Patch-Based Image Warping for Content-Aware Retargeting

Shih-Syun Lin, I-Cheng Yeh, Chao-Hung Lin, *Member, IEEE*, and Tong-Yee Lee, *Senior Member, IEEE*

***Abstract*—Image retargeting is the process of adapting images to fit displays with various aspect ratios and sizes. Most studies on image retargeting focus on shape preservation, but they do not fully consider the preservation of structure lines, which are sensitive to human visual system. In this paper, a patch-based retargeting scheme with an extended significance measurement is introduced to preserve shapes of both visually salient objects and structure lines while minimizing visual distortions. In the proposed scheme, a similarity transformation constraint is used to force visually salient content to undergo as-rigid-as-possible deformation, while an optimization process is performed to smoothly propagate distortions. These processes enable our approach to yield pleasing content-aware warping and retargeting. Experimental results and a user study show that our results are better than those generated by state-of-the-art approaches.**

***Index Terms*—Image retargeting, image warping, optimization.**

I. Introduction

With the development of mobile devices, image retargeting has become an active research topic in the fields of image processing and computer graphics. The naive approaches, that is, linear scaling and uniform cropping, have been proven inappropriate when the aspect ratio of an image is changed significantly. Thus, considerable research efforts have been devoted to the technique of content-aware retargeting [1]–[14]. The key to this technique is resizing images to arbitrary aspect ratio while keeping visually salient objects in similar aspect ratio. Thus, minimizing visual distortion is the requirement of content-aware retargeting.

Seam carving [3] and image warping [7] are recent techniques of content-aware retargeting. Seam carving iteratively removes or inserts a seam passing through unimportant regions. This approach may generate jagged edges because of the removal of discontinuous seams. In contrast, image warping offers a better possibility of producing a continuous deformation for content-aware retargeting. Wang *et al.* [7] propose an optimized scale-and-stretch warping using a quad mesh as a control mesh. This approach has the advantage of distributing distortions to homogeneous regions, as it forces quads with significant content to scale uniformly and distorts quads with homogeneous content. This approach can preserve the aspect ratios of local objects. However, for an object occupying many quads, an inconsistent deformation may occur because of the inconsistent scaling factors of quads [8]. This inconsistency may result in an unnatural deformation, causing significant visual distortion even though a well-defined significance map is used (Fig. 1(c)–(h)). To solve this problem, a patch-based image warping is proposed. In preprocessing, the input image is segmented into several patches and each patch is assigned a significance value. Then, the patches with high significance are forced to undergo as-rigid-as-possible deformation by similarity transformation constraints while smoothly propagating distortions through an optimization process. With the aid of patch-based significance map and warping technique, the proposed approach can consistently preserve visually salient content (Fig. 1(b)).

Compared with recent studies on image retargeting, the proposed approach offers the following contribution. A patch-based image retargeting approach is proposed to ease unpleasant visual distortions caused by inconsistent warping. Thus, our approach can yield better results in terms of shape and structure line preservation compared with related approaches. The remainder of this paper is organized as follows. Section II reviews the related work. Section III presents the proposed approaches. Section IV discusses the experimental results, and Section V presents the conclusions, limitation, and future work.

II. Related Work

Many content-aware image retargeting techniques have recently been proposed. Following the categorization suggested by Shamir and Sorkine [16], the approaches are classified into two categories: *discrete* and *continuous*. In discrete approaches, an input image is resized by cropping [17]–[21] or seam carving [3], [5], [6], [11], [14]. In cropping, an optimal rectangle is selected on image, and the regions outside this rectangle are cut from the image. In seam carving, a seam is iteratively removed to preserve visually salient content. A seam in an image refers to a continuous path from top to bottom or from left to right with minimal significance. Recently, Rubinstein *et al.* [22] present a multi-operator algorithm that combines cropping, linear scaling, and seam carving to optimally resize images. Similarly, Pritch *et al.* [23] suggest an approach that discretely

Manuscript received November 29, 2011; revised March 04, 2012 and June 05, 2012; accepted June 20, 2012. Date of publication November 20, 2012; date of current version January 15, 2013. This work was supported in part by the National Science Council (contracts NSC-99-2221-E-006-066-MY3, NSC-100-2628-E-006-031-MY3, NSC-100-2221-E-006-188-MY3, and NSC-100-2119-M-006-025), Taiwan. The associate editor coordinating the review of this manuscript and approving it for publication was Chong-Wah Ngo.

S.-S. Lin, I.-C. Yeh, and T.-Y. Lee are with the Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 701, Taiwan (e-mail: catchylss@hotmail.com; ichenyeh@gmail.com; tonylee@mail.ncku.edu.tw).

C.-H. Lin is with the Department of Geomatics, National Cheng Kung University, Tainan 701, Taiwan (e-mail: linhung@mail.ncku.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2012.2228475

removes repeated patterns in homogenous regions instead of scaling and stretching images. Although these approaches can generate pleasing results for many cases, it should be noted that cropping is inappropriate for the case that visually salient objects are near the borders of images, and the seam removal may result in discontinuous artifacts.

In contrast to discretely removing pixels in homogenous regions, the continuous approaches optimize a mapping or warping using several deformation and smoothness constraints [1], [2], [4], [7]–[10]. In the work of Liu *et al.* [1], a non-linear warping is proposed to preserve important content. However, image features outside the region of interest may suffer from significant distortions. Gal *et al.* [2] suggest a texturing approach based on Laplacian editing. This approach sets similarity transformation constraints on local objects in warping, leading to content-aware resizing. Wolf *et al.* [4] retarget an image by merging less important pixels to reduce distortion. However, the distortion can only be propagated along the resizing direction. To refine the distortion propagation, Wang *et al.* [7] propose an optimized scale-and-stretch approach, which iteratively warps local regions to match the optimal scaling factors as close as possible. This approach can distribute distortions in all directions, and each local region, i.e., each quad, has an acceptable homogeneous scaling. However, due to the different scaling factors of quads, an object occupying several quads may suffer from inconsistent scaling and deformation. To ease this problem, Zhang *et al.* [8], Guo *et al.* [9], Jin *et al.* [10], and Hung *et al.* [24] propose to force visually salient objects to undergo similarity transformations during resizing. These approaches perform very well on shape preservation of local objects. However, inconsistent deformations may still occur on structure lines, as they rely on accurate handle setting [8] or rigidity map generation [24]. In contrast, the proposed approach can preserve both the visually salient objects and structure lines while smoothly propagating distortions, resulting in pleasing retargeting results.

The proposed approach is inspired by the work [7], and the goal of shape preservation is similar to [8]–[10]. However, there are substantial differences between our approach and these approaches. First, a patch-based retargeting scheme with a patch-based significance map is proposed to ease inconsistent deformations. Second, a similarity transformation constraint is introduced to preserve visually salient content. Third, instead of using a triangle mesh as a control mesh in warping, which may have the problem of inconsistency in triangle orientations [9], [10], following the approach [7], we use a quad mesh with grid orientation constraints to avoid skew artifacts.

III. Content-Aware Image Resizing

A. System Overview

Fig. 2 schematically illustrates the proposed method that consists of two main steps: *preprocessing* and *image warping*. In the preprocessing step, the input image is segmented into several homogenous patches using the graph-based segmentation approach [25]. A significance measurement for the segmented patches is then performed to generate a significance map. In the generation of significance map, the context-aware saliency estimation [15] is adopted to generate a saliency value for each pixel. Afterward, each segmented patch is assigned an averaged saliency value to ease the problem of inconsistent deformation. In the warping step, a quad mesh is created to cover the image, and the proposed patch-based warping is performed to force quads within a high-significance patch to undergo as-rigid-as-possible transformation.

