# Patch-based Image Warping for Content-aware Retargeting

Shih-Syun Lin, I-Cheng Yeh, Chao-Hung Lin and Tong-Yee Lee, *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 global and local structure lines, which are sensitive to the human visual system. In this paper, we preserve the shapes of both highly salient objects and structure lines while minimizing visual distortions. To ease the unpleasant visual distortions caused by inconsistent warping of visually salient content, a patch-based image-resizing scheme with an extended significance value measurement is adopted. In this scheme, a similarity transformation constraint is used to force an as-rigid-as-possible deformation on the visually salient content, while an optimization process is performed to smoothly propagate the distortions. This scheme enables our approach to yield agreeable content-aware image warping and retargeting. The experimental results show that the proposed approach can effectively preserve the shapes of both highly salient objects and structure lines, and thus a satisfying retargeting result is obtained. The conducted user study also shows that the results generated by the proposed approach 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 and has drawn significant attention in the fields of image processing and computer graphics. The naive approaches, namely, linear scaling and uniform cropping, have been proven inappropriate when the aspect ratio is changed significantly. Thus, considerable research efforts have been devoted to content-aware retargeting [1]–[12]. The key point of content-aware retargeting is to resize images to arbitrary aspect ratio while keeping visually salient objects in similar aspect ratio. Thus, minimizing visual distortion is the main requirement of image retargeting.

Seam carving [1] and image warping [9] are recently proposed to resize images nonhomogeneously. Seam carving iteratively removes or inserts a seam passing through unimportant regions. This approach may generate jagged edges because of the removal of discontinuous seam. In contrast, image warping offers a better possibility of producing a continuous deformation for content-aware retargeting. Wang et al. [9] proposed an optimized scale-and-stretch warping using a quad mesh as a control mesh. This approach has the advantage of distributing distortions caused by warping to homogeneous regions, as it forces quads with highly important content to scale uniformly and distorts quads with homogeneous content. Thus, the approach can efficiently preserve the shapes of local objects. However, for a highly salient object or a structure line that occupies many quads, an inconsistent deformation may occur because of the inconsistent scaling factors of quads [10]. This inconsistency may result in an unnatural deformation, causing significant visual distortion even if a good significance map is used (Figure 1 (c)-(h)). To solve this problem and to preserves the shape of visually salient objects, a patch-based image warping approach is proposed. First, the source image is segmented into several patches in preprocessing, and each patch is assigned a significance value. Then, the significant patches 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 similarity transformation constraints, the proposed approach has a good chance of preserving the visually salient content (Figure 1 (b)). In addition, the use of distortion propagation and patch-based significance map reduces the need of perfect object segmentation, making our approach feasible to cope with various cases.

Compared with the recent studies on image retargeting,

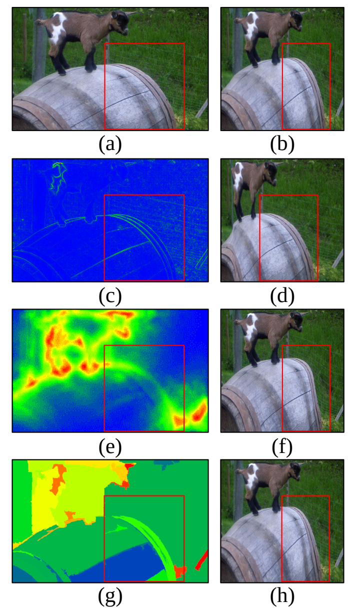



Fig. 1. Comparisons of image retargeting with various significance maps. (a) Source image; (b) result generated by the proposed patch-based image retargeting approach; (c), (e), (g) significance maps generated by the approaches [9], [13], and our approach. Pixel significance is visualized by colors ranging from red (the highest significance) to blue (the lowest significance) ; (d), (f), (h) the corresponding retargeting results generated by the warping approach [9].

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

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

the proposed approach offers the following contributions. A patch-based image retargeting approach with optimal distortion propagation is proposed to ease the unwanted visual distortion 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

Content-aware image retargeting has drawn significant attention, and many studies have been proposed in recent years. Following the categorization suggested by Shamir and Sorkine [14], the image retargeting approaches are classified into two categories: *discrete* and *continuous*. In the discrete approaches, the image is resized by cropping [15]–[19] or seam carving [1]–[5]. In cropping, an optimal rectangle is selected on the image, and the regions outside this rectangle are cut from the image. In seam carving, a seam is iteratively removed to preserve image contents. 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. [20] presented a multi-operator algorithm that combines cropping, linear scaling, and seam carving to optimally resize images. Similarly, Pritch et al. [21] suggested an approach that discretely 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 cases that the highly salient objects are near the borders of images, and the discontinuous seam removal may result in undesirable artifacts.

