Balanced Error Distribution for Internet Video Retargeting

Hon-Hang Chang[1], Shwu-Huey Yen[2], Timothy K. Shih[1], Wee-Kheng Leow[3], Shih-Jung Wu[2]

[1] Dept. of Computer Science and Information Engineering, National Central University, Taiwan
[2] Dept. of Computer Science and Information Engineering, Tamkang University, Taiwan
[3] Dept. of Computer Science, School of Computing, National University of Singapore, Singapore
sicachang@gmail.com, 105390@mail.tku.edu.tw, timothykshih@gmail.com, leowwk@comp.nus.edu.sg

Abstract

With the progress of network technology, online video is widely used by users and popular. Image and video retargeting techniques have been proposed in the recent past by researchers around the world, low-complexity algorithms allow video to be transmitted more quickly. Our Balanced Error Distribution (BED) retargeting method follows the same concept like uniform scaling, but only removes unimportant regions, balancing main objects' error distribution. Priority map is proposed in this paper to maintain the structure of straight lines and irregular shape of objects without deforming complex image contents, which may be altered in traditional seam carving for complex image or video. In addition, the proposed mechanism adopts the general temporal coherence algorithm with strict condition (two continuous seams must be close as possible) to maintain visual continuity, such that the resulting video will not look shaky due to sudden changes in the background. The proposed method not only resizes the video with the retention of important information but also maintains the structural properties of objects in different kinds of videos.

Keywords: Video retargeting, Seam carving, Energy map, Saliency map

1 Introduction

The screen size of 3C devices is getting smaller than that of a standard laptop in recent years. Smaller screen allows the users to carry and use 3C devices everywhere, but makes it difficult for them to browse image and watch movie clearly as the image contents are scaled down. Image retargeting solves this problem by removing unimportant parts of the image contents while making the important objects more significant, so that the users can see the important objects more clearly. So, image retargeting can be applied to any 3C devices with small screen size, such as tablet PCs and mobile phones.

Video retargeting has 5 key difficulties that need to be addressed (see in Table 1). Image retargeting must address the first four difficulties D1-D4: OS (Maintain main object size), NO (Maintain number of main objects), GD (Reduce geometry distortion), and LD (Reduce linear structure distortion), whereas video retargeting has to address all 5 difficulties including TC (Maintain temporal coherence). So, video retargeting is more difficult than image retargeting because it has to address both spatial and temporal difficulties. Several improved video retargeting methods have been proposed in recent years to show better video retargeting results. Rubinstein et al. [18] proposed an improved video retargeting method based on seam carving. The algorithm can preserve the size of the main object but cannot avoid distorting geometry and linear structure. Wolf et al. [23] proposed another video retargeting method based on warping operator. It cannot fully maintain the size of the main object, the number of main objects, and it distorts geometry and linear structure. Another scaling based method is proposed by Yen et al. [24, 25], can minimize distortion of geometry and linear structure, and maintain temporal coherence. But it cannot avoid changing the size of the main object, because warping or scaling based retargeting methods cannot preserve the original shape of main object on some video with complicated background. Two examples with complicated background are shown in Figure 1. Suppose the initial mesh (the first frame) is perfectly defined (can predict the global saliency location in following frame), the object cannot be guaranteed which always in sparse grid as the mesh can be adjusted within a narrow range by 1-ring rule. So, no existent work can address all five difficulties (Table 2).

Table 1. Five difficulties in image/video retargeting

No.	Difficulty	Depiction
D1	OS	Maintain main object size
D2	NO	Maintain number of main objects
D3	GD	Reduce geometry distortion
D4	LD	Reduce linear structure distortion
D5	TC	TC: Maintain temporal coherence

*Corresponding Author: Timothy K. Shih; E-mail: timothykshih@gmail.com
DOI: 10.3966/160792642018091905018

(a) Source I: Rate (b) SCL (c) Targeted result [24]

(d) Source II: Woman (e) SCL (f) Targeted result [23]

Figure 1. Comparison of the size of main object between uniform scaling and two warping based method (Wolf et al.*'s* method [23] and Wang *et al's* method [24])

Table 2. Difficulties addressed by various video retargeting methods

Method	References	D1	D2	D3	D4	D5
Scaling	1-4, 24, 25	×	✓	✓	✓	✓
Cropping	20	✓	×	✓	✓	✓
Warping	6, 23	×	×	×	×	✓
Content-aware Warping	9-17, 27-28	×	✓	×	✓	✓
Seam Carving	7, 8, 18, 22, 25	✓	✓	×	×	✓
Our Proposed Method		✓	✓	✓	✓	✓

Our video retargeting method considers highest priority in temporal coherence part. The damaged sections of the main objects after the removal of the seams will be labeld with high enery to reduce the chance of the seams are defined in the same import regions.

One example of a Ball moving in a large motion is shown in Figure 2. (a), (d) and (g) are the retargeting results in the *k*-th frame f^k using warping, scaling, and proposed BED method. And, (b-c), (e-f) and (g-h) represents the retargeting results in the following two frames f_{k+1} and f_{k+2}. In order to satisfy the temporal coherence, the same spatial location or 1-ring neighborhood between consecutive frames which used in warping based (a-c) and scaling based (d-f) methods. Obviously, the ball will be severely deformed using restricting grid (could only deform in 1-ring neighborhood in consecutive frames) using warping and scaling based methods respectively. Also, it is possible the main object become too smaller in the 1-ring restriction especially in moving camera (see in Figure 1).

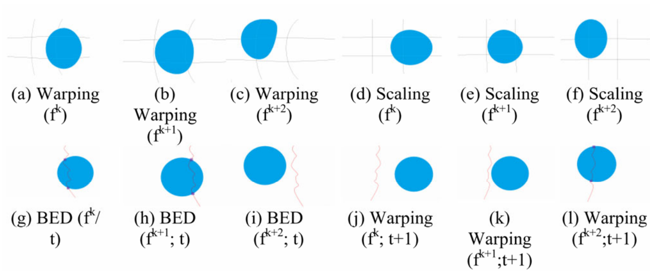


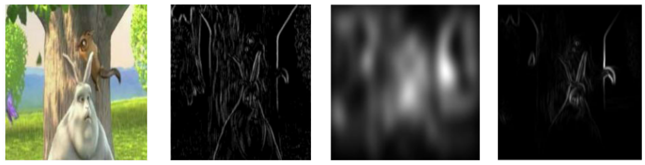
Figure 2. An example of video retargeting for large mortion of video (Fast-moving ball) using warping based (a-c), scaling based (d-f) and Proposed BED method (g-i) in consecutive frames f^k- f^{k+2}, respectively; (j-l) represents our results in next iteraciton t+1 of seam carving phase

In contrast, our proposed algorithm can address the spatial difficulty after following the stringent restrictions (1 pixel neighborhood between consecutive frames, the same with Rubinstein et al.'s method [18]). Even though it is possible the seam to pass through the object (Ball) under the temporal coherence, the proposed priority map in this paper records the damage of the object (purple block) in the t-th iteration, to avoid object deformation again in next iteration t+1 of seam carving phase (j-l), addressing the problem of some object are being damaged by a similar seam [18].

