Computer Graphics International 2012

Tearable Image Warping for Extreme Image Retargeting

Lai-Kuan Wong · Kok-Lim Low

Abstract We introduce a new image geometric transformation, named *tearable image warping*, for content-aware image retargeting and recomposition. In tearable image warping, an object is allowed to *partially detached* from the background during warping, which allows the background to warp more freely without distorting the object. The part of the object that is still attached to the background ensures that spatial semantics between the object and its environment is preserved. Compared to traditional non-homogeneous image warping, tearable warping produces retargeted results with less severe distortions, especially in cases of extreme retargeting. Compared to scene carving, tearable warping can better preserve the global context and achieve better scene consistency by preserving the semantic connectedness between an object and its environment. In addition, tearable warping is also able to alter the object-background relationship within the image, which makes it quite suitable for aesthetics-based image recomposition. We also show that tearable image warping is a unified approach of traditional non-homogeneous image warping and the cut-and-paste techniques.

Keywords tearable image warping · image retargeting · image recomposition · interactive digital photo editing

1 Introduction

Digital image warping has many and diverse applications, including image distortion correction, image registration, and image morphing. In the recent years, an area of application that has received much research focus is image retargeting. Due to the ubiquitous use of mobile devices and wide variety of display devices, there is a need to retarget images to different resolutions and aspect ratios. Many researchers [5, 6, 7, 21] have attempted content-aware retargeting by using non-homogeneous warping such that unimportant regions are allowed to distort more and important regions are preserved as uniformly as possible. Different constraints and optimization methods are used to warp images. Some researchers [8, 9] have also attempted to enhance aesthetics of the retargeted images by modifying the composition based on selected photographic rules.

Compared to other competing image retargeting approaches, such as cropping [16, 17, 22], seam carving [1, 14] and cut-and-paste [18], image warping has the advantage of retaining all the content and image topology (homeomorphism) of the original image, and thus is better at preserving the original spatial semantics among objects in the image. However, maintaining the original image topology can limit the flexibility of the warping, where non-salient regions have to be severely distorted in order to minimize distortion in the salient regions.

We observed that in general, image topology does not need to be preserved everywhere to maintain semantics correctness. For example, the connectedness between an object's boundary and its adjacent background can be disregarded if it does not correspond to actual physical contact in the 3D world. On the other hand, the connectedness with the part of the environment where it comes into actual physical contact (e.g. the ground) should be preserved. In this paper, we introduce a new image warping method, named *tearable image warping*, that capitalizes on this idea for scene-consistent image retargeting. In addition to the two scene consistency properties defined by Mansfield et al. [10]—zero object distortion and correct scene occlusion—our approach is able to further enhance scene consistency by ensuring that objects maintain consistent physical contacts with their environment in the retargeted image.

In tearable image warping, we divide each selected object's boundary into *tearable* and *non-tearable* segments. Normally, the tearable segments correspond to where depth discontinuity occurs, and non-tearable segments to parts of

Lai-Kuan Wong
Department of Computer Science
National University of Singapore
E-mail: lkwong@comp.nus.edu.sg

Kok-Lim Low
Department of Computer Science
National University of Singapore
E-mail: lowkl@comp.nus.edu.sg

the object boundary that have actual physical contacts with the environment or other objects. Conceptually, during warping, we allow the object's boundary to tear along the tearable segments. This allows the background to partially break away from the object and be warped more independently, which often can distribute warping more evenly to avoid local distortion. Meanwhile, the object is kept undistorted and the non-tearable segments help to preserve image semantics by constraining the object to maintain the real contacts in the 3D world. Any hole left behind after the warping is automatically inpainted.

Our target application is interactive image retargeting and recomposition. Our system provides a semi-automatic interface to allow users to easily segment the important objects, and to approximately mark each non-tearable segment (using a polyline that we call *object handle*). The system then automatically retargets the image to the desired aspect ratio, and can also allow the user to interactively recompose the retargeted image, both using tearable image warping.