In contrast to discretely removing pixels in homogenous regions, the continuous approach optimizes mapping or warping using several deformation and smoothness constraints to preserve image content [6]–[12]. In the work of Liu et al. [6], non-linear warping is used to preserve important content. However, image features outside the region of interest may suffer from significant distortions. Gal et al. [7] suggested a texturing approach based on Laplacian editing. This approach warps images with similarity transformation constraints on local objects, leading to content-aware resizing. Wolf et al. [8] retargeted an image by merging less important pixels to reduce distortion. However, the distortion can only be propagated along the resizing direction. To improve the distortion propagation, Wang et al. [9] proposed 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 distortion 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. [10], Guo et al. [11], Jin et al. [12], and Hung et al. [22] force highly salient objects to undergo similarity transformations when resizing images. These approaches perform very well on the shape preservation of local objects. However, inconsistent deformations may still occur on structure lines, as they mainly rely on accurate handle setting [10] or rigidity map generation [22]. By contrast, the proposed approach can effectively preserve both the visually salient objects and structure lines while smoothly propagating the distortions, resulting in agreeable content-aware image retargeting.

Fig. 3. Significance map generation. From left to right: the source image, the segmentation result (each color represents a patch), the segmented patches displayed by the representative colors, and the merge result and the significance map.

The proposed approach is inspired by the previous work [9], and the goal of shape preservation is similar to [10]–[12]. However, there are substantial differences between our approach and these approaches. First, a patch-based image-retargeting scheme with a patch-based significance value measurement is adopted to ease the unpleasant visual distortions caused by inconsistent warping. Second, a similarity transformation constraint is used to force as-rigid-as-possible deformation on the visually salient content, while an optimization process is performed to smoothly propagate the distortions. Third, instead of using a triangle mesh as a control mesh in image warping, which may have the problem of inconsistent triangle orientations [11], [12], following the approach [9], we use a quad mesh with grid orientation constraints to avoid skew artifact.

## III. Content-aware Image Resizing

### A. System overview

Figure 2 schematically illustrates the proposed method that consists of two main steps: *preprocessing* and *image retargeting*. In the preprocessing step, the input image is segmented into several homogenous patches by the graph-based image segmentation approach [23]. A significance value measurement for the segmented pathces is then performed to generate a significance map for retargeting. In the generation of significance map, the context-aware saliency estimation [13] is adopted to generate a saliency value for each pixel. Afterwards, each segmented patch is assigned an averaged saliency value to ease the problem of inconsistent deformation. In the image retargeting step, a quad mesh is created to cover the image, and patch-based image warping is performed to force the quads of each significant patch to undergo an as-rigid-as-possible deformation in the resizing process, and an optimization solver is used to smoothly propagate the distortion.

### B. Object-based significance map generation

Many measurements on the significance value determination of pixels have been proposed in image retargeting approaches. Pixels with large gradient magnitudes are generally considered significant pixels [2], [8], and pixel gradient is combined with the pixel saliency value to determine the pixel significance value [9]. However, inconsistent deformation may arise when retargeting is performed using such pixel-based measurements. This inconsistency leads to an unnatural deformation, especially