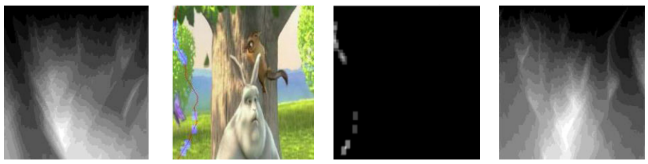
Seam carving is used to remove unimportant parts to reduce the image size. In warping and scaling based methods, some pixels of the main object are merged together to reduce the number of pixels belongs to the body. If we try to make the main object larger in the merging phase, it does not easily overcome the 3rd difficulty GD (it is easy to distort the shape of main object).

However, seam carving by itself may distort lines and shapes after removing the seams. To overcome this problem, our method labels the distorted regions of main objects after seam removing in a map called Priority Map (see in Figure 3 (g)) to avoid creating seams that pass through these damaged regions repeatedly.

The remaining sections are organized as follows. The methods used in the related works are described in Section 2 to highlight existing retargeting methods. Our proposed image and video retargeting methods are discussed in detail in Section 3. To verify the practical applicability of the proposed method, test results and detailed comparisons with other available methods are presented in the Section 4. Finally, Section 5 concludes this paper and highlights directions for future research.



(a) Original Fance I (b) Gradient Map G_k (c) Saliency Map S_k (d) Feature Map F_k



(e) Ek ($N = k$) (f) Optimal Seam (g) Priority Map (h) Energy Map

Figure 3. Proposed **BED** retargeting phase

2 Related Works

Scaling and cropping are the earliest methods to be used for image and video re-targeting [1-4]. In Kopf et al.'s method [4], saliency map [5] and face detection are used to identify region of interest (ROI) to be preserved in retargeting. Liu et al. [6] proposed an image retargeting method based on a fish eye warping technique that preserved important image contents after warping.

Saliency can be used as an important tool to define the important objects during image resizing. Wang et al. [8] combined both the gradient map and saliency map. Gradient map was used to preserve the structure of the image and saliency map to preserve important objects in the image, resulting in a better-retargeted image. Fang et al. [40] proposed a saliency model in the compressed domain. Fang et al.'s retargeting algorithm effectively preserves the visually important region in the image. Their method is only suitable for the images with limited main objects. It fails to maintain all the structure of the image content if the image contains too many main objects.

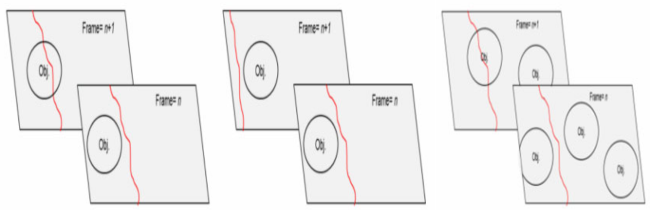
In addition, there are several improved retargeting methods [9-17, 36-37] based on content-aware mechanisms. Rubinstein et al. [19] proposed a novel retargeting technique by combing three operators, namely scaling, cropping, and seam carving, to achieve high-quality retargeted results. In recent years, a content-aware image retargeting method called Patch-Based Image warping is proposed by S. Lin [29]. A similarity transformation constraint is used to force visually salient content to undergo as-rigid-as-possible deformation. But, these methods cannot maintain the main object size after retargeting because of the use of warping and scaling in some iterations within retargeting procedure.

An advanced retargeting technique based on seam carving was originally proposed by Avidan et al. [7]. Seam carving is different from content-aware cropping, scaling and warping since it attempts to remove only unimportant pixels. The concept behind seam carving is to remove optimum seams vertically or horizontally to match the height and width of the target image size. However, it may distort line structure and geometry after removing the seams. In recent years, a great number of retargeting methods introduced by researchers can produce acceptable results only for specific image or video with few linear structures in the image. To solve this problem, Johannes et al. [26] proposed a tailored retargeting method for line preservation using seam carving operator. The limitation of Kiess et al.'s method is that it cannot handle a large number of linear structures in the input image. Therefore, the proposed method in this paper applies careful considerations on maintaining temporal coherence and preserving the structure of image contents to make the target image or video more pleasing to the viewers. Besides, Battiato et al. [30] proposed a content-aware image resizing technique allow to take into account the visual content of image during the resizing processes, then Battiato et al.'s method can preserve salient region better than other state-of-the art regions.

For video retargeting, special attention must be placed on the continuity of its consecutive frames in order to maintain the smoothness of the video. Keeping that in mind, Ariel et al. [18] proposed an improved video seam carving technique using the surface-carving approach to resize video with high temporal coherence. Showing that strict geometric continuity between the seams in two consecutive frames may not be necessary for achieving temporal coherence, Grundmann et al. [21] proposed a discontinuous seam carving approach. This strategy could maintain the temporal coherence by calculating the difference with two adjacent frames. In subsequent years, Rubinstein et al. [22] proposed a video seam carving method, which is capable of detecting whether the seams pass through the main object.

Some of the warping or scaling-based video retargeting methods [23-25, 38-39] could re-target the video with a higher temporal and spatial coherence but the retargeted results look similar to the results of uniform scaling. It means that the size of the main object cannot be preserved well. Li et al. proposed a video retargeting scheme that satisfy the two conditions, temporal coherence and undistorted-shapes of salient objects, simultaneously [39]. However, this method does not guarantee maintaining the original size of the objects. Some of the main objects in the Li et al.'s results become smaller in size, which is similar to the results of uniform scaling method. Liu et al. [20] proposed a window motion-based retargeting technique in contrast to other video retargeting methods. Therefore, the main object is located at the center of each frame in the retargeted video. But only part of ROI is retained after crop-ping. In recent year, Wang et al. [27] proposed a motion-based video retargeting technique with optimized crop-and-warp, to further improve the technique of [25]. In [25], the difficulty NO cannot be addressed as only the region of interesting (ROI) will be retain. So, some main objects will be ignored in resultant video. Wang et al. try to warp all information inside the ROI. To reduce motion distortion in resizing, an object-preserving warping is proposed by Lin et al. [28]. The basic idea is to measure content significance and resize videos by utilizing information of object motions rather than pixel motions.

So far, all existing retargeting methods produce fair results, but they cannot be applied to all available aspect ratios in different types of videos. Three scenarios related to temporal coherence and spatial difference with respect to seam carving, are shown in Figure 4. If temporal coherence is given a higher priority over spatial difference (a), the seam in frame n+1 will be identical to that in frame n. So, it is clear that, this method maintains temporal coherence in adjacency frames but cannot avoid placing the seam through the main object, thus changing the original size of the object. On the other hand, if careful consideration is applied to maintain spatial coherence (b), the seam can avoid going through the main object, but it causes shaking of the resultant video due to temporal incoherence.



(a) Temporal coherence with higher priority than spatial difference

(b) Spatial difference with higher priority than temporal coherence

(c) Weak case for maintaining spatial coherence

Figure 4. Comparison of the order of priority of temporal coherence and spatial difference

To address the problems illustrated in Figure 4, our video retargeting method considers highest priority in temporal coherence part. All seams in neighboring frame must be close. That is why we use the regular concept of temporal coherence (see in Figure 4 (a)). Our method marks all damaged main object part during the seam removal. Therefore, it avoids the main body to change smaller after multiple seams removal. It means that our method solves the 5th difficulty **TC** first, then used proposed priority map ***P*** to preserve the main body (**D1-D4**). However, it overcomes the 5 difficulties in our method as shown in Table 2.