B. Object-Based Significance Map Generation

Many pixel significance measurements have been proposed in image retargeting approaches. Pixels with large gradient magnitudes are generally considered significant pixels [4], [5], and thus, gradient is combined with pixel saliency to determine pixel significance [7]. However, inconsistent deformation may arise when a retargeting is performed using such pixel-based measurement. This inconsistency leads to an unnatural deformation, especially in visually salient content. For instance, the retargeting results shown in Fig. 1(d) and (f) are generated using the approach [7] with the pixel-based significance measurements presented by Wang *et al.* [7] and Goferman *et al.* [15], respectively. The results show that the barrel has unnatural deformation caused by using inconsistent significance values in warping. Artifacts occurring in visually salient objects are sensitive to human visual system. Thus, this unnatural deformation will lead to considerable visual distortion.

With the aim of preserving visually salient objects and structure lines, we adopt a patch-based significance measurement instead of a pixel-based one. In our approach, the input image is initially segmented into several homogenous patches, after which each patch is assigned a significance value. In this manner, a patch containing an object or structure lines can be deformed consistently. In the preprocessing, the segmentation approach [25] is performed to partition images. In general, many patches are generated, and over-segmentation occurs as shown in Fig. 3. To address this problem, a commonly used merge process is performed. Each patch is assigned an average color to roughly represent this patch, and the neighboring patches are merged according to the following criteria. First, small patches are merged with their neighboring patches. If the area of patch, i.e., the number of pixels in the patch, is smaller than a defined threshold (set to 0.01% of the image area), the patch is merged to its neighboring patch that has similar representative color. Second, if the adjacent patches have similar representative colors, these patches are merged together. In the implementation, the color similarity is defined as the Euclidean distance in RGB color space, and the merge threshold is set to 20.

The saliency detection proposed by Goferman *et al.* [15] is performed after the segmentation. The following is a brief introduction to this approach. In [15], the saliency detection follows four basic principles: 1) low-level considerations including color and contrast; 2) global considerations including suppressing frequently occurring features and maintaining features that deviate from the norm; 3) visual organization rules stating that visual forms may possess one or several centers of gravity; and 4) high-level consideration that takes human faces into account. In accordance with principle (1), a

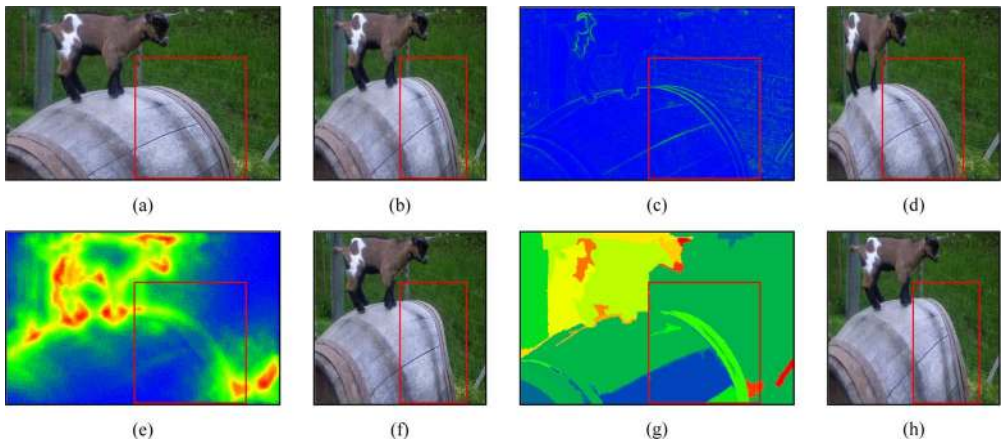

Fig. 1. Comparisons of image retargeting using various significance maps. (a) Original image; (b) our result; (c), (e), (g) significance maps generated by the approaches [7], [15], and our approach, respectively. Pixel significance is visualized by colors ranging from red (highest significance) to blue (lowest significance); (d), (f), (h) the corresponding retargeting results generated by the warping approach [7].

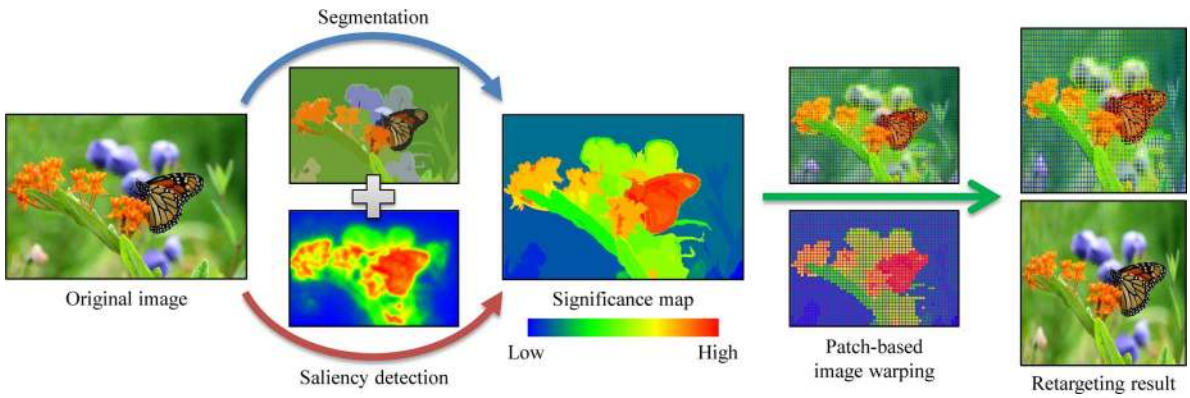

Fig. 2. Overview of the proposed approach. Left: original image; middle: significance map generated by combining segmented objects and detected saliency values; right: our retargeting result. In the result, the quads within a high-significance patch, e.g., the patches of butterfly and flower, are consistently preserved.

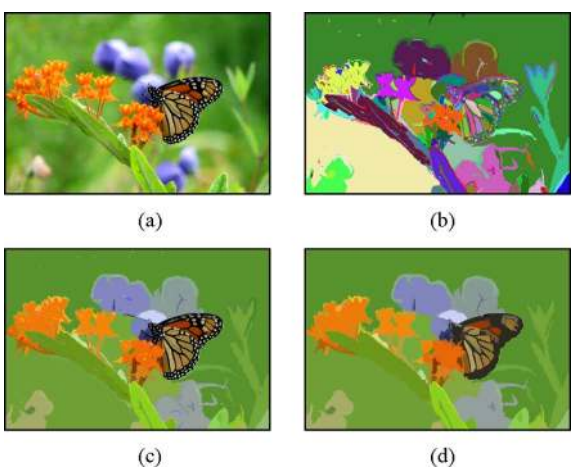

Fig. 3. Significance map generation. Top: original image (left) and segmentation result (right) (the patches are represented by colors); bottom: segmented patches displayed by their representative colors (left) and merge result (right).

high saliency value is assigned to nonhomogeneous regions, whereas a low saliency value is assigned to homogeneous regions. To meet principle (2), frequently occurring features are suppressed, and according to principle (3), the salient pixels are grouped together. Finally, principle (4) is considered in post-processing. In our approach, each segmented patch is assigned a significance value by averaging the saliency values of pixels within this patch. With the aid of this significance map, our approach can avoid inconsistent deformation, and thus can generate satisfactory retargeting results as shown in Fig. 1(b). Note that generating a perfect segmentation is sometimes difficult even when a state-of-the-art approach is adopted. Fortunately, by utilizing the distortion propagation and patch-based significance map, our approach can address the problems caused by imperfect segmentation. For instance, in Fig. 1(g), the sheep in the image is partitioned into several patches with imperfect boundaries. In this case, the effect of inconsistent deformation is reduced as shown in Fig. 1(b), and thus, the result is better than those generated by using pixel-based significance measurement (Fig. 1(d) and (f)).