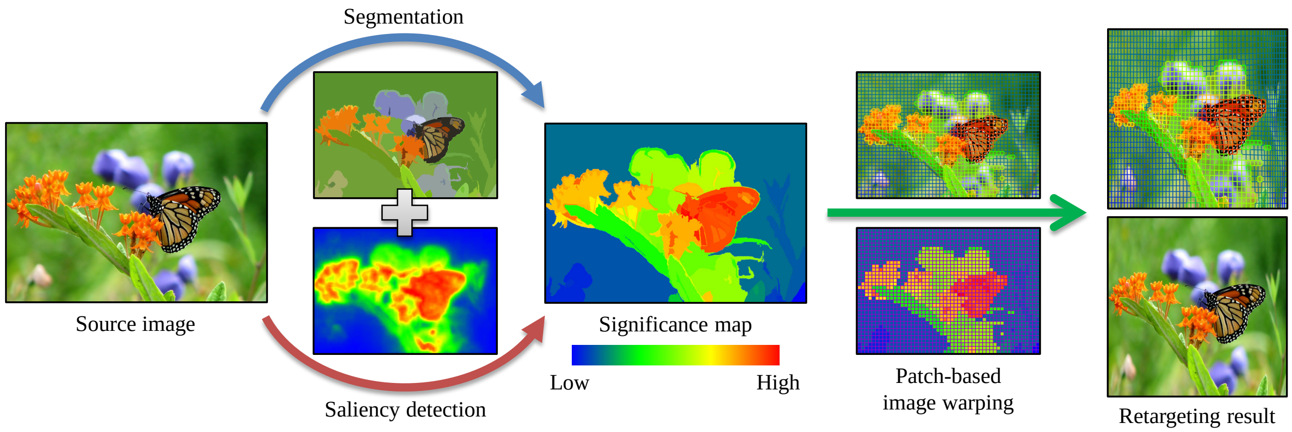



Fig. 2. Overview of the proposed approach. Left: the source image; Middle: the significance map generated by combining the object segmentation and the saliency detection; Right: the retargeting result generated by the proposed patch-based image warping. In the result, the quads of significant patches, e.g., the patches of butterfly and flower, are well preserved.

in the visually salient content. For instance, the retargeting results shown in Figures 1(d) and (f) are generated using the approach presented in [9] with the pixel-based significance measurements presented by Wang et al. [9] and Goferman et al. [13], respectively. The results show that the barrel has unnatural deformation caused by the inconsistent significance values. Artifacts occurring in highly salient objects are sensitive to the human visual system; thus, this unnatural deformation will lead to considerable visual distortion. In this study, with the aim of preserving entire visually salient objects, we adopt a patch-based significance value determination approach instead of a pixel-based one. In the proposed approach, the source image is initially segmented into several homogenous patches, after which each patch is assigned a significance value. Therefore, a patch containing an object or some structure lines can potentially be deformed consistently. In the first step, the graph-based image segmentation approach [23] is adopted to partition images. In general, numerous patches are generated, and over-segmentation occurs (Figure 3). To address this problem, a commonly used merge process is performed. Each patch is assigned an average pixel 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 in all experiments), the patch is merged to the 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 threshold for this criterion is set to 20 in all experiments. Once segmentation is obtained, the saliency detection proposed by Goferman et al. [13] is adopted. The following is a brief introduction to saliency detection. In [13], the context-aware 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 high saliency value is assigned to nonhomogeneous regions while 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 path is simply assigned an average significance value after saliency detection. By utilizing this measurement, the proposed approach can generate satisfying results (Figure 1(b)). Note that generating perfect segmentation for all cases is impossible even when a state-of-the-art segmentation approach is adopted. Fortunately, with the aid of smooth distortion propagation and the generated significance map, the proposed approach can address the problems caused by imperfect object segmentation. For instance, in Figure 1(g), the sheep in the image is partitioned into many patches with imperfect boundaries. In this case, the effect of inconsistent deformation is reduced (Figure 1 (b)). Thus, the result is better than those generated by the approaches that use a pixel-based significance measurement (Figures 1 (d) and (f)).

### C. *Patch-based image warping*

Following the warping approach presented by Wang et al. [9], a quad mesh $\mathbf{M} = (\mathbf{V}, \mathbf{E}, \mathbf{Q})$ containing a vertex set $\mathbf{V} = \{v_1, ..., v_{n_v}\}$ , an edge set $\mathbf{E} = \{e_1, ..., e_{n_e}\}$, and a quad face set $\mathbf{Q} = \{q_1, ..., q_{n_q}\}$ 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 $\mathbf{P} = \{patch_1, ..., patch_{n_p}\}$ and its corresponding significance values $\mathbf{S} = \{s_1, ..., s_{n_p}\}$ generated by the approach mentioned in Section III-B are used in warping. Here, $n_p$ represents the number of patches. Assume that the source image of $m \times n$ pixels is resized into a new image of $m' \times n'$ pixels. The proposed image warping aims at finding a deformed mesh geometry $\mathbf{V}' = \{v'_1, ..., v'_{n_v}\}$ in which the quads in a highly salient patch are forced to undergo a similarity transformation, while the distortions are smoothly propagated. To achieve this goal, two soft 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.