After the removal of all defined unimportant regions, it is impossible that all the main objects will not be damaged if we use strict the temporal coherence (see in Figure 4 (a)), seams in the main body area will begin to remove.

3 Seam Carving with Error Distribution

Our proposed method is an extension of the original seam carving algorithm for image resizing [18, 26]. Compared to existing methods, our method can address all five difficulties in image and video retargeting discussed in Section 1. In comparison, no existing methods address all five difficulties. Before discussing the details of our method, we first review the original seam carving algorithm (Section 3.1). Next, we describe how our method of imposing spatial coherence with error distribution (Section 3.2) and temporal coherence (Section 3.3) address the five difficulties D1 to D5.

3.1 Seam Carving

A vertical seam S of an input image $I(x,y)$ is a set of connected pixels s_y, one at each row y of the image. Similarly, a horizontal seam is a set of connected pixels, one at each column. For the remainder of this paper, we will explain only for vertical seams. A vertical seam S is defined in terms of the *gradient map* $G(x,y)$ of image I. The *energy map* $E(x,y)$ of image I is computed iteratively from the gradient map $G(x,y)$ top down as follows.

$$E(x,y)=\begin{cases} G(x,y) & for\ y=1, \\ G(x,y)+\min_{i=-1}^{1} E(x+i,y-1) & for\ 1<y\leq H \end{cases} \quad \textbf{(1)}$$

where H is the height of the image. Then, the seam S of image I is the set of pixels $s_y = (x_y,y)$ that minimize the total energy E_T.

$$E_T=\sum_{y=1}^{H} E(x_y,y) \quad \textbf{(2)}$$

The seam S subject observes the connectivity constraint $|x_y - x_{y-1}| \leq 1$, for each $y > 1$. Therefore, the seam S contains pixels of least significance in the image. Energy minimization can be computed efficiently starting at the bottom of the image:

$$For\ y=H, s_H=(x_H,H),\ such\ that\ x_H=\arg\min_x E(x,H).$$

$$For\ y<H, s_H=(x_{y+1}+d,y),\ such\ that\ d=\arg\min_{i=-1}^{1} E(x_{y+1}+i,y)$$

$$s_y=\begin{cases} (x_H,H) & with\ x_H=\arg\min_x E(x,H)\quad for\ y=H, \\ (x_{y+1}+d,y)\ with\ d=\arg\min_{i=-1}^{1} E(x_{y+1}+i,y) & for\ y<H, \end{cases} \quad \textbf{(3)}$$

After computing the seam S, pixels along the seam in the image I are removed, reducing the width of I by one pixel. The pixels along the seam in the gradient map G are also removed, giving G the same width as I. To perform image retargeting by seam carving, the input image I is first scaled to fit the target height. Next, the above seam carving algorithm is performed multiple times, iteratively reducing the image's width by one pixel at a time until the target width is achieved. Existing seam carving methods for image and video retargeting differ in the technique of computing the gradient map $G(x,y)$. For example, Avidan et al. [7] use edge maps computed by cross-correlation of Sobel masks M and image I.

$$G(x,y)=\left\{\sum_{i=-w/2}^{w/2}\sum_{j=-w/2}^{w/2}M(i,j)I(x+i,y+j)\right. \quad \textbf{(4)}$$

where w is the width of the masks.

3.2 Spatial Coherence and Error Distribution

In general Sobel masks, do not consider the visual distortion caused by seam removal. Instead of using Sobel masks, we define three different 5×5 masks, namely vertical mask M_v, left diagonal mask M_l, and right diagonal mask M_r (Figure 5). Vertical seams cutting across horizontal lines produce no distortion on the quality of horizontal lines and edges after seam carving (Figure 6). So, horizontal mask, present in Sobel masks for detecting horizontal edges, is omitted and replaced by two diagonal masks. Moreover, seam carving that includes horizontal mask and equal weights of 2.5 produces significant shape distortion. On the other hand, omitting the horizontal mask and applying higher weight to the vertical mask than the diagonal masks (w_v = 6, w_l = w_r = 2) results in less shape distortion (5). So that the improved mask can reduce the distortion caused by vertical seam removal. Also, the horizontal seam removal works the same way.

The proposed gradient map G is defined by the difference of adjacent pixel in different directions (Upper right, Upper and Upper left). Higher the value of G the strength of the edge pixel is greater. Cross-correlations between the three masks and the input image we produce three corresponding oriented edge maps G_v, G_l, and G_r. These oriented edge maps are com-bined into a single map $G_c(x,y)$ after normalizing their contents to the range of [0, 1] so that their weights w_v, wl, and w_r can be more robustly set to work well for all images:

$$G_c(x,y)=w_vG_v(x,y)+w_lG_l(x,y)+w_rG_r(x,y) \quad \textbf{(5)}$$

Next, Canny edge detector is applied on $G_c(x,y)$ to obtain binaries edge map $G(x,y)$, which is used in (1) to compute the energy map for seam carving. As discussed in the previous sections, the key idea of our method is to avoid cutting the same object in consecutive iterations. To achieve this goal and minimize distortions of object ge-ometry and line

-1	-2	0	2	1
-2	-4	0	4	2
-4	-8	0	8	4
-2	-4	0	4	2
-1	-2	0	2	1

(a) Vertical mask

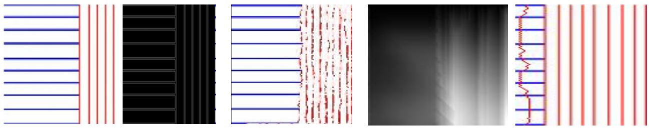
-1	-2	-2	-2	0
-2	-4	-4	0	2
-2	-4	0	4	2
-2	0	4	4	2
0	2	2	2	1

(b) Left diagonal mask

0	2	2	2	1
-2	0	4	4	2
-2	-4	0	4	2
-2	-4	-4	0	2
-1	-2	-2	-2	0

(c) Right diagonal mask

Figure 5. Edge masks for detecting oriented edges



(a) Input image (b) Edge map (c) Removal of vertical line (d) Targeted image with optimal (e)

Figure 6. Example of retargeting line patterns

structures, two key elements are added to the method. First, a feature map $F(x,y)$ is used to encode important features or objects in the input image I. The feature map F considers both gradient and saliency in image. Then a threshold TH is used to separate background (0) and foreground (main object) (1).

$$F(x,y)=w_gG(x,y)+w_cC(x,y) \quad \textbf{(6)}$$

where w_g and w_c are positive weights, and $C(x,y)$ is the saliency map as defined in [5]. The saliency map C is defined to represent the significant parts in the source image. Main purpose of the saliency map is to determine the main object's energy. Hence the highest energy area more likely represents the main subject. Saliency map is used as a tool for main object detection in our method. So we choose an efficient method [31] that can define more main objects as the tool. We purpose to preserve all the main objects (not only for several main objects) using our feature map (composed with proposed edge map and saliency map). For an example in [32], several important main objects can be defined, but some secondary important region is ignored.