C. Patch-Based Image Warping

Following the warping approach presented by Wang *et al.* [7], a quad mesh $\mathbf{M} = (\mathbf{V}, \mathbf{E}, \mathbf{Q})$ containing a vertex set $\mathbf{V} = \{v_1, \ldots, v_{n_v}\}$, an edge set $\mathbf{E} = \{e_1, \ldots, e_{n_e}\}$, and a quad face set $\mathbf{Q} = \{q_1, \ldots, q_{n_q}\}$ are created for image warping, where n_v, n_e, and n_q represent the number of vertices, edges, and quads, respectively. In addition, a set of patches $\mathbf{P} = \{patch_1, \ldots, patch_{n_p}\}$ and its corresponding

significance values $\mathbf{S} = \{s_1, \ldots, s_{n_p}\}$ generated by the approach mentioned in Section III-B are used in warping. Here, n_p represents the number of patches. Assume that the original image of $m \times n$ pixels is resized into a new image of $m' \times n'$ pixels. The proposed warping aims at finding a deformed mesh geometry $\mathbf{V}' = \{v'_1, \ldots, v'_{n_v}\}$ in which the quads in a visually salient patch are forced to undergo a similarity transformation. To achieve this goal, two constraints, namely, *patch transformation constraint* and *grid line constraint*, are applied to the quads and the patches with an optimization solver. These two constraints are described as follows.

1) Patch Transformation Constraint: To preserve the patches with high significance and to avoid over-deformation on patches with low significance, two energy terms are introduced into the quad mesh optimization, namely, *rigid transformation* and *linear scaling*. To calculate these two energy terms, a representative edge is selected for each patch. The edge closest to the center of the patch is selected as the representative edge. Given an arbitrary edge $\mathbf{e} = (v_a, v_b)$, the geometry transformation $\mathbf{T}$ (containing a scale factor s and a rotation factor r) between edge $\mathbf{e}$ and representative edge, denoted by $\mathbf{C} = (v_1, v_2)$, can be formulated as follows:

$$\mathbf{e} = \mathbf{TC} \Rightarrow \begin{bmatrix} e_x \\ e_y \end{bmatrix} = \begin{bmatrix} s & r \\ -r & s \end{bmatrix} \begin{bmatrix} C_x \\ C_y \end{bmatrix} \Rightarrow \begin{bmatrix} s \\ r \end{bmatrix} = \begin{bmatrix} C_x & C_y \\ C_y & -C_x \end{bmatrix}^{-1} \begin{bmatrix} e_x \\ e_y \end{bmatrix}, \quad (1)$$

where

$$\begin{cases} C_x = v_{1_x} - v_{2_x} \\ C_y = v_{1_y} - v_{2_y}, \end{cases} \text{and} \begin{cases} e_x = v_{a_x} - v_{b_x} \\ e_y = v_{a_y} - v_{b_y}. \end{cases}$$

To address the problem of inconsistent deformation, the rigid transformation term is formulated as the rigidity of patches in quad mesh warping:

$$D_{ST}(patch_i) = s_i \times \sum_{\mathbf{e}'_j \in \mathbf{E}(patch_i)} \left\| \mathbf{e}'_j - \mathbf{T}_{ij}\mathbf{C}'_i \right\|^2, \quad (2)$$

where s_i is the significance value of patch i, and its range is $[0, 1]$. $\mathbf{e}'_j$ and $\mathbf{C}'_i$ represent the deformed edge in $patch_i$ and the deformed representative edge of $patch_i$, respectively. $\mathbf{T}_{ij}$ calculated by (1) is the geometry transformation between $\mathbf{e}_j$ and $\mathbf{C}_i$ in the original image, and this transformation is fixed in warping. Note that $\mathbf{T}_{ij}$ is computed for each edge individually within a patch. Therefore, the visually salient objects are forced to undergo as-rigid-as-possible transformation rather than rigid transformation in warping. In this manner, the preservation of visually salient content and the over-deformation avoidance of homogenous content can potentially be balanced. To further avoid over-deformation in low-significance patches, an energy term with respect to linear scaling is included:

$$D_{LT}(patch_i) = (1 - s_i) \times \sum_{\mathbf{e}'_j \in \mathbf{E}(patch_i)} \left\| \mathbf{e}'_j - \mathbf{L}\mathbf{T}_{ij}\mathbf{C}'_i \right\|^2, \quad (3)$$

where

$$\mathbf{L} = \begin{bmatrix} \frac{m'}{m} & 0 \\ 0 & \frac{n'}{n} \end{bmatrix}$$

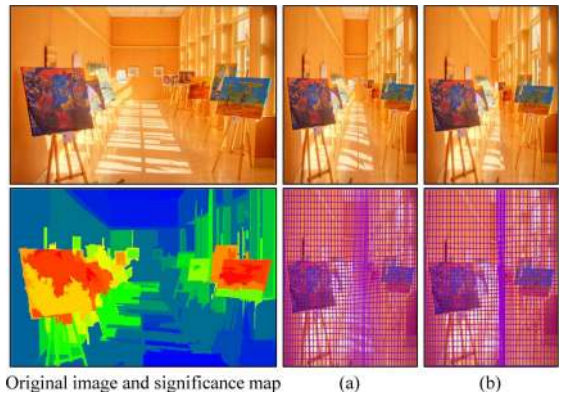

Fig. 4. The original image is narrowed by the proposed approach with the linear scaling energy term (a) and without this energy term (b).

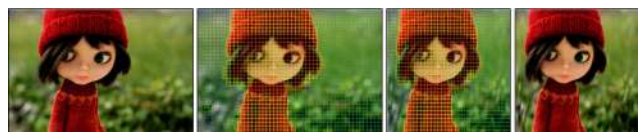

Fig. 5. Result of patch-based image warping. From left to right: original image, significance map with quad mesh (the quad color is the quad's significance value), deformed quad mesh, and retargeting result.

is the scaling matrix for linear scaling $(m \times n) \rightarrow (m' \times n')$. This energy term is used to measure the difference between a deformed patch and a linear-scaling patch. The weighting factor is set to $(1 - s_i)$, and thus, a low-significance patch has a large weight to avoid over-deformation. Fig. 4 shows a comparison of the proposed approach with and without the energy term of linear scaling. Without this term, some low-significance patches are greatly squeezed, and artifacts occur, as shown in Fig. 4(b). The total patch transformation energy is defined by summing up the individual patch energy term with its weight:

$$D_{TF}(\mathbf{P}) = \sum_{k=1}^{n_P} \Big(\alpha \times D_{ST}(patch_k) + (1 - \alpha) \times D_{LT}(patch_k) \Big), \quad (4)$$

where α is the weighting factor for the energy terms D_{ST} and D_{LT}. To preserve the shapes of high-significance patches, a large value is assigned to this parameter $(\alpha = 0.8)$.