**Patch transformation constraint**. To preserve the shapes of patches with high significance values and to avoid over-deformation of patches with low significance values, 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 the edge $\mathbf{e}$ and the representative edge, denoted by $\mathbf{C} = (v_1, v_2)$, can be formulated as follows:

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

where

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

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

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

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 Eq. 1 is the geometry transformation between $\mathbf{e}_j$ and $\mathbf{C}_i$ in the source image; these transformations are fixed in warping. The geometric meaning of this term is to measure the changes of edge geometric relations, i.e., $\mathbf{T}_{ij}$, in warping. To avoid over-deformation in low-significance patches, an energy term with respect to linear scaling is added:

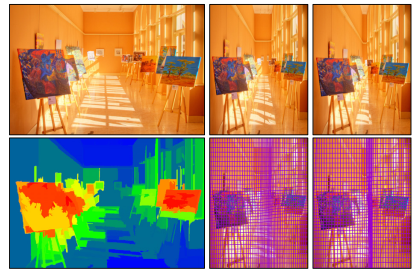
$$D_{LT}(patch_i) = (1 - s_i) \times \sum_{\mathbf{e}'_j \in E(patch_i)} \left\| \mathbf{e}'_j - \mathbf{L}\mathbf{T}_{ij}\mathbf{C}'_i \right\|^2, \tag{3}$$

where

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

is the scaling matrix for linear scaling $(m \times n) \to (m' \times n')$. 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-significant patch has a large weight to avoid over-deformation. In Figure 4, a comparison of the proposed approach with and without the energy term of linear scaling is given. Without the term of linear scaling, some low-significant patches may disappear, and artifacts may occur, as shown in Figure 4(b). The total patch transformation energy is defined by summing up the individual patch energy term with its significance weight:

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



Source image and significance map (a) (b)

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

where $\alpha$ is the weighting factor for the energy terms $D_{ST}$ and $D_{LT}$. To preserve the shapes of objects in highly significant patches, a large value is assigned to $\alpha$. In all experiments, $\alpha$ is set to 0.8.

**Grid orientation constraint.** The goal of grid orientation constraint is to avoid further skew artifacts. Following the measurement of orientation distortion presented by Shirley et al. [24], this energy term is defined by measuring the grid line blending. Assume that 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 energy term is defined as the 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:

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

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

**Optimization and boundary conditions.** To find an optimized mesh geometry $\mathbf{V}' = (v'_1, ..., v'_{n_v})$ , the sum of the patch transformation energy and the quad orientation energy is minimized:

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

subject to the following 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 Eq. 6 yields a linear least-squares system $\mathbf{AV}' = \mathbf{b}$:

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

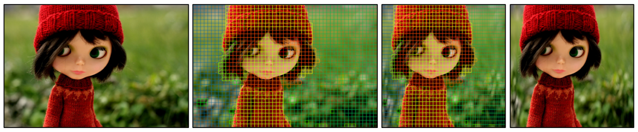


Fig. 5. Result of patch-based image warping. From left to right: the source image, the significance map with quad mesh (the quad color is the quad's significance value), the deformed quad mesh, and the retargeting result.

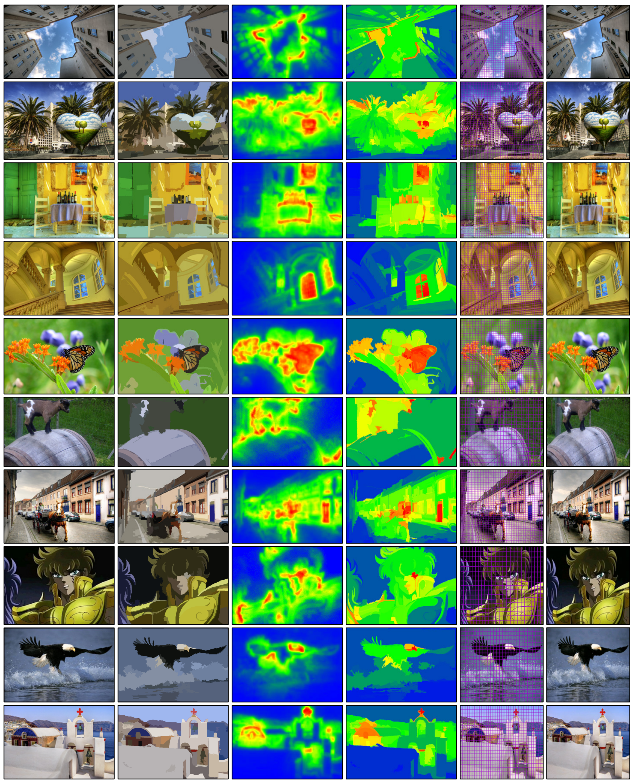


Fig. 6. The results generated by our approach. From left to right: the source images, the segmentation results, the saliency maps, the significance maps used in the retargeting, the deformed quad meshes, and the retargeting results.