Results show that our proposed tearable warping algorithm in general produces less distortions than the traditional non-homogeneous warping methods [7]. Compared to scene carving [10], our approach can better preserve scene consistency by maintaining the desired connectedness between objects and background. In summary, the main contributions of our work include

- the concept of tearable/non-tearable object boundary, which leads to more flexible warping without sacrificing image semantics preservation;
- a novel retargeting method that can preserve all three scene consistency properties—object protection, correct depth order, and semantic connectedness—simultaneously in extreme retargeting cases; and
- a practical algorithm to implement our tearable image warping idea for image retargeting and recomposition.

2 Related Work

There are three popular classes of content-aware image retargeting operators: cropping, seam-carving, and warping. Each of these approaches utilizes some region-of-interest (ROI) extraction methods, such as saliency detection or gradient energy, to compute an importance map and tries to minimize distortion in the important regions. None of these operators can completely solve the image retargeting problem [20]. Each operator has its own advantages and disadvantages for different applications.

Basically, content-aware cropping methods [16, 17] search for the best cropping window that contains all the important objects. Some methods [12, 22] try to incorporate aesthetics measures into its cropping optimization to enhance the aesthetics of the retargeted images. Cropping is artifact- and distortion-free but it may destroy the global context and is highly inflexible. Seam carving [1, 14] is an elegant approach that removes the least important vertical or horizontal seams, measured by a computed importance map. Although seam carving can produce some good and interesting results, it is only effective for images with sufficient homogenous regions. For more complex images, results are prone to distortions and artifacts when the seams cut through high-frequency content [20].

Various variants of the image warping approach have been proposed for image retargeting. The more recent warping methods represent an input image as a quad mesh [21] or a triangle mesh [5, 7] and perform optimization to find the new locations for the mesh control points, keeping the ROIs as rigid as possible. Jin et al. [7] went a step further to enhance image aesthetics of the retargeted images by including a set of selected photographic rules in the warping optimization. However, due to the non-changeable object-background relationship, there is limited flexibility to enhance image aesthetics using warping.

More recent approaches use a combination of multiple operators. The challenge is to find the best way to combine them. A few researchers [3, 15] have attempted to combine seam carving with other retargeting operators such as scaling and cropping. These methods may reduce distortions caused by pure seam carving but in cases of extreme resizing, severe distortions are largely unavoidable. Liu et al. [8] combined cropping and warping to retarget and optimize photo composition simultaneously to produce some interesting results. However, the algorithm is slow due to its search in a 4D space and it is unable to preserve global context.

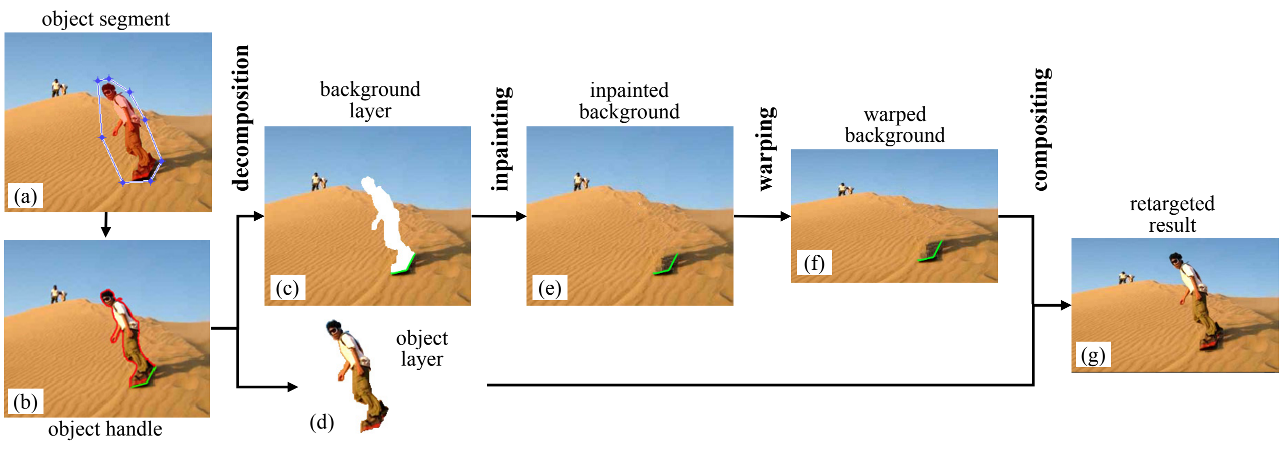

Fig. 1 Steps in the tearable image warping algorithm. In (b), the object boundary is shown in red and the object handle in green.