Second, a *priority map* $P(x,y)$ (see Figure 3 (g))is used to encode the priority of objects to be protected from being cut by seam carving. Some objects with lower saliency in the source image are also can be defined by the new feature map F. Then a threshold TH is used to separate background (bg) and foreground (main object) (mb). As in Figure 7, the secondary main object (tree leaf) cannot be detected by only saliency map. After the k-th seam pass through the main object, the value in our Priority map is incresed to protect this main object (see Figure 8 (e)). Initially, the priority map $P_0(x,y)$ is initialized to zero. At the beginning of iteration k = 0, the priority map P is added to the edge map G to produce the *prioritized edge map* G' that indicates the regions to be protected in iteration k:

$$G'_k(x,y) = G_{k-1}(x,y) + P_{k-1}(x,y) \qquad \textbf{(7)}$$

Then, the seam S_k is computed from G'_k (instead of G_k) through energy minimization discussed in Section 3.1. Next, the feature map $F_k(x,y)$ and the priority map $P_k(x,y)$ are computed to identify the regions with significant features that are intersected by seam S_k:

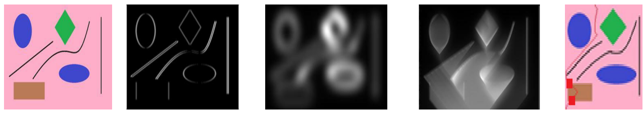
$$P_k(x,y) = \begin{cases} \{P_{k-1}(x,y) + \gamma & \text{if } (x^*,y^*) \in S_k \wedge |x - x^*| \leq \wedge \\ & |y - y^*| \leq r \wedge F_k(x^*,y^*) < \tau \\ P_{k-1}(x,y) & \text{otherwise} \end{cases} \qquad \textbf{(8)}$$

where, $\gamma > 1$ is a constant parameter, τ is a constant threshold, and (y^*, x^*) is the location of seam. After computing $P_k(x,y)$, its contents are normalized to [0, 1]. So, regions to be protected in the next iteration have priotiry values $P_k(x,y)$ close to 1 while the other regions have priority values less than 1. That is, regions with significant features that are cut in iteration k will not be cut again in the next iteration. Moreover, equation (8) with normalization rotates the priority among the regions in the image (Figure 9). In this way, the pixels will be removed evenly over the regions, thus distributing distortion errors that result from pixel removal. All seams will be avoided to define in the same region if the value of significant map is higher, that's why our method called BED (balance error distribution), and it can solve the two difficulties D3: **GD** and D4: **LD** specially. As in Figure 7 (b) [7] cannot avoid to meet the difficulty **GD** and **LD**. But, our BED method can balance the error distribution in line and contour to reduce the visual distortion.

(a) Input image (b) Seam carving with horizontal mask and equal weights causes significant shape distortion (c) Seam carving without horizontal mask and higher weight for vertical mask results in less shape distortion

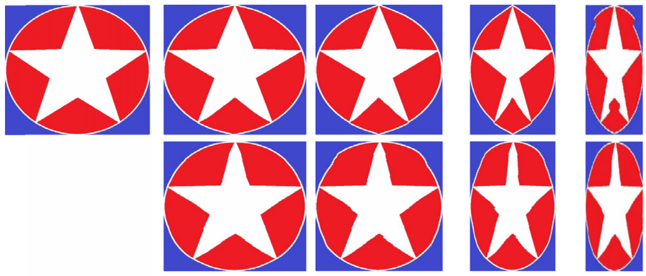
Figure 7. Shape preservation without horizontal mask

In the final step, pixels along the seam S_k are removed from the image and the maps to yield truncated image I_k and truncated maps G_k, C_k, F_k, and P_k.



(a) (b) (c) (d) (e)

Figure 8. The seam removal process of proposed method



(a) Input image (b) Seams computed in the first three iterations (c) Seams computed in the first three iterations (d) Seams computed in the first three iterations (e) Retargeted image

Figure 9. Rotation of priority of regions with important features. Top row: without priority rotation, seams in successive iterations are placed at the same region. Bottom row: with priority rotation, seams in successive iterations are placed in different regions

This step completes the operation at iteration k. The whole process is iterated until the target width of the image is achieved. The above algorithm can be summarized as follows:

(1) Initialization
  Let $I_0 \leftarrow I$ (Figure 3 (a)), and compute edge map G_0 (Eq. 5) (Figure 3 (b)), saliency map C_0 (Figure 3 (c)), feature map F_0 (Eq. 6) (Figure 3 (d)), and priority map P_0=0.
(2) For each iteration k = 1,…, K, until target width is satisfied
  (a) Determine optimal seam:
    Compute prioritized edge map $G'_k \leftarrow G_{k-1} + P_{k-1}$ (Eq. 7) and optimal seam S_k (Figure 3 (f)) from E_k (Eq. 1-3) (Figure 3 (c)).
  (b) Update priority:
    Compute priority map P_k (Eq. 8 with normalization) (Figure 3 (g))
  (c) Remove seam:
    $I_k \leftarrow I_{k-1} \setminus S_k$, $G_k \leftarrow G_{k-1} \setminus S_k$,
    $C_k \leftarrow C_{k-1} \setminus S_k$, $F_k \leftarrow F_{k-1} \setminus S_k$,
    $P_k \leftarrow P_k \setminus S_k$,
    Where \ denote set subtraction
(3) Output retargeted image I_K

Note that the image *I*, edge map *G*, saliency map *C*, and feature map *F* are computed only once at Step 1 of the algorithm. In subsequent iterations, they are simply truncated by removing pixels along the seam. In each iteration *k*, only the prioritized edge map G'_k, seam S_k, and priority map P_k need to be computed. So, our algorithm is very efficient compared to existing seam carving methods that attempt to preserve lines and shapes [18]. As an example in [18], Rubinstein et al., found the surface of seam in 3D plane with higher time complexity. Our method only considers each frame to find out the optimal seam to reduce the searching time for optimal seams.

3.2 Temporal Coherence

Our video retargeting method considers highest priority in temporal coherence part. In our method, temporal coherence is enforced by constraining the seams in successive frames to lie near each other (see in Figure 4 (a)). To achieve this goal, a *temporal priority map* $T_k^t(x, y)$ is defined for iteration *k* at frame *t* > 1:

$$T_k^t(x, y) = \begin{cases} 0 & \text{if } |x - x^*| < D \wedge s_k^{t-1} = (x^*, y), \\ \rho & \text{otherwise.} \end{cases} \quad \textbf{(9)}$$

where, *D* is a constant distance and $\rho > 1$ is a constant parameter. For frame 1, $T_k^1(x, y) = 0$ for all *k*. This temporal priority is added to the energy map *E*, modifying Eq. (1) into the following:

$$E_k^t(x, y) = \begin{cases} G'^t_k(x,y) + T_k^t(x,y) & \text{for } y = 1, \\ G'^t_k(x,y) + T_k^t(x,y) + \min_{i=-1}^{1} E_k^t(x+i, y-1) & \text{for } y > 1 \end{cases} \quad \textbf{(10)}$$

Note that in the above equation, the edge map *G* is replaced by the prioritized edge map *G′* as discussed in Section 3.2. During video retargeting, in each iteration *k*, the video frames are retargeted starting from frame 1 in the order of the frame sequence. For each frame *t*, the seam carving algorithm described in Section 3.2 is executed. The temporal priority T_k^t is computed based on the seam S_k^{t-1} obtained in the previous frame *t*-1. Next, the modified energy map E_k^t is computed according to Eq. 10 (instead of Eq. 1), which is then used to determine the optimal seam S_k^t for seam carving. Then, the algorithm is repeated for all subsequent video frames, until the satisfied width is achieved. Our approach addresses the difficulty (D5: **TC**) since the modified energy map considers spatial domain even in temporal coherence.