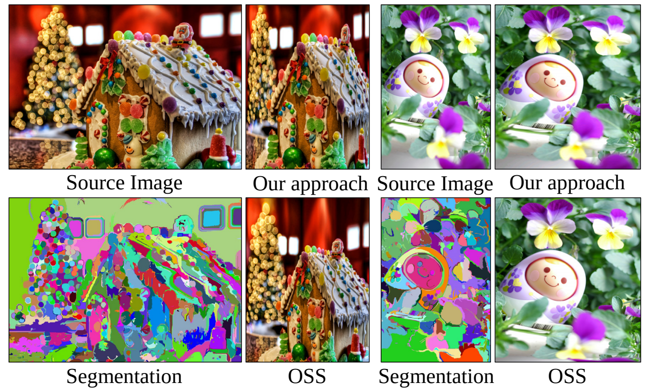



Fig. 7. Robustness demonstration. The images with serious over-segmentation (those that skip the merge process) are tested. The segmented patches are visualized by colors. The retargeting results generated using the over-segmented patches in our approach are compared with those generated by OSS [9].

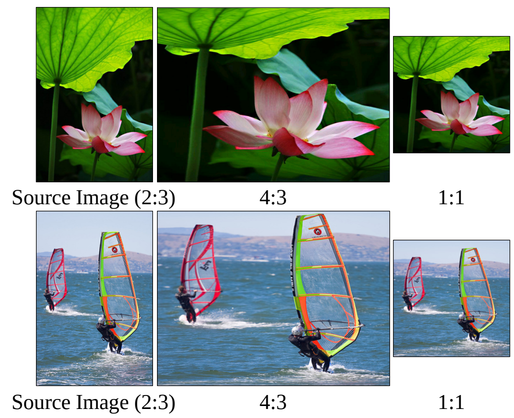



Fig. 8. Results of resizing images horizontally and vertically. The aspect ratio of source images is altered from 2:3 to 4:3 and 1:1, respectively.

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 $w$ to enforce the minimization to meet the boundary conditions. In the experiment, $w$ is set to $n_v$.This least-squares system has an unique solution $\mathbf{V}' = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$, and thus the deformed mesh geometry $\mathbf{V}' = (v_1', ..., v_{n_v}')$ can be obtained. To make sure that the deformed mesh is strictly constrained by the boundary conditions, this equation is solved iteratively until the movements of boundary vertices and internal vertices are smaller than 0.5 pixels. The result of patch-based image warping is shown in Figure 5. Each patch is warped to undergo a similarity transformation to preserve the shape of local objects and structure lines.

## IV. Experimental Results and Discussion

To give a fair comparison, most of the 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 Figures 6-13; the others are attached to the accompanying files. The experimental results are evaluated using a PC with 3.4 GHz CPU and 4 GB memory. On average, for a $1024 \times 768$ image, the computation time for resizing is 0.063s, and that for preprocessing is 10.07s (including 0.06s for segmentation and 10.01s for saliency detection). To demonstrate the ability of shapes preservation, some images containing dense information or evident foreground objects are tested. The segmented objects, saliency maps, significance maps, deformed quad meshes, and retargeting results are shown in Figure 6. 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 objects and structure lines can be preserved efficiently. To demonstrate the robustness of our approach, an image with serious over-segmentation (those that skip the process of patch merge) is tested (see Figure 7). In this extreme case, a satisfying retargeting result can still be obtained, indicating that our approach can cope with various cases. Moreover, in Figure 8, the image is resized horizontally and vertically, and the aspect ratio of the source images are altered from 2:3 to 4:3 and 1:1, respectively. This experiment demonstrates that our approach can adapt images to fit display with various aspect ratios and sizes.

The quad resolution and weighting factor $\alpha$ in Eq. 4 are

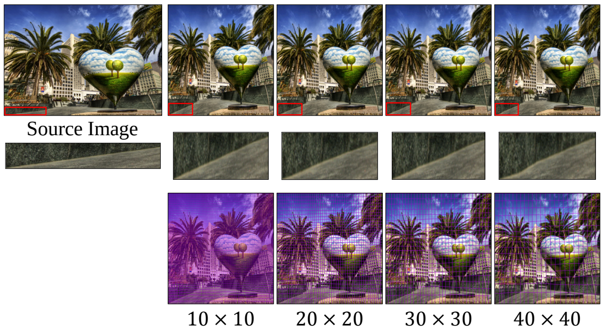



Fig. 9. Retargeting results using various quad resolutions, including 10 (pixels) x 10 (pixels), 20 x 20, 30 x 30, and 40 x 40. The images in the middle row are the close-up views of the retargeting results.


Fig. 10. Retargeting results using different values of parameter $\alpha$ .