2) Grid Orientation Constraint: The goal of grid orientation constraint is to avoid skew artifacts. Following the measurement of orientation distortion presented by Shirley *et al.* [26], this term is defined by measuring the grid line blending. A quad q has four vertices $\{v_a, v_b, v_c, v_d\}$ with two horizontal edges (v_a, v_b), (v_d, v_c) and two vertical edges (v_a, v_d), (v_b, v_c). This term is defined as quad deformation, and it is formulated as the distance of the y component between the end vertices of the deformed horizontal edges and the distance of the x component between the end vertices of the deformed vertical edges:

$$D_{OR}(\mathbf{M}) = \sum_{q \in \mathbf{Q}} \Big(\left\| v'_{a_y} - v'_{b_y} \right\|^2 + \left\| v'_{d_y} - v'_{c_y} \right\|^2 + \left\| v'_{a_x} - v'_{d_x} \right\|^2 + \left\| v'_{b_x} - v'_{c_x} \right\|^2 \Big), \quad (5)$$

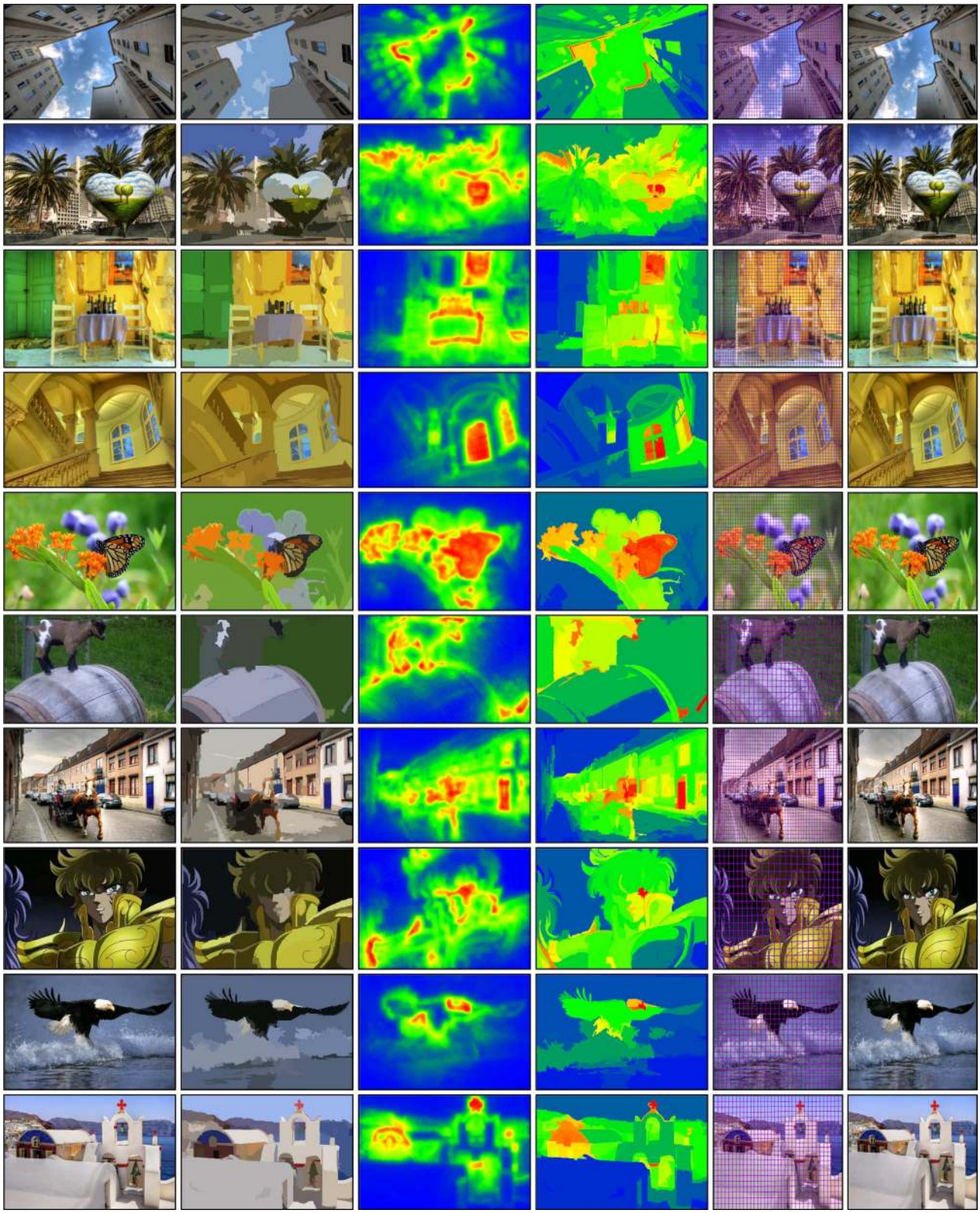

Fig. 6. Results of our approaches. From left to right: original images, segmentation results, saliency detection results, significance maps used in retargeting, deformed quad meshes, and retargeting results.

where the suffixes x and y represent the $x-$ and $y-$ component of the vertex position, respectively.

3) Optimization and Boundary Conditions: To find an optimized mesh geometry $\mathbf{V}' = \{v'_1, \ldots, v'_{n_v}\}$, the sum of patch transformation energy and quad orientation energy is minimized:

$$D(\mathbf{M}, \mathbf{P}) = D_{TF}(\mathbf{P}) + D_{OR}(\mathbf{M}), \tag{6}$$

subject to the boundary conditions:

$$v'_{i_y} = \begin{cases} 0 & \text{if } v_i \text{ is on the top boundary} \\ m' & \text{if } v_i \text{ is on the bottom boundary} \end{cases}$$

$$v'_{i_x} = \begin{cases} 0 & \text{if } v_i \text{ is on the left boundary} \\ n' & \text{if } v_i \text{ is on the right boundary} \end{cases}$$

The minimization of (6) yields a linear least-squares system $\mathbf{AV}' = \mathbf{b}$:

$$\begin{bmatrix} \mathbf{A}_{D_{TF}} \\ \mathbf{A}_{D_{OR}} \\ w \times \mathbf{v}_{left} \\ w \times \mathbf{v}_{right} \\ w \times \mathbf{v}_{top} \\ w \times \mathbf{v}_{bottom} \end{bmatrix} \begin{bmatrix} v'_{x_1} \\ \vdots \\ v'_{x_{n_v}} \\ v'_{y_1} \\ \vdots \\ v'_{y_{n_v}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ 0 \\ w \times n' \\ 0 \\ w \times m' \end{bmatrix}, \tag{7}$$

where the sub-matrices $\mathbf{A}_{D_{TF}}$ and $\mathbf{A}_{D_{OR}}$ contain n_v rows, respectively, for the patch transformation energy and the quad orientation energy; $\mathbf{v}_{left}$, $\mathbf{v}_{right}$, $\mathbf{v}_{top}$, and $\mathbf{v}_{bottom}$ containing n_{bv} rows (where n_{bv} is the number of boundary vertices) with a large weight w $(w = n_v)$ to enforce the minimization to meet the boundary conditions. This least-squares system has the unique solution $\mathbf{V}' = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$, and thus, the deformed mesh geometry $\mathbf{V}' = \{v'_1, \ldots, v'_{n_v}\}$ can be obtained. To ensure that the obtained mesh meets the boundary conditions and to handle a large system, the conjugate gradient method is adopted. This method solves the system iteratively until the movements of vertices are smaller than 0.5 pixels. The result of patch-based image warping is shown in Fig. 5. The high-significance patches are consistently deformed in warping to preserve the visually salient content.