In our method, seam carving is used to remove relatively unimportant pixels along the seams in the image and video. We use the Feature map and Priority map to prevent seam removal in the same object to maintain the original shape of the object (addressing the difficulties **D1-D4**). Hence, our method overcomes both the geometry and linear structures distortion. Finally, temporal coherence (address the 5th difficult **TC**) is achieved by constraining the seams in consecutive video frames to lie near each other. These properties will be demonstrated and discussed in Section IV.

4 Experimental Results

In this section, our experimental results of the proposed method are discussed. Experimental results demonstrate that our algorithm can address the 5 difficulties D1-D5 and can be applied to images or videos with or without complex image content. In addition, the results are compared with several other retargeting methods to demonstrate that the proposed algorithm is more practical for different kinds of images and videos.

4.1 Body Maintenance

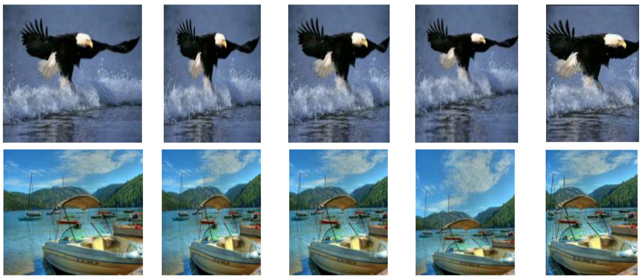
The image retargeting methods discussed above are compared in Figure 10. The source image is retargeted to 50% of width. Retargeting operators used are: simple scaling operator (SCL), manually chosen cropping windows (CR), Nonhomogeneous warping (WARP) [23], Energy-based deformation (LG) [39], Shift-maps (SM) [38], Scale-and-Stretch (SNS) [8], Streaming Video (SV) [5], Seam-Carving (SC) [18], Multi-operator (MULTIOP) [19], and. the proposed method (PROP). In this case we can see that (i) [18] cannot maintain all important object, another tree is disappear. (j) [19] retarget the image using multiple retargeted operators liked bi-cubic scaling (SCL), cropping (CR) and seam carving (SC) to perform better results than other seam carving, cropping, scaling, warping based methods, but our main object size can be maintain better. Obviously, our method performs well than other retargeting methods (b-j), because both two difficulties **NO** and **SO** are addressed after using our **BED** protecting mechanism (Priority Map *P*).

See in Figure 11, *eagle* and *boat* are both important main object. In image retargeting phase, we have to consider more on maintain the size of most important object *eagle* and *boat*. Lin *et. al* [29]'s retargeting results almost the same with only using uniform scaling (b). Rubinstein et al.'s method [19] (c) performs well in **SO** in this case than Lin et al.'s method [19] (d), but cannot address the difficulty **GD**, since the boat with heavily deformation in shape. Our **BED** result (e) performs well than other result, because we maintain the original size of the main object *eagle* and *boat*.

(a) ORIGINAL (b) SCL (c) CR (d) WARP [23] (e) LG [39]

(f) SM [38] (g) SNS [8] (h) SV [40] (i) SC [18] (j) MULTI [19] (k) PROP

Figure 10. Comparison of body maintenance

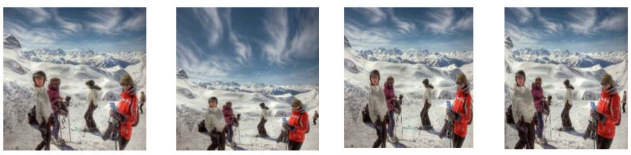


(a) Source (b) SCL (c) [19] (d) [29] (e) **BED**

Figure 11. Comparison of body size

In Figure 12 (b-f) results well maintained all important regions like all people in foreground and all mountains in background. But the 4 people became smaller in size than the original image. Rubinstein et al.'s results perform worse than other retargeting techniques (b-f) in two difficulties: **SO** and **NO**. The 4 people in foreground became smaller in size and the fifth person in background is disappeared. Our **BED** preserves important regions (people and mountain) in size and counts.

(a) ORIGINAL (b)SCL (c) WARP (d) LG



(e) SC (f) SV (g) MULTI [19] [h] BED

Figure 12. Comparison of number of body

4.2 Maintenance of Linear/Geometry Structures

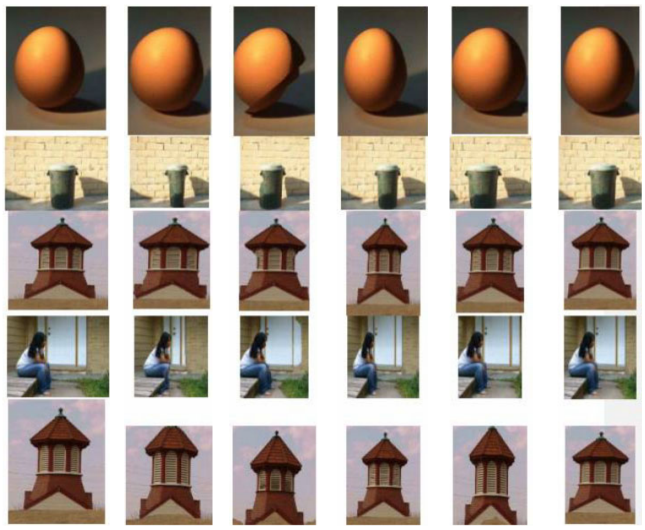
The line structure is maintained using seam carving in Johannes et al.'s [26] method as shown in Figure 13 (c). Our **BED** protecting mechanism performs better than Johannes et al.'s method as can be seen in (d). The second row shows limitation cases of Johannes et al.'s [26] method. In images with too complex line structures like in the image (e), all the lines cannot be protected. Our method performs better (h) than Johannes et al.'s method (g). It means that our **BED** can protect the lines well in **LD** than Johannes et al.'s method.

(a) source image (b) uniform scaling (c) K. Johannes et al.'s method [26] (d) proposed method

(e) source image (f) uniform scaling (g) K. Johannes et al.'s method [26] (h) proposed method

Figure 13. Comparison of the image results. Here all images are resized to 60 % of the original width

More results compared in **GD** aspect are introduced here. Figure 14 (b), (c) and (e) are the results of the seam carving based methods and (d) is the result of scaling based method. Seam carving based method can maintain the original size of the main object (egg). Scaling based method make the main object smaller in size as some of the original pixels are merged. Obviously, this example shows that Battiato et al.'s method [30] perform better than [7, 33-35], but the edge of the egg is distorted (left side). Our approach can preserve the contour of the shape even the targeted width is smaller than the width of egg. The bucket in the second raw demonstrates that our method can maintain original geometry structure than others. Also, our results demonstrate that our **BED** approach maintains the Geometry Distortion **GD** in vertical retargeting (see the 3rd -4th row).