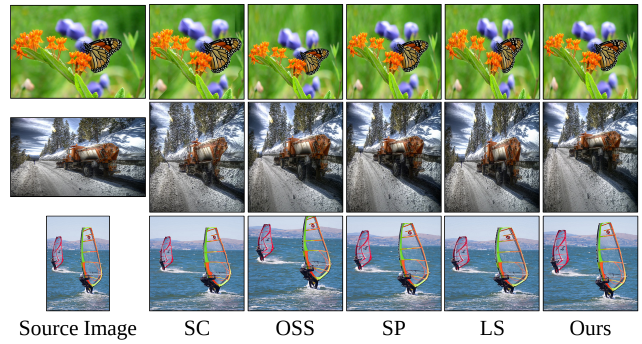



Fig. 11. Comparisons with the related approaches including seam carving (SC) [1], optimized scale-and-stretch (OSS) [9], shape preserving (SP) [10], and linear scaling (LS) by using the images that contain evident foreground objects.

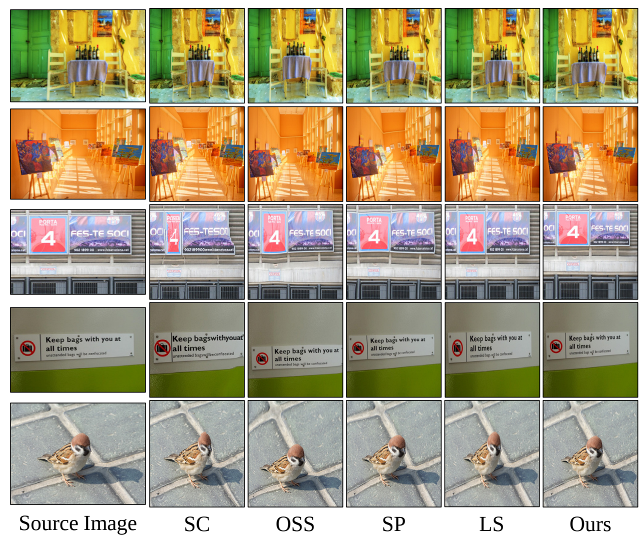



Fig. 12. Comparisons with the related approaches including SC [1], OSS [9], SP [10], and LS by using the images that contain structure lines.

the main parameters in our approach. The quad mesh is used as a control mesh in warping, and parameter $\alpha$ is the weighting factor of the energy terms, rigid transformation and linear scaling. To test how sensitive the retargeting results are to these parameters, various quad resolutions, including 10 (pixels) x 10 (pixels), 20 x 20, 30 x 30, and 40 x 40, and various values of the weighting factor $\alpha$ are tested. The results are shown in Figures 9 and 10. Clearly, 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 a red rectangle). To consider both quality and efficiency, the quad resolution is set to 20 (pixels) x 20 (pixels) in the experiments. As for the parameter $\alpha$, it can control how rigid the highly salient patches are in warping. Setting a larger value forces the patches to be more rigid in warping, whereas setting a smaller value forces the patches to be more like linear scaling. To preserve the shapes of visually salient objects, a large value is assigned. In all experiments, $\alpha$ is set to 0.8.

The image retargeting approaches can be categorized into discrete and continuous. Therefore, we compare our approach with the standard discrete resizing approach, i.e., seam carving (SC) [1], the standard continuous resizing approach, i.e., optimized scale-and-stretch (OSS) [9], and the linear scaling (LS). In addition, we also compare with the most related approaches including, quad-mesh-based approach (SP) [10], and triangle-mesh-based approach (MP) [11]. 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 Figures 11 and 12, and the comparisons with MP are shown in Figure 13. To give a fair comparison with the OSS approach, the same quad resolution is used (quad resolution is set to $20 \times 20$). The SC approach [1] allows for higher flexibility in pixel removal, and thus can be applied to some interesting applications such as object removal. However, the comparison indicates that the discrete SC sometimes generates artifacts on visually salient objects, producing noticeable visual distortion. The OSS approach [9] has the advantage of absorbing distortion by homogeneous regions. However, the inconsistent deformation on the visually salient content is sensitive to human visual system. The SP approach [10] has significant improvements in shape preservation. However, the structure lines sometimes cannot be preserved well because of the inaccuracy of edge detection. The MP approach [11] has good performance in shape preservation. However, the inconsistency of triangle orientations generates significant distortions in the regions near the visually salient objects, resulting in undesirable artificial effects. By contrast, our approach can ease the problem of inconsistent deformation through the proposed patch-based image warping. In addition, using quad mesh with orientation constraints can significantly reduce the artificial effects caused by the inconsistent orientations. These properties enable our approach to generate better results in terms of visual quality compared with the re-

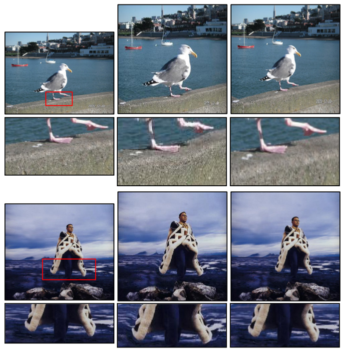



Fig. 13. Comparisons with the triangle-mesh-based approach (MP) [11].