Notably, only a few retargeting methods can protect objects to avoid unpleasant distortion of objects in extreme retargeting cases. Setlur et al. [18] and Mansfield et al. [10] have presented retargeting methods that use a model decomposition approach to protect objects. Setlur et al. [18] proposed a non-photorealistic retargeting method that identifies the ROIs, removes them, inpaints the holes left by the ROIs, resizes the inpainted image and re-inserts the ROIs. Relatively, this approach is highly flexible but is not popular due to its dependence on inpainting and its susceptibility to semantics violation. It also has no flexibility in repositioning the objects. Mansfield et al. [10] proposed scene carving, a layered approach to seam carving to prevent seam from eliminating important objects. Given an image with object segments and their respective depth order, the scene carving algorithm first decomposes the image into layers, followed by finding the best seams to remove from the background layer and positioning the objects so that their visibility is maximized. This method successfully reduces disturbing distortion to objects and can be used for extreme retargeting by allowing objects to overlap in the correct depth order. However, scene consistency may still be violated because this approach does not guarantee consistency of semantic connectedness between an object and its environment, especially when shadows or reflections exist. In addition, this approach is still prone to visual distortion inherent to the seam carving method.

3 Scene Consistency

Mansfield et al. [10] defined that a retargeted image is scene consistent if objects (1) are not distorted but kept as in the original image, and (2) are placed in their correct depth ordering. We find these two properties insufficient to ensure scene consistency, and the definition could be made more useful by including a third property: (3) objects must maintain consistent physical contacts with their environment. This is especially important, for example, if the object causes shadows, reflections or ripples in the environment.

To achieve zero object distortion and scene consistent occlusion, we adopt Mansfield's model decomposition approach, whereby an image can be described by a relative depth map comprising of *object segments* and their relative *depth order*. To support the third property of scene consistency, we divide the boundary of each object into tearable and non-tearable segments. Tearable segments represent boundary sections where depth discontinuity occurs and non-tearable segments correspond to boundary sections that have physical contact with environment and must therefore be preserved.

We use an *object handle* to represent the non-tearable segment of an object. An object handle is a polyline drawn by the user to specify a part of an object's boundary that is non-tearable. In our implementation, an object handle marks a local area that has to be kept rigid, therefore the polyline does not need to be as precise as the object boundary. In general, an object handle can be anywhere in the object segment, and need not even be near the object boundary. For good retargeting results, an object handle must satisfy two criteria:

- an object handle must be preserved as rigid as possible, ideally without rotation; and
- the object must be able to be re-inserted to coincide precisely with its handle at a new location.

Each object can have multiple object handles as shown in the examples in Figure 5. Multiple handles of an object are combined as one object handle in the warping process.

4 Algorithm

Given an image with a set of object segments and their respective object handles (for specifying non-tearable segments), our tearable image warping algorithm performs the following three main steps: (1) decomposition, (2) warping and (3) image compositing.

Figure 1 shows an overview of the algorithm steps and the intermediate inputs/outputs. In the decomposition step, the image is first decomposed into a background layer and potentially multiple object layers. Holes left by the objects in the background layer are automatically inpainted. In the warping step, traditional non-homogeneous warping is applied only to the background layer, always keeping the object handles as rigid as possible. In the image compositing step, the objects are pasted back to the warped background based on the new positions of their respective handles. Warping the background without the need to protect the entire objects gives more room to distribute the distortion more uniformly. The rigidity of the handles ensures that the objects can be seamlessly restored back to their original contact locations with the background.

4.1 Step 1: Decomposition

The decomposition step uses the object segmentations supplied by the user to decompose the input image into a background layer and one or more object layers. A feathered mask is created for each object layer so that the object's boundary can be blended smoothly with the background when the object is re-inserted back to form the final image in the image compositing stage. Holes left by cut-out objects in the background layer will be automatically inpainted using exemplar-based inpainting methods [2, 11].