(a) Source (b) [34] (c) [34] (d) 35 (e) [30] (f) **BED**

Figure 14. The Geometry comparison.

4.3 Video Test Results

All video results are compared with the results of other important retargeting techniques namely, seam carving [18], discontinuous seam carving [21], and other recent warping-based retargeting methods [23-25, 28], to demonstrate that our **BED** also can be applied to video to overcome the **TC** difficulty. In order to make the comparison convincing, all retargeted results are resized around 50% of the original width. All selected videos were appeared in the recent well-known papers in the literature. As our seam carving based method use strict temporal coherence (see in Figure 4 (a)), the resultant video will not cause shaking affect. The seam is possible to pass through the main object **OB**, but the seam will not pass through the same region of the object too many times, as our Priority Map records these regions as important region, to maintain the main object size after seam carving. It is due to all compared retargeting results are addressed temporal coherence, we still compared our results with other works in 4 difficulties: **SO**, **NO**, **GD** and **LD**.

Selected video results are used to demonstrate that the proposed method can preserve the main object size well in our strict temporal coherence.

Figure 15 shows the comparison of our retargeting results with uniform scaling and Grundmann et al.'s [21] method. Grundmann et al.'s result can preserve most of the important regions while maintaining the temporal coherence. Comparing to uniform scaling and Grundmann et al.'s method [21], we can see that our method can maintain the object in its original size.

(a) the original	(b) are video retargeting results using uniform scaling	(c) Grundmann et al.'s method [21]	(d) the proposed method, respectably

Figure 15. Comparison of the video retargeting results for a selected frame

Rubinstein et al.'s method [18] generates good results for some videos. As shown in Figure 16, our **BED** method shows that we can use a lower time complexity to achieve the acceptable result than using graph cut in 3D plane. Also, we can see that the line stricture can be maintained slightly well with difficulties **SO** and **LD** than Rubinstein et al.'s [18], Wang et al.'s [35], Lin et al.'s [28] methods. And, the difficulty of **SO** is addressed compared to the result of uniform scaling.

Figure 16. Comparison of the video retargeting results with uniform scaling, Rubinstein et al.'s [18], Wang et al.'s [35], Lin et al.'s [28] and our BED method. The first row is for frame 46 and second row is for frame 72 respectively

Figure 17 (c) is Grundmann et al.'s [21] retargeting result. They use a user interface to mark important region, since the girl is dressed a similar color to the background which is difficult to track as main object. Our method performs well without enough saliency information. Because this image will some strong line to make our feature map more strength, then some important region could be detected and protect well, so the difficulty **GD** also can be addressed than Grundmann et al.'s method in this case.

(a) is the input video with user marked important area

(b) are video retargeting results using uniform scaling

(c) Grundmann et al.'s method [21]

(d) with pre-marked important region and the proposed method with no pre-marked region

Figure 17. Comparison of the video retargeting results for a selected frame

In Figure 18, we compare our results with selected classic retargeting methods (*i.e., Rubinstein* et al.'s [*18*]). The strategy of seam surface caused lots of seams pass through the main object. In the results in the first and the second rows, we can notice serious damages to major objects (*i.e., man and woman*). Wolf et al.'s method [23] also cannot preserve some main objects. Wang et al. [24] and Yen et al. [25] can preserve structure of the main object but cannot preserve the original size of the main object, which makes the results almost look similar to uniform scaling. The 3rd raw show that the proposed approach can avoid the line structure be damaged (white line), but also maintain the main object size (girl).

Figure 18. Comparison of the video results from left to right with uniform scaling, Rubinstein et al. [18], Wolf et al. [23], Wang et al. [24], Yen et al. [25], and the proposed method

Wang et al.'s method [27] proposed a motion-based video retargeting method using warping operator, further improving previous work Wang *et al's* [24] method. As shown in Figure 19, compared to cropping, [27]'s method can preserve more important objects (tree in background). But in this case we can see that the main object (bear) is heavily deformed (thinner). Our approach better maintain the main object shape (**SO** and **GD** are addressed).

(a) Source (b) Cropping (c) [27] (d) **BED**

Figure 19. Comparrison of main object size **SO** and Geometry distortion **GD** between cropping and Wang *et. al's* method [27]

Here, we show two videos with complex linear and geometry structure. See in Figure 20, we used red circle and red rectangle to highlight the difficulty of **LD** and **GD**. The first row shows that our method can maintain the line distortion well than Wang et al.'s [24] and Wang et al.'s [27] method. Also, we can maintain the geometry distortion (see second row in (c)).

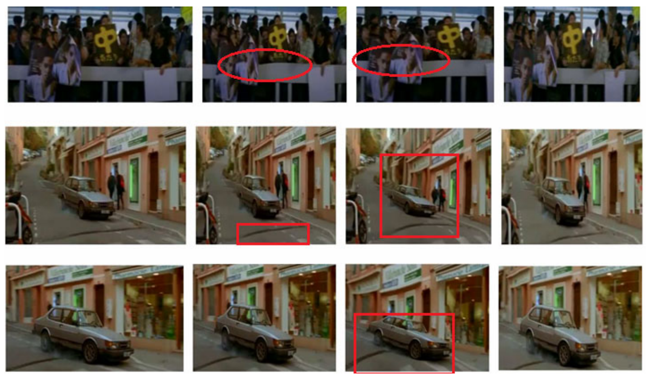


Figure 20. Comparrison the geometric distortion between [24, 27] and our **BED** method. Rectangles represent significant distortion

Besides, two warping based techniques [28, 35] are compared in Figure 21. In this case we can see that, if we use warping to change object large, it will easily cause the geometry distortion.

Figure 21. Comparrison between [28, 35] and our BED method

Finally, we demonstrate that our method is not only capable of preserving structure of lines and shapes, but also retains the size of objects even for the cases which are difficult to retarget using seam carving. In Figure

22, we compare the video results using Rubinstein et al.'s [18] forward energy method and the proposed method to show the problem. In these 6 cases of videos, all videos contain complex image content in each frame. It is the limiting case for almost all the seam carving methods since these cases cannot maintain the spatial and temporal coherence at the same time. If we consider spatial domain in each image, all major objects can be retained but the video result will loose temporal coherence; and hence, the retargeted result will look shaky in visual. In other words, if we focus on maintaining the temporal coherence, the optimal seam may go through the moving object.