Fig. 14. Comparisons with the multi-operator approach [20].

lated approaches. Note that performing exact comparisons without obtaining the codes of the related approaches is difficult. Thus, the images tested in the work [11] are used in the comparison instead of re-implementing this approach.

Combining cropping with image warping or seam carving for image retargeting has been proven to great improve visual quality [20], [25], [26]. 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 combined with the cropping operation and the multi-operator approach [20], which combines the operators of cropping, scaling, and seam carving, is given in Figure 14. The results show that multi-operator approach can have great improvements in content-aware resizing. However, some discontinuity artifacts can still occur in the structure lines.

**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. [26] was used. In this system, the paired comparison was adopted, in which the participants were shown two retargeted images side by side at a time and asked to choose the one they liked better (for more details, refer to our survey website http://140.116.246.46/survey/index.php?mode=0). Following the work [26], the images having the attributes that can be mapped to the major retargeting objectives, namely, *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 more of these attributes, and the dataset is manually classified into two groups by attribute. Each group has eight images having the same attribute. The goodness-of-fit test based on the chi-square distribution was conducted for each image group. In this test, the statistical significance level was set to 0.05, i.e., $\alpha = 0.5$, 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 Tables III. For the images having the attribute of structure lines (group $A_1$), $\chi_0^2$ is greater than $\chi^2_{0.05,7}$, indicating that our approach is better than the related approaches for the images in group $A_1$. As for the images having the attribute of evident foreground objects (group $A_2$), $\chi_0^2$ is close to $\chi^2_{0.05,7}$ in the approaches OSS and SP, and $\chi_0^2$ is greater than $\chi^2_{0.05,7}$ in the approaches SC and LS. This indicates that our approach is better than SC and LS, and slightly better than OSS and SP for the images in group $A_2$. The user study shows the clear superiority of our approach over the related approaches, especially for the images containing structure lines.

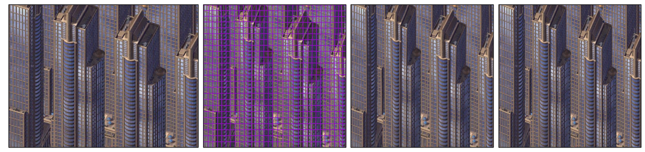



Fig. 15. Retargeting result for the image containing objects with similar significance values. (a) The deformed quad mesh; (b) our result; (c) the linear scaling.

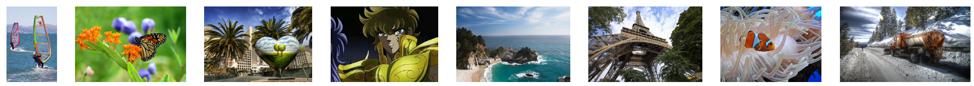
## V. Conclusions, Limitation, and Future Work

A novel patch-based warping was presented for image retargeting. In our approach, the energy term of similarity transformation can force an as-rigid-as-possible deformation on the visually salient content and the energy term of linear scaling can avoid over-deformation. Moreover, the optimization process with the estimated significance values can smoothly propagate distortions. These processes can ease unwanted deformations caused by inconsistent warping, 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 structure line preservation. Nevertheless, our approach has the following limitation. Our approach cannot work well for images that contain objects with similar significant values. The retargeting results in this case are similar to those generated by linear scaling, as shown in Figure 15, 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.