IV. Experimental Results and Discussion

For a fair comparison, most images used in the related works were tested in our experiments. Some representative cases that images contain evident foreground objects or structure lines are shown in Figs. 6–13, and the others are attached to the supplementary documents. In the experiment, our approach is evaluated using a PC with 3.4 GHz CPU and 4 GB memory. On average, for a 1024×768 image, the computation time for warping is 0.063 seconds, and that for preprocessing is 10.07 seconds, including 0.06 seconds for segmentation and 10.01 seconds for saliency detection.

In Fig. 6, we show the results of the proposed processes, including image segmentation, saliency detection, significance map generation, quad mesh warping, and image retargeting. In our approach, each patch is assigned a significance value, and the quads within a patch are forced to undergo as-rigid-as-possible transformation. Thus, the shapes of visually salient objects and structure lines can be well preserved.

To demonstrate the robustness of our approach, an over-segmentated image generated by skipping the merge step in segmentation was tested (see Fig. 7). In this extreme case, a satisfactory retargeting result can still be obtained, indicating that our approach is only slightly sensitive to segmentation quality. Moreover, in Fig. 8, the images are resized horizontally and vertically. The aspect ratio of images are altered from 2:3 to 4:3 and 1:1. This experiment demonstrates that our approach can adapt images to fit display with various aspect ratios while preserving content.

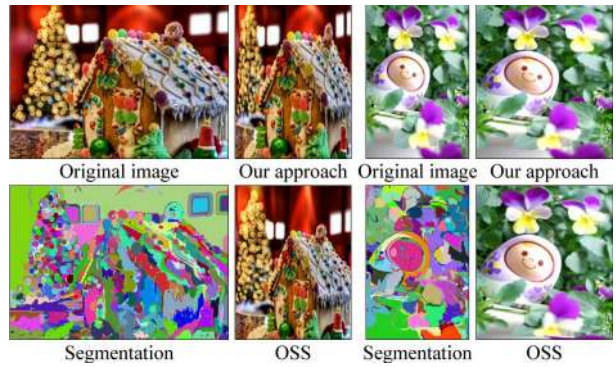

Fig. 7. Robustness demonstration. The retargeting results generated using over-segmented patches in our approach are compared with those generated by OSS [7]. The segmented patches are visualized by colors.

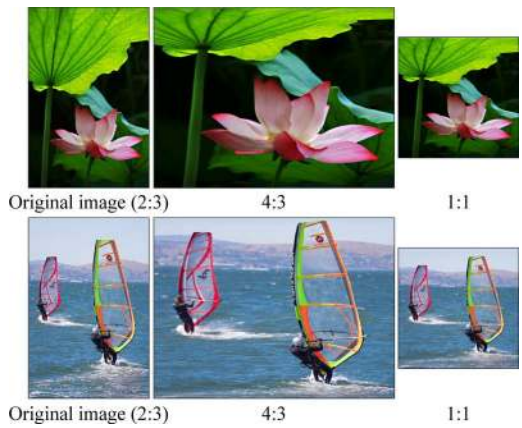

Fig. 8. Results of resizing images horizontally and vertically. The aspect ratio of original images is altered from 2:3 to 4:3 and 1:1.

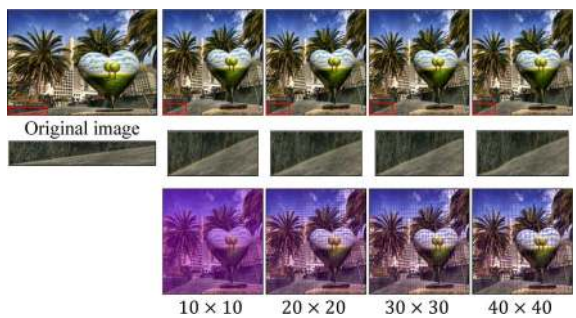

Fig. 9. Retargeting results using various quad resolutions, including 10 $(pixels) \times 10$ $(pixels)$, 20×20, 30×30, and 40×40. The images in the middle row are the close-up views of the retargeting results.

Fig. 10. Retargeting using various settings of parameter α.

The quad resolution and the weighting factor α in (4) are the main parameters in our approach. The quad mesh is used as the control mesh in warping, and the parameter α is the weighting factor for the energy terms of rigid transformation

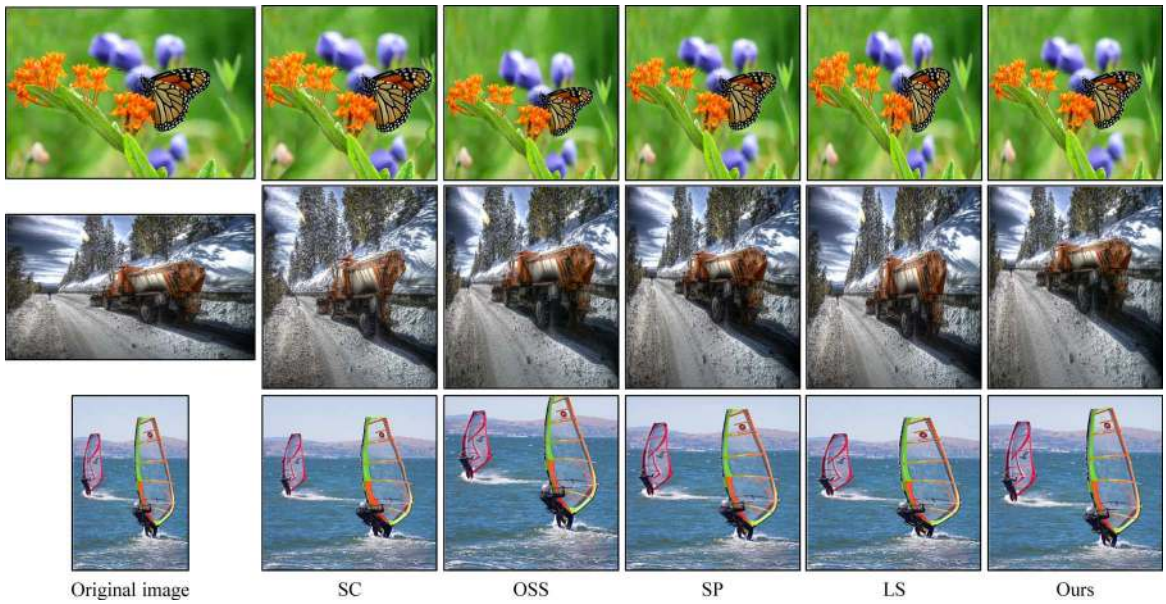

Fig. 11. Comparisons with the related approaches including seam carving (SC) [3], optimized scale-and-stretch (OSS) [7], shape preserving (SP) [8], and linear scaling (LS) by using images that contain evident foreground objects.

