n+w]
3: for i:=1 to m do
4:    for j:=1 to n do
5:       N_{i,j}:=1
6:       G_{i,j}:=Gradient of I_{i,j}
7: call ENLARGEONE w times
8: procedure ENLARGEONE
9:    compute PE with Eq. (3)
10:   use dynamic programing to search min−seam
11:   for each pixel(u,v)in the seam do
12:      duplicate I_{u,v} on its right
13:      duplicate G_{u,v} on its right
14:      duplicate N_{u,v} on its right
15:      for all I_{u,x} identical to I_{u,v}
16:         N_{u,x} = N_{u,x} + 1
17: end procedure
```

Appendix C. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.imavis.2012.06.008.

References

[1] M. Rubinstein, A. Shamir, S. Avidan, Multi-operator media retargeting, ACM Trans. Graphics. 28 (3).
[2] P. Krähenbühl, M. Lang, A. Hornung, M. Gross, A system for retargeting of streaming video, ACM Trans. Graph. 28 (2009) 126:1–126:10.
[3] L. Wolf, M. Guttmann, D. Cohen-Or, Non-homogeneous content-driven video-retargeting, In: International Conference on Computer Vision, 2007, pp. 1–6.
[4] M. Rubinstein, A. Shamir, S. Avidan, Improved seam carving for video retargeting, ACM Trans. Graphics. 27 (3).
[5] F. Liu, M. Gleicher, Automatic image retargeting with fisheye-view warping, In: ACM Symposium on User Interface Software and Technology, 2005, pp. 153–162.
[6] M. Rubinstein, D. Gutierrez, O. Sorkine, A. Shamir, A comparative study of image retargeting, ACM Trans. Graphics. 29 (5).
[7] V. Daniel, T. Matthew, P. Kari, M. T., G. Natasha, A survey of image retargeting techniques, In: Proc. SPIE, , 2010.
[8] H. Liu, X. Xie, W.-Y. Ma, H.-J. Zhang, Automatic browsing of large pictures on mobile devices, In: ACM Multimedia, , 2003, pp. 148–155.
[9] B. Suh, H. Ling, B.B. Bederson, D.W. Jacobs, Automatic thumbnail cropping and its effectiveness, In: UIST'03: Proceedings of the 16th annual ACM symposium on User interface software and technology, ACM, New York, NY, USA, 2003, pp. 95–104.

[10] A. Santella, M. Agrawala, D. DeCarlo, D. Salesin, M. Cohen, Gaze-based interaction for semi-automatic photo cropping, In: Proc. the SIGCHI conference on Human Factors in computing systems, 2006, pp. 771–780.
[11] G. Ciocca, C. Cusano, F. Gasparini, R. Schettini, Self-adaptive image cropping for small displays, IEEE Trans. Consum. Electron. 53 (4) (2007) 1622–1627.
[12] R. Gal, O. Sorkine, D. Cohen-Or, Feature-aware texturing, In: Proc. Eurographics Symposium on Rendering, 2006, pp. 297–303.
[13] Y.F. Zhang, S.M. Hu, R.R. Martin, Shrinkability maps for content-aware video resizing, Comput. Graphics Forum 27 (7) (2008) 1797–1804.
[14] Y.-S. Wang, C.-L. Tai, O. Sorkine, T.-Y. Lee, Optimized scale-and-stretch for image resizing, ACM Trans. Graphics. 27 (5) (2008).
[15] G.-X. Zhang, M.-M. Cheng, S.-M. Hu, R.R. Martin, A shape-preserving approach to image resizing, Comput. Graphics Forum 28 (7) (2009) 1897–1906.
[16] Y. Guo, F. Liu, J. Shi, Z.-H. Zhou, Image retargeting using mesh parameterization, IEEE Trans. Multimedia 11 (5) (2009) 856–867.
[17] D. Simakov, Y. Caspi, E. Shechtman, M. Irani, Summarizing visual data using bidirectional similarity, In: CVPR, 2008.
[18] S. Avidan, A. Shamir, Seam carving for content-aware image resizing, ACM Trans. Graphics. 26 (3) (2007).
[19] W. Dong, N. Zhou, J.-C. Paul, X. Zhang, Optimized image resizing using seam carving and scaling, ACM Trans. Graphics. 29 (5) (2009).
[20] Y.-S. Wang, H.-C. Lin, O. Sorkine, T.-Y. Lee, Motion-based video retargeting with optimized crop-and-warp, ACM Trans. Graphics. 29 (4) (2010).
[21] Y.-F. Ma, H.-J. Zhang, Contrast-based image attention analysis by using fuzzy growing, In: MULTIMEDIA'03: Proceedings of the Eleventh ACM International Conference on Multimedia, ACM, 2003, pp. 374–381.
[22] L. Chen, X. Xie, X. Fan, H.-J. Zhang, H.-Q. Zhou, A visual attention model for adapting images on small displays, Multimedia Syst. 9 (4) (2003) 353–364.
[23] Y. Hu, D. Rajan, L.-T. Chia, Robust subspace analysis for detecting visual attention regions in images, In: Proceedings of ACM Multimedia, 2005.
[24] T. Bosse, P.-P. van Maanen, J. Treur, A cognitive model for visual attention and its application, In: The 2006 IEEE/WIC/ACM Int. Conf. on Intelligent Agent Technology (IAT-06), 2006, pp. 255–262.
[25] Z. Karni, D. Freedman, C. Gotsman, Energy-based image deformation, Comput. Graphics Forum 28 (5) (2009) 1257–1268.
[26] P. Viola, M.J. Jones, Robust real-time face detection, Int. J. Comput. Vis. 57 (2) (2004) 137–154.
[27] L. Itti, C. Koch, E. Niebur, A model of saliency-based visual attention for rapid scene analysis, IEEE Trans. Pattern Anal. Mach. Intell. 20 (11) (1998) 1254–1259.
[28] Y. Pritch, E. Kav-Venaki, S. Peleg, Shift-map image editing, In: ICCV, 2009.
[29] M. Grundmann, V. Kwatra, M. Han, I.A. Essa, Discontinuous seam-carving for video retargeting, In: CVPR, 2010, pp. 569–576.
[30] T. Deselaers, P. Dreuw, H. Ney, Pan, zoom, scan — time-coherent, trained automatic video cropping, In: CVPR, 2008.
[31] H. Sakoe, IEEE transactions on acoustics, Speech Signal Process. 26 (1978) 43–49.
[32] X. Hou, L. Zhang, Saliency detection: a spectral residual approach, In: CVPR, 2007.
[33] C. Guo, Q. Ma, L. Zhang, Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform, In: CVPR, 2008.
[34] R. Achanta, S. Hemami, F. Estrada, S. S., Frequency-tuned salient region detection, In: CVPR, 2009.