4.2 Step 2: Warping

The choice of a suitable base image warping method is important. We chose a triangle-mesh-based warping method for its efficiency and ease of representation. In a triangle mesh, an object handle can be easily made into edges of the triangle mesh using constrained Delaunay triangulation. Figure 2 shows an example initial triangle mesh for an input image, in which the green edges represent an object handle. In principle, any befitting mesh parameterization [5, 7] can be used to find the destination locations for the triangle nodes. We adapted the non-homogenous scaling optimization method by Jin et al. [7] to warp the background image, but without the saliency constraints and the weights to preserve salient content.

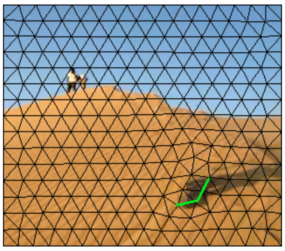

Fig. 2 A triangle mesh used for warping. The object handle is part of the mesh (highlighted in green).

Instead, we apply our handle shape constraint as a hard constraint to preserve the shapes and orientations of object handles. This method constrains the transformation of triangles to scaling and translation, without rotation, and is thus an ideal method to preserve object handles.

Given a source mesh M for the input background image and the object handles, the warping process is the problem of mapping M to a target mesh M' that still keeps all the objects handles rigid. For image retargeting, M' must fit the target image aspect ratio. For image recomposition, the object handles in M' must be positioned at their new target locations. After we have computed M', the new warped background image is obtained by an inverse piecewise affine mapping for each triangle in the mesh. The following subsections describe the warping energy and constraints that are used in the computation of the target mesh M'.

4.2.1 Warping Energy

Scale transformation error. For each triangle $t \in M$, we constrain the transformation to non-uniform scaling [7], denoted by $G_t = \begin{pmatrix} s_t^x & 0 \\ 0 & s_t^y \end{pmatrix}$. However, in general, for each triangle in M, there is an affine mapping that maps it to its corresponding triangle in M' and the linear portion of the affine mapping can be represented by a 2×2 Jacobian matrix J_t. The scale transformation error is defined as

$$E_w = \sum_{t \in T} A_t \|J_t - G_t\|_F^2 \tag{1}$$

where A_t is the area of triangle t in M' and $\|\cdot\|_F^2$ is the Frobenius norm.

Smoothness error. To avoid discontinuity in the resulting image, we enforce a smoothness term that tries to minimize the scale difference between neighboring triangles [7]:

$$E_s = \sum_{\substack{s,t \in M \\ s \text{ and } t \text{ are adjacent}}} A_{st} \|G_t - G_s\|_F^2 \tag{2}$$

where $A_{st} = (A_s + A_t)/2$.

Total error. The total warping energy is defined as the weighted sum of the scale transformation and smoothness errors:

$$E = \alpha E_w + \beta E_s \tag{3}$$

where α and β are the weights. Minimizing the total error function will try to constrain the warping of all triangles to non-uniform scaling, without rotation.

4.2.2 Handle Shape Constraint

To ensure that an object can be re-inserted seamlessly to its object handle in the warped background, we must preserve the shape and orientation of the handle during the warping process. Here, we assume that all object handles of an object has been combined into one. Suppose the object handle consists of n vertices, $v_1, v_2, \cdots, v_n$, in M, and they are being mapped to vertices $u_1, u_2, \cdots, u_n$ in M'.

To preserve the shape of an object handle, for each vertex v_i, we preserve two distance measures: (1) distance between v_i and v_1, and (2) distance between v_i and v_{i+1}. To preserve the orientation of the object handle, each of the above distances is computed as a signed x-distance and a signed y-distance separately. Consequently, for each object k, the handle shape constraints are

$$\begin{aligned} u_{i,x} - u_{1,x} &= s_k(v_{i,x} - v_{1,x}) \\ u_{i,y} - u_{1,y} &= s_k(v_{i,y} - v_{1,y}) \end{aligned} \tag{4}$$

and

$$\begin{aligned} u_{j+1,x} - u_{j,x} &= s_k(v_{j+1,x} - v_{j,x}) \\ u_{j+1,y} - u_{j,y} &= s_k(v_{j+1,y} - v_{j,y}) \end{aligned} \tag{5}$$

where $i = 2, \cdots, n$, $j = 2, \cdots, n-1$, and s_k is a scale factor that allows the object handle (and the object) to undergo uniform scaling. The scale factor is specified by the user or set to 1 by default. Each object can have a different scale factor.