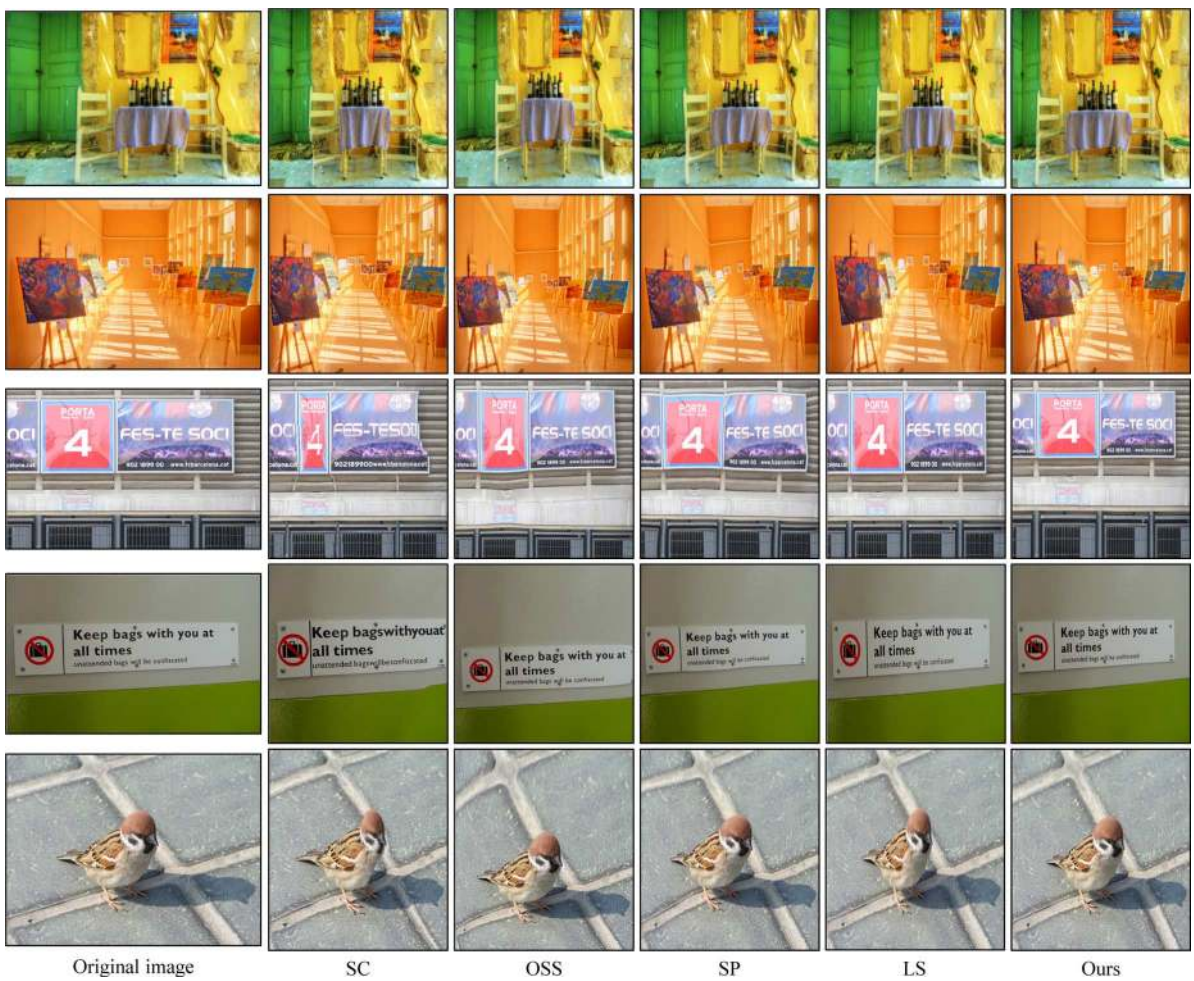

Fig. 12. Comparisons with the related approaches including SC [3], OSS [7], SP [8], and LS by using images that contain structure lines.

and linear scaling. To test how sensitive the retargeting results are to these parameters, various quad resolutions, including $10\ (pixels) \times 10\ (pixels)$, 20×20, 30×30, and 40×40, and various settings of the parameter α were tested. The results are shown in Figs. 9 and 10. Obviously, the higher the quad resolution we set, the higher the retargeting quality but the lower the computation efficiency we obtain (see the region marked by red rectangle). To consider both the quality and efficiency,

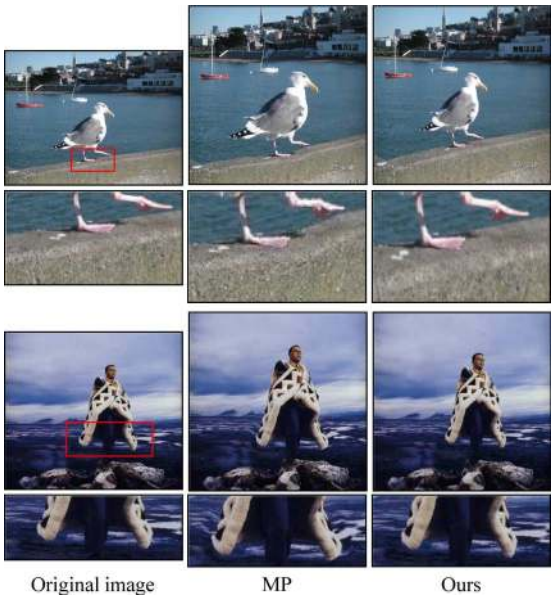

Fig. 13. Comparisons with the triangle-mesh-based approach (MP) [9].

the quad resolution is set to $20\,(pixels) \times 20\,(pixels)$. The parameter α controls how rigid the visually salient patches are in warping. Setting a larger weight forces the patches to be more rigid in warping, whereas setting a smaller weight forces the patches to be more like linear scaling. To preserve visually salient content, a large value ($\alpha = 0.8$) is assigned.

We compare our approach with the standard discrete retargeting approach, i.e., seam carving (SC) [3], the standard continuous retargeting approach, i.e., optimized scale-and-stretch (OSS) [7], and the linear scaling (LS). In addition, we also compare with the most related approaches including, quad-mesh-based approach (SP) [8], and triangle-mesh-based approach (MP) [9]. These two related approaches preserve both visually salient objects and structure lines, which are similar to ours. The comparisons with SC, OSS, SP, and LS are shown in Figs. 11 and 12, and the comparisons with MP are shown in Fig. 13. For a fair comparison with the OSS approach, the same quad resolution is used (quad resolution is set to 20×20). The SC approach [3] allows for higher flexibility in pixel removal, and thus, can be applied to some interesting applications such as object removal. However, the comparisons indicate that the discrete SC sometimes generates discontinuous artifacts, producing noticeable visual distortion. The OSS approach [7] has the advantage of absorbing distortion by homogeneous regions. However, the inconsistent deformation is sensitive to human visual system. The SP approach [8] and MP approach [9] have good improvements in shape preservation. However, the SP approach sometimes cannot well preserve structure lines because of inaccuracy of edge detection, and the MP approach may has the problem of inconsistency in triangle orientations. In contrast, our approach can ease the problem of inconsistent deformation through the proposed patch-based warping. In addition, using quad mesh with orientation constraints can significantly reduce the artificial effects caused by inconsistent orientations. These properties enable our approach to generate better retargeting results in terms of visual quality compared with the related approaches.

Combining cropping with image warping or seam carving for image retargeting has been proven to greatly improve visual quality [12], [22], [27]. Integrating cropping into the proposed scheme is easy. The optimal cropping is performed first, and our approach is then applied to resize the cropped results. A comparison between our approach combining with the cropping operation and the multi-operator approach [22] which combines the operators of cropping, scaling, and seam carving is given in Fig. 14. The results show that the multi-operator approach has a good improvement on shape preservation. However, discontinuity artifacts sometimes occur for the images having structure lines.