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 $A_1$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $O_i$ | SC [1] | 3 | 6 | 1 | 6 | 6 | 18 | 12 | 9 |
| | OSS [9] | 41 | 14 | 5 | 20 | 21 | 13 | 24 | 41 |
| | SP [10] | 1 | 11 | 5 | 9 | 26 | 27 | 32 | 7 |
| | L.S. | 14 | 23 | 9 | 16 | 11 | 42 | 27 | 29 |
| $E_i$ | – | 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 $A_2$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $O_i$ | SC | 16 | 40 | 6 | 41 | 26 | 31 | 28 | 4 |
| | OSS [9] | 25 | 18 | 23 | 34 | 25 | 25 | 46 | 41 |
| | SP [10] | 21 | 34 | 46 | 20 | 25 | 40 | 51 | 39 |
| | L.S. | 7 | 11 | 14 | 9 | 34 | 37 | 8 | 13 |
| $E_i$ | – | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 38 |

TABLE III
THE VALUES OF $\chi_0^2$ AND $\chi_{0.05,7}^2$, AND THE RESULTS OF GOODNESS-OF-FIT TEST.

| | | SC [1] | OSS [9] | SP [10] | L.S. |
|---|---|---|---|---|---|
| $\chi_0^2$ | $A_1$ | 199.55263 | 82.026314 | 139.21053 | 106.93420 |
| | $A_2$ | 78.157898 | 32.131580 | 27.263159 | 146.67105 |
| $\chi_{0.05,7}^2$ | – | 14.07 | 14.07 | 14.07 | 14.07 |
| Acceptance /Rejection ($A_1$/ $A_2$) | | r/r | r/r | r/r | r/r |

## REFERENCES


[1] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM Trans. Graph.*, vol. 26, no. 3, July 2007.

[2] 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, August 2008.

[3] A. Shamir and S. Avidan, "Seam carving for media retargeting," *Commun. ACM*, vol. 52, no. 1, pp. 77–85, January 2009.

[4] 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, June 2010.

[5] M. Frankovich and A. Wong, "Enhanced seam carving via integration of energy gradient functionals," *Signal Processing Letters*, vol. 18, no. 6, pp. 375–378, 2011.

[6] F. Liu and M. Gleicher, "Automatic image retargeting with fisheye-view warping," in *Proceedings of the 18th annual ACM symposium on User interface software and technology*, 2005, pp. 153–162.

[7] R. Gal, O. Sorkine, and D. Cohen-Or, "Feature-aware texturing," in *Proceedings of Eurographics Symposium on Rendering*, 2006, pp. 297–303.

[8] L. Wolf, M. Guttmann, and D. Cohen-Or, "Non-homogeneous content-driven video-retargeting," in *Proceedings of the Eleventh IEEE International Conference on Computer Vision (ICCV-07)*, 2007, pp. 1–6.

[9] 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, December 2008.

[10] 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.

[11] Y. Guo, F. Liu, J. Shi, Z.-H. Zhou, and M. Gleicher, "Image retargeting using mesh parametrization," *IEEE Trans. Multi.*, vol. 11, no. 5, pp. 856–867, August 2009.

[12] Y. Jin, L. Liu, and Q. Wu, "Nonhomogeneous scaling optimization for realtime image resizing," *Vis. Comput.*, vol. 26, no. 6-8, pp. 769–778, June 2010.

[13] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," in *IEEE CVPR*, 2010, pp. 2376–2383.

[14] A. Shamir and O. Sorkine, "Visual media retargeting," in *ACM SIGGRAPH ASIA 2009 Courses*, 2009, pp. 11:1–11:13.

[15] B. Suh, H. Ling, B. B. Bederson, and D. W. Jacobs, "Automatic thumbnail cropping and its effectiveness," in *Proceedings of the 16th annual ACM symposium on User interface software and technology*, 2003, pp. 95–104.

[16] 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.

[17] F. Liu and M. Gleicher, "Video retargeting: automating pan and scan," in *Proceedings of the 14th annual ACM international conference on Multimedia*, 2006, pp. 241–250.

[18] A. Santella, M. Agrawala, D. DeCarlo, D. Salesin, and M. Cohen, "Gaze-based interaction for semi-automatic photo cropping," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 2006, pp. 771–780.

[19] C. Tao, J. Jia, and H. Sun, "Active window oriented dynamic video retargeting," in *Proceedings of the Workshop on Dynamical Vision*, 2007, pp. 1–12.

[20] M. Rubinstein, A. Shamir, and S. Avidan, "Multi-operator media retargeting," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 23:1–23:11, July 2009.

[21] Y. Pritch, E. Kav-Venaki, and S. Peleg, "Shift-map image editing," in *IEEE ICCV'09*, Sept 2009, pp. 151–158.

[22] Q. xing 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.

[23] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vision*, vol. 59, no. 2, pp. 167–181, September 2004.

[24] P. Shirley, M. Ashikhmin, M. Gleicher, S. Marschner, E. Reinhard, K. Sung, W. Thompson, and P. Willemsen, *Fundamentals of Computer Graphics, Second Ed.* A. K. Peters, Ltd., 2005.

[25] 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, July 2010.

[26] 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, December 2010.