4.2.3 Boundary Positional Constraint

To keep the problem well-posed, we constrain the boundary vertices in the input mesh M to the boundary of the output mesh M'. For each vertex v on the left, right, top or bottom border of M', we apply the positional constraints $v_x = 0$, $v_x = W$, $v_y = 0$, and $v_y = H$, respectively, where W and H are the width and height of the output image, respectively.

4.2.4 Retargeting-Specific Constraints

Considering that image retargeting is an automatic process and objects are not warped together with the background, we chose to be more conservative in modifying the image. Therefore, we apply two additional hard constraints: (1) *object boundary constraint* to ensure that objects are not cropped off, and (2) *non-overlap constraint* to ensure that objects that do not overlap in the input image will not overlap in the output image. However, if a depth order is given, our algorithm can relax the non-overlap constraint to allow occlusion of objects by re-inserting the objects in the final image compositing step based on the depth order.

Object boundary constraint. We first compute an axis-aligned bounding box around the object. During optimization, we compute the new position of the bounding box based on the new position of the object handle. To

Fig. 3 Image retargeting with and without the non-overlap constraint. (a) Input image, (b) retargeted with non-overlap constraint, (c)–(e) retargeted without the non-overlap constraint.

enforce this constraint, we disallow any part of the object's bounding box to move outside the target image region.

Non-overlap constraint. This is similar to the object boundary constraint, but we keep track of the bounding boxes of all objects and enforce that they do not overlap. Using axis-aligned bounding boxes for overlap testing is efficient but not accurate, since it may give false hits when the objects are actually not overlapping. For higher accuracy, complex polygons or hierarchical bounding boxes can be used. Figure 3 compares some results of tearable image warping with and without the non-overlap constraint.

4.2.5 Recomposition-Specific Constraints

Handle positional constraint. For recomposition, the user can interactively move any selected object, and even move it to occlude other objects. Due to the interactive nature, object boundary and non-overlap constraints are not necessary. The only required constraint is the handle positional constraint. We set the position of the selected object based on the user's mouse movement and fix the other objects at their current positions.

Suppose each object's handle consists of n vertices, $v_1, v_2, \cdots, v_n$, in M, and they are being mapped to vertices $u_1, u_2, \cdots, u_n$ in M'. Let $d = (d_x, d_y)^{\mathrm{T}}$ be the translation computed from the mouse movement, the handle positional constraint for the selected object is

$$u_1 = v_1 + d, \tag{6}$$

and for every of the unselected objects, the constraint is

$$u_1 = v_1. \tag{7}$$

Interactive background warping. For recomposition, in addition to moving the objects around, we allow the user to interactively warp the background by simply dragging any part of the background. This extra feature allows the user to have more control over the composition of the output image. For example, the user can move the horizon to a desired location by simply dragging it to a new position as shown in the example in Figure 4.

The method to enable interactive background warping is similar to that of relocating the objects in recomposition. In this case, the background is warped in the direction of the mouse movement while keeping all the object handles fixed at their current locations. More specifically, we first find the mesh vertex nearest to the mouse click position, and enforce the hard constraint that it must be repositioned at the new mouse location.

4.3 Step 3: Image Compositing

The warped background is combined with all the object layers to form the final output image in the image compositing step. Each cut-out object image is first scaled by its respective scale factor s_k before it is re-inserted onto the warped background at its new object handle location. If object overlap is allowed, we re-insert the object according to the given depth order. For seamless blending of the objects into the warped background, we use feathered masks of the objects for the compositing.

5 Implementation

We provide the implementation details on the user input and interaction interface, and the optimization of the warping process.

5.1 User Input and Interaction

We use GrabCut [13] to allow the user to segment the objects easily. Very often, the user only needs to draw a polygon around the object with a few clicks. Similarly, the users can specify a polyline to define an object handle.

For image retargeting, the user is prompted to provide the target scale factors s_x and s_y for the x and y directions respectively. If occlusion is toggled on, the user can choose to use the default depth order derived from the order of object segmentation or to provide the depth order for the objects. To recompose a retargeted image, the user can click on any object and drag it within the retargeted image.

Fig. 4 Result of interactive background warping, where the horizon has been moved but the object's position fixed.

5.2 Optimization Details

Note that the energy function in Eq. (3) is a convex quadratic function. The energy weights α and β