1) User Study: To evaluate our approach, a user study involving 76 participants with age ranging from 20 to 39 years old was conducted. The survey system provided by Rubinstein *et al.* [12] was used, in which the participants were shown two retargeted images side by side at a time and asked to choose the one they liked better. Following the work [12], the images having the attributes that can be mapped to the major retargeting objectives, *preserving content*, *preserving structure*, and *preventing artifact*, are used. The attributes are: *structure lines* and *evident foreground objects*. The image dataset is made up of 16 images having one or two attributes. We manually classify the dataset into two groups, and each group has eight images having the same attribute. Images having structure lines and foreground objects are grouped into A_1 and A_2, respectively. The goodness-of-fit test based on the chi-square distribution was conducted for each image group. In this test, the statistical significance level is set to 0.05, and null hypothesis is that the retargeting results generated by the approaches have the same quality. Thus, the expected frequency, i.e., the number of expected votes, is 38 for 76 participants, and $\chi^2_{0.05,7} = 14.07$ with seven degrees of freedom. The observed frequencies, i.e., the number of votes, are shown in Tables I and II; the results of the goodness-of-fit tests are shown in Table III. The survey and the chi-square test indicate that the retargeting results generated by SC, OSS, SP and our approach have similar quality for the images in group A_2 (χ^2_0 is close to $\chi^2_{0.05,7}$), and our approach is better than the test approaches for the images in group A_1 (χ^2_0 is greater than $\chi^2_{0.05,7}$), especially for the images with evident structure lines such as the 2nd, 3rd, 4th, and 5th test images. Therefore, our approach can preserve not only the foreground objects but also the structure lines, and our approach can potentially generate better retargeting results compared with the related approaches.

V. Conclusions, Limitation, and Future Work

A novel patch-based image warping is presented for retargeting. In our approach, the energy term of similarity transformation can force the visually salient content to undergo as-rigid-as-possible deformation, and the energy term of linear scaling can avoid over-deformation. Moreover, the mesh warping with the proposed significance map can smoothly propagate distortions. These processes efficiently ease the inconsistent deformations, enabling our approach to cope with images containing dense information and structure lines. The conducted comparisons and user study show the clear superiority of our approach over the related approaches in terms of visually salient

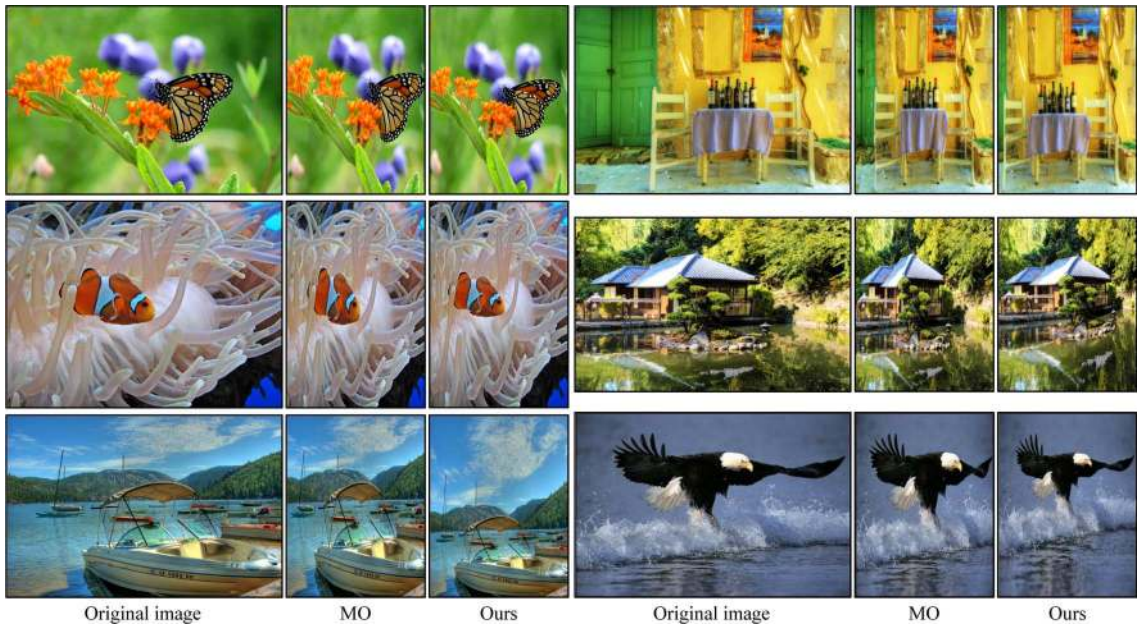

Fig. 14. Comparisons with the multi-operator approach (MO) [22].

TABLE I
OBSERVED FREQUENCY (DENOTED BY O_i) AND EXPECTED FREQUENCY (DENOTED BY E_i) FOR EACH TEST IMAGE HAVING THE ATTRIBUTE OF STRUCTURE LINES (GROUP A_1)

Group A_1										# of votes (the related approach : ours)
O_i	SC	3	6	1	6	6	18	12	9	61 : 547
	OSS	41	14	5	20	21	13	24	41	179 : 429
	SP	1	11	5	9	26	27	32	7	118 : 490
	LS	14	23	9	16	11	42	27	29	171 : 437
E_i	--	38	38	38	38	38	38	38	38	--

TABLE II
OBSERVED FREQUENCY AND EXPECTED FREQUENCY FOR THE TEST IMAGE HAVING THE ATTRIBUTE OF EVIDENT FOREGROUND OBJECTS (GROUP A_2)

Group A_2										# of votes (the related approach : ours)
O_i	SC	16	40	6	41	26	31	28	4	192 : 416
	OSS	25	18	23	34	25	25	46	41	237 : 371
	SP	21	34	46	20	25	40	51	39	276 : 332
	LS	7	11	14	9	34	37	8	13	133 : 475
E_i	--	38	38	38	38	38	38	38	38	--

TABLE III
THE VALUES OF χ_0^2 AND $\chi_{0.05,7}^2$, AND THE RESULTS OF GOODNESS-OF-FIT TEST

		SC	OSS	SP	LS
χ_0^2	A_1	199.55	82.03	139.21	106.93
	A_2	78.16	32.13	27.26	146.67
$\chi_{0.05,7}^2$	--	14.07	14.07	14.07	14.07
Acceptance / Rejection (A_1 / A_2)		r / r	r / r	r / r	r / r

preservation. Nevertheless, our approach has the limitation that it cannot work well for images containing objects with similar significance values. The retargeting results in this case are similar to those generated by linear scaling, as shown in Fig. 15,

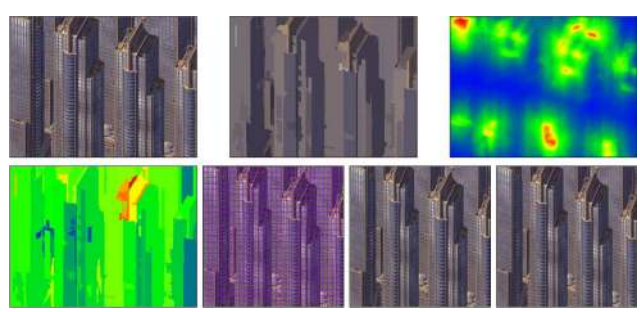

Fig. 15. Retargeting result for the image containing objects with similar significance values. Top: original image (left), segmentation result (middle), and saliency detection result (right); bottom (from left to right): significance map, deformed quad mesh, our result, and linear scaling.

since all the objects are deformed with similar weights. In the future, we plan to integrate linear scaling into our system to deal with this special case. In addition, we plan to extend our method to cope with video or other temporal and geospatial data.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their insightful comments. Our special thanks go to O. Sorkine for providing the user study system, to Y. Guo and F. Liu for their experimental data, and to Y.-S. Wang for his helpful suggestions on the implementation and draft of this paper.