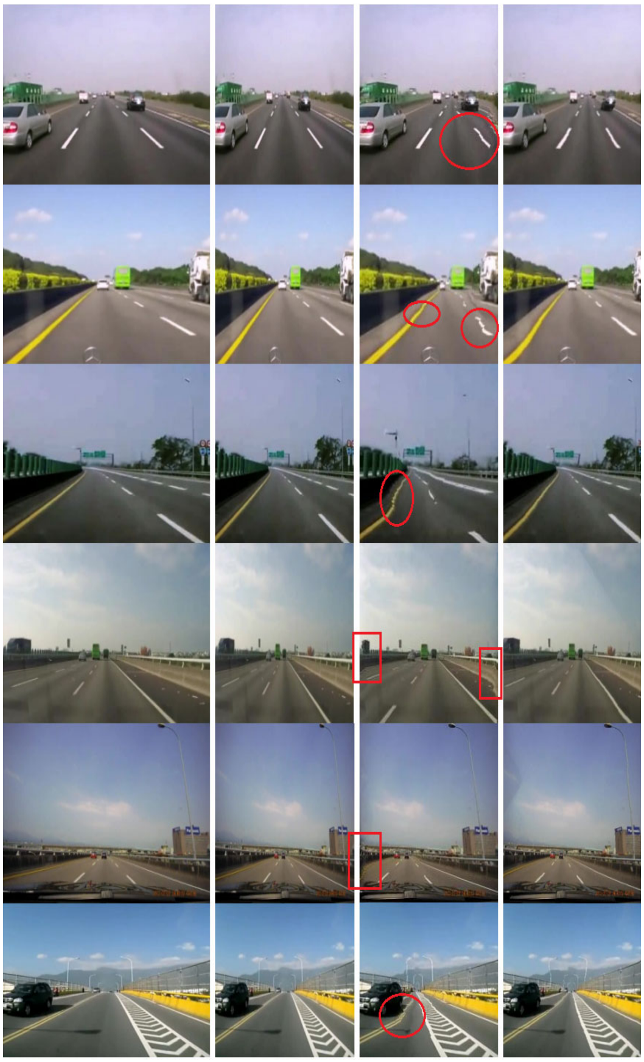
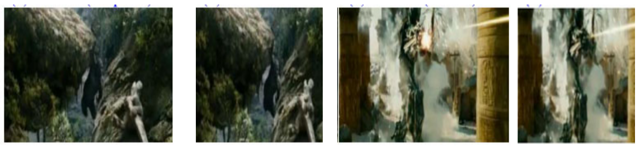


Figure 22. Comparison of the video results. The input videos are shown in first column, the second column to the fourth column are the video retargeting results with uniform scaling, Rubinstein et al.'s method [18] and the proposed method

As shown in Figure 23, the BED method could be applied for low motion videos (a-d) or high motion videos (e-g). From this example we can see that the man, woman, and orangutan are almost maintained with original size. It means that only the important regions are disappeared by the defined multiple seams. In other words, it means our BED method performs normal even with different motion video.

(a) Source (Airplain) (b) BED (c) Source (Dance) (d) BED



(e)Source (Orangutan) (f) BED (g) Source (Robot) (g) BED

Figure 23. The video results with large motion videos using our BED method

5 Conclusion

This paper proposes an efficient retargeting method that the online video could be watched or transferred each other using mobile phone, which can be applied to image and video with complex structures of lines and shapes. The contributions are in the achievement to keep three important factors in video seam carving: to preserve the size and structure of multiple main objects, to maintain the continuation of straight lines, and to allow temporal coherence. The proposed method is developed as a prototype tool to be used commercially and now we work on improving the performance and to implement the proposed method to resize and retarget videos for smart phone applications.

References

[1] A. Santella, M. Agrawala, D. DeCarlo, D. Salesin, M. Cohen, Gaze-based Nteraction for Semi-automatic Photo Cropping, *Proc. SIGCHI Conference on Human Factors in Computing Systems*, Montreal, Canda, 2006, pp. 771-780.

[2] B. Suh, H. Ling, B. Bederson, D. Jacobs, Automatic Thumbnail Cropping and Its Effectiveness, *Proc. 16th Annual ACM Symposium on User Interface Software and Technology*, Vancouver, Canada, 2003, pp. 95-104.

[3] H. Liu, X. Xie, W.-Y. Ma, H.-J. Zhang, Automatic Browsing of Large Pictures on Mobile Devices, *ACM Internal Conference on Multimedia*, Berkeley, CA, 2003, pp. 148-155.

[4] S. Kopf, B. Guthier, H. Lemelson, W. Effelsberg, Adaptation of Web Pages and Images for Mobile Applications, *Proc. of IS&SPIE Conference on Multimedia on Mobile Devices*, San

Jose, CA, 2009, pp. 1-12.

[5] L. Itti, C. Koch, N. Ernstr, A Model of Saliency Based Visual Attention for Rapid Scene Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No.11, pp. 1254-1259, November, 1998.

[6] F. Liu, M. Gleicher, Automatic Image Retargeting with Fisheye-view Warping, *Proc. ACM Symposium on User Interface Software and Technology*, Seattle, WA, 2005, pp. 153-162.

[7] S. Avidan, A. Shamir, Seam Carving for Content Aware Image Resizing, *ACM Transactions on Graphics*, Vol. 26, No. 3, Article No. 10, July, 2007.

[8] Y. S. Wang, C. L. Tai, O. Sorkine, T. Y. Lee, Optimized Scale-and-stretch for Image Resizing, *ACM Transactions. on Graphics (TOG)-Proc. of ACM SIGGRAP*, Vol. 27, No. 5, Article No. 118, December, 2008.

[9] S. Battiato, G. M. Farinella, N. Grippaldi, G. Puglisi, Content-based Image Resizing on Mobile Devices, *International Conference on Computer Vision Theory and Applications (VISAPP)*, Rome, Italy, 2012, pp. 87-90.

[10] C. Tao, J. Jia, H. Sun, Active Window Oriented Dynamic Video Retargeting, *ICCV Proc. Workshop on Dynamic Vision*, Rio de Janeiro, Brazil, 2007, pp. 1-12.

[11] S. Cho, H. Choi, Y. Matsushita, S. Lee, Image Retargeting Using Importance Diffusion, *Proceeding IEEE International Conference on Image Processing*, Cairo, Egypt, 2009, pp. 977-980.

[12] L. Chen, X. Xie, X. Fan, W. Ma, H. Zhang, H. Zhou, A Visual Attention Model for Adapting Images on Small Displays, *Multimedia Systems*, Vol. 9, No. 4, pp. 353-364, October, 2003.

[13] X. Fan, X. Xie, H. Zhou, W. Ma, Looking into Video Frames on Small Displays, *Proc. Eleventh ACM International Conference on Multimedia*, Berkeley, CA, 2003, pp. 247-250.

[14] J. Xiao, X. Zou, Z. Liu, X. Guo, A Novel Adaptive Interpolation Algorithm for Image Resizing, *International Journal Innovative Computing, Information and Control*, Vol. 3, No. 6(A), pp. 1335-1345, December, 2007.

[15] Y. Zhang, S. Gao, C. Zhang, J. Chi, Application of a Bivariate Rational Interpolation in Image Zooming, *International Journal of Innovative Computing, Information and Control*, Vol. 5, No. 11(B), pp. 4299-4307, November, 2009.

[16] P. Lin, H. Chu, T. Lee, Smooth Shape Interpolation for 2D Polygons, *International Journal Innovative Computing Information and Control*, Vol. 4, No. 9, pp. 2405-2417, September, 2008.