REFERENCES

[1] F. Liu and M. Gleicher, "Automatic image retargeting with fisheye-view warping," in *Proc. 18th Annual ACM Symp. User Interface Software and Technology*, 2005, pp. 153–162.
[2] R. Gal, O. Sorkine, and D. Cohen-Or, "Feature-aware texturing," in *Proc. Eurographics Symp. Rendering*, 2006, pp. 297–303.
[3] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM Trans. Graph.*, vol. 26, no. 3, Jul. 2007.
[4] L. Wolf, M. Guttmann, and D. Cohen-Or, "Non-homogeneous content-driven video-retargeting," in *Proc. Eleventh IEEE Int. Conf. Computer Vision (ICCV-07)*, 2007, pp. 1–6.
[5] M. Rubinstein, A. Shamir, and S. Avidan, "Improved seam carving for video retargeting," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 16:1–16:9, Aug. 2008.
[6] A. Shamir and S. Avidan, "Seam carving for media retargeting," *Commun. ACM*, vol. 52, no. 1, pp. 77–85, Jan. 2009.
[7] Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee, "Optimized scale-and-stretch for image resizing," *ACM Trans. Graph.*, vol. 27, no. 5, pp. 118:1–118:8, Dec. 2008.
[8] G.-X. Zhang, M.-M. Cheng, S.-M. Hu, and R. R. Martin, "A shape-preserving approach to image resizing," *Comput. Graph. Forum*, vol. 28, no. 7, pp. 1897–1906, 2009.
[9] Y. Guo, F. Liu, J. Shi, Z.-H. Zhou, and M. Gleicher, "Image retargeting using mesh parametrization," *IEEE Trans. Multimedia*, vol. 11, no. 5, pp. 856–867, Aug. 2009.
[10] Y. Jin, L. Liu, and Q. Wu, "Nonhomogeneous scaling optimization for realtime image resizing," *Vis. Comput.*, vol. 26, no. 6–8, pp. 769–778, Jun. 2010.
[11] D. Han, M. Sonka, J. Bayouth, and X. Wu, "Optimal multiple-seams search for image resizing with smoothness and shape prior," *Vis. Comput.*, vol. 26, no. 6–8, pp. 749–759, Jun. 2010.
[12] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir, "A comparative study of image retargeting," *ACM Trans. Graph.*, vol. 29, no. 6, pp. 160:1–160:10, Dec. 2010.
[13] Y.-J. Liu, X. Luo, Y.-M. Xuan, W.-F. Chen, and X.-L. Fu, "Image retargeting quality assessment," *Comput. Graph. Forum*, vol. 30, pp. 583–592, 2011.
[14] M. Frankovich and A. Wong, "Enhanced seam carving via integration of energy gradient functionals," *Signal Process. Lett.*, vol. 18, no. 6, pp. 375–378, 2011.
[15] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," in *Proc. IEEE CVPR*, 2010, pp. 2376–2383.
[16] A. Shamir and O. Sorkine, "Visual media retargeting," in *ACM SIGGRAPH ASIA 2009 Courses*, 2009, pp. 11:1–11:13.
[17] B. Suh, H. Ling, B. B. Bederson, and D. W. Jacobs, "Automatic thumbnail cropping and its effectiveness," in *Proc. 16th Annual ACM Symp. User Interface Software and Technology*, 2003, pp. 95–104.
[18] L.-Q. Chen, X. Xie, X. Fan, W.-Y. Ma, H. Zhang, and H.-Q. Zhou, "A visual attention model for adapting images on small displays," *Multimedia Syst.*, vol. 9, no. 4, pp. 353–364, 2003.
[19] F. Liu and M. Gleicher, "Video retargeting: Automating pan and scan," in *Proc. 14th Annual ACM Int. Conf. Multimedia*, 2006, pp. 241–250.
[20] A. Santella, M. Agrawala, D. DeCarlo, D. Salesin, and M. Cohen, "Gaze-based interaction for semi-automatic photo cropping," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, 2006, pp. 771–780.
[21] C. Tao, J. Jia, and H. Sun, "Active window oriented dynamic video retargeting," in *Proc. Workshop on Dynamical Vision*, 2007, pp. 1–12.
[22] M. Rubinstein, A. Shamir, and S. Avidan, "Multi-operator media retargeting," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 23:1–23:11, Jul. 2009.
[23] Y. Pritch, E. Kav-Venaki, and S. Peleg, "Shift-map image editing," in *Proc. IEEE ICCV'09*, Sep. 2009, pp. 151–158.
[24] Q. X. Huang, R. Mech, and N. Carr, "Optimizing structure preserving embedded deformation for resizing images and vector art," *Comput. Graph. Forum*, vol. 29, no. 7, pp. 1887–1896, 2009.
[25] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, Sep. 2004.
[26] P. Shirley, M. Ashikhmin, M. Gleicher, S. Marschner, E. Reinhard, K. Sung, W. Thompson, and P. Willemsen, *Fundamentals of Computer Graphics*, 2nd ed. Abington, U.K.: A. K. Peters, 2005.
[27] Y.-S. Wang, H.-C. Lin, O. Sorkine, and T.-Y. Lee, "Motion-based video retargeting with optimized crop-and-warp," *ACM Trans. Graph.*, vol. 29, no. 4, pp. 90:1–90:9, Jul. 2010.

Shih-Syun Lin received the B.S. degree in Applied Mathematics from Providence University, Taiwan in 2007 and the M.S. degree in Graduate Institute of Educational Measurement and Statistics from National Taichung University, Taiwan in 2010. Recently, he is a Ph.D. student in Department of Computer Science and Information Engineering, National Cheng-Kung University, Taiwan. His research interests include video retargeting, mesh deformation, pattern recognition, and computer graphics.

I-Cheng Yeh received the B.S. degree from the Department of Computer and Information Science, National Chiao Tung University, Taiwan in 2004. Following the studies, he received the Ph.D. degree in Computer Science and Information Engineering from National Cheng Kung University, Taiwan in 2011. Now, he is performing compulsory military service in the Republic of China Marine Corps. His research interests include discrete surfaces parameterization and manipulation, human motion visualization and digital image manipulation.

Chao-Hung Lin (M'10) received his M.S. and Ph.D. degree in computer science and information engineering from National Cheng-Kung University, Taiwan in 1998 and 2004, respectively. Since 2006, he has been a member of the faculty of the Department of Geomatics at National Cheng-Kung University. He currently is an associate professor. He leads the Digital Geometry Processing Laboratory (http://dgl.geomatics.ncku.edu.tw) and co-leads the Computer Graphics Laboratory, National Cheng-Kung University (http://graphics.csie.ncku.edu.tw). His current research interests include remote sensing, point cloud processing, digital map generation, information visualization, and computer graphics. He served as an editorial board member of the *International Journal of Computer Science and Artificial Intelligence*, and served as a member of the international program committees of Pacific Graphics 2011. He is a member of the ACM.

Tong-Yee Lee (SM'10) received the Ph.D. degree in computer engineering from Washington State University, Pullman, in May 1995. He is currently a distinguished professor in the Department of Computer Science and Information Engineering, National Cheng-Kung University, Tainan, Taiwan. He leads the Computer Graphics Group, Visual System Laboratory, National Cheng-Kung University (http://graphics.csie.ncku.edu.tw/). His current research interests include computer graphics, non-photorealistic rendering, medical visualization, virtual reality, and media resizing. He is a member of the ACM.