[17] Y. Tian, C. Zhang, F. Peng, S. Zheng, An Iterative Hybrid Method for Image Interpolation, *Proc. International Conference on Intelligent Computing (ICIC 2005)*, Vol. 1, Heifei, China, 2005, pp. 10-19.

[18] M. Rubinstein, A. Shamir, S. Avidan, Improved Seam Carving for Video Retargeting, *ACM Transactions on Graphics (TOG)- Proc. ACM SIGGRAPH 2008*, Vol. 27, No. 3, Article No. 16, August, 2008.

[19] M. Rubinstein, A. Shamir, S. Avidan, Multioperator Media Retargeting, *ACM Transactions on Graphics*, Vol. 28, No. 3, Article No. 23, August, 2009.

[20] F. Liu, M. Gleicher, Video Retargeting: Automating Pan and Scan, *Proc. ACM International Conference on Multimedia*, Santa Barbara, 2006, pp. 241-250.

[21] M. Grundmann, V. Kwatra, M. Han, I. Essa, Discontinuous Seam-carving for Video Retargeting, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, CA, 2010, pp. 596-576.

[22] B. Yan, K. Sun, L. Liu, Matching Area Based Seam Carving for Video Retargeting, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 23, No. 2, pp. 302-310, February, 2013.

[23] L. Wolf, M. Guttmann, D. Cohen-Or, Non-homogeneous Content-driven Video-retargeting, *Proc. IEEE International Conference on Computer Vision*, Rio de Janeiro, Brazil, 2007, pp. 1-6.

[24] Y. S. Wang, H. Fu, O. Sorkine, T. Y. Lee, H. P. Seidel, Motion-aware Temporal Coherence for Video Resizing, *ACM Transactions on Graphics*, Vol. 28, No. 5, Article No. 127, December, 2009.

[25] T. C. Yen, C. M. Tsai, C. W. Lin, Maintaining Temporal Coherence in Video Retargeting Using Mosaic-Guided Scaling, *IEEE Transactions on Image Processing*, Vol. 20, No. 8, pp. 2339-2351, August, 2011.

[26] J. Kiess, S. Kopf, B. Guthier, W. Effelsberg, Seam Carving with Improved Edge Preservation, *SPIE Electronic Imaging*, Vol. 7542, San Jose, CA, 2010, pp. 1-11.

[27] Y.-S. Wang, H.-C. Lin, O. Sorkine, T.-Y. Lee, Motion-based Video Retargeting with Optimized Crop-and-warp, *ACM Transactions on Graphics (TOG)- Proceedings of ACM SIGGRAPH 2010 TOG Homepage*, Vol. 29, No. 4, Article No. 90, July, 2010.

[28] S. S. Lin, C. H. Lin, I. C. Yeh, S. H. Chang, C. K. Yeh, T. Y. Lee, Content-Aware Video Retargeting Using Object-Preserving Warping, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 10. pp. 1677-1686, October, 2013.

[29] S. S. Lin, I. C. Yeh, C. H. Lin, T. Y. Lee, Patch-based Image Warping for Content-Aware Retargeting, *IEEE Transactions on Multimedia*, Vol. 15, No. 2, pp. 359-368, February, 2013.

[30] S. Battiato, G. M. Farinella, G. Puglisi, D. Ravì, Saliency-Based Selection of Gradient Vector Flow Paths for Content Aware Image Resizing, *IEEE Transactions on Image Processing*, Vol. 23, No. 5, pp. 2081-2095, May, 2014.

[31] X. Hou, J. Harel, C. Koch, Image Signature: Highlighting Sparse Salient Regions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 1, pp. 194-201, January, 2012.

[32] A. Borji, M. M. Cheng, H. Jiang, J. Li, Salient Object Detection: A Benchmark, *IEEE Transactions on Image Processing*, Vol. 24. No. 12, pp. 5706-5722, December, 2015.

[33] Y. Ding, J. Xiao, J. Yu, Importance Filtering for Image Retargeting, *IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, Colorado Springs, CO, 2011, pp. 89-96.

[34] R. Achanta, S. Süsstrunk, Saliency Detection for Content-aware Image Resizing, *16th IEEE International Conference on Image Processing (ICIP)*, Cairo, Epypt, 2009, pp. 1005-1008.

[35] Y. S. Wang, J. H. Hsiao, O. Sorkine, T. Y. Lee, Scalable and Coherent Video Resizing with Per-Frame Optimization, *ACM Transactions on Graphics (TOG)- Proceedings of ACM SIGGRAPH 2011*, Vol. 30, No. 4, Article No. 88, July, 2011.

[36] G. Roberto, E. Ardizzone, R. Pirrone, Real-time Content-Aware Image Resizing Using Reduced Linear Model, *Proc. IEEE International Conference on Image Processing*, Hong Kong, China, 2010, pp. 2813-2816.

[37] Z. Karni, D. Freedman, C. Gotsman, Energy-based Image Deformation, *Computer Graphics Forum*, Vol. 28, No. 5, pp. 1257-1268, July, 2009.

[38] P. Krähenbühl, M. Lang, A. Hornung, M. Gross, A System for Retargeting of Streaming Video, *ACM Transactions on Graphics (TOG)*, Vol. 28, No. 5, Article No. 126, December, 2009.

[39] B. Li, L. Duan, J. Wang, R. Ji, C.-W. Lin, W. Gao, Spatiotemporal Grid Flows for Video Retargeting, *IEEE Transactions on Image Processing*, Vol. 23, No. 4, pp. 1615-1628, April, 2014.

[40] Y. Fang, Z. Chen, W. Lin, C.-W. Lin, Saliency Detection in the Compressed Domain for Adaptive Image Retargeting, *IEEE Transactions on Image Processing*, Vol. 21, No. 9, pp. 3888-3901, September, 2012.

Biographies

Hon-Hang Chang is Ph.D. student at the Department of Computer Science and Information Engineering, National Central University (NCU), Taiwan (R.O.C.). He acquired his Master's degree in Department of Photonics and Communication Engineering of Asia University of Taiwan in 2011. His research fields are image processing, information hiding and water marking.

Shwu-Huey Yen is Associate Professor at the Department of Computer Science and Information Engineering, Tamkang University (TKU), Taiwan (R.O.C.). His research interests are in the area of image/video processing and automata theory.

Timothy K. Shih is Professor at Department of Computer Science and Information Engineering, National Central University, Taiwan. Prof. Shih is a Fellow of the Institution of Engineering and Technology (IET). In addition, he is a senior member of ACM and a senior member of IEEE. Prof. Shih also joined the Educational Activities Board of the Computer Society. He acquired his MS in Computer Engineering from California State University, Chico, USA in 1985 and Ph.D. in Computer Engineering from Santa Clara University, USA in 1993. His current research interests are mainly in the area of Computer Vision, Interactive Multimedia, Multimedia Processing, Human Computer Interaction and e-learning.

Wee-Kheng Leow is Associate Professor at Department of Computer Science, National University of Singapore, Singapore. His research interests are in the area of Computer Vision, Interactive Multimedia and multimedia processing.

Shih-Jung Wu is Associate Professor at the Department of Innovative Information and Technology, Tamkang University (TKU), Taiwan (R.O.C.). His research interests are in the area of Wi-Fi, parallel algorithm, graph theory and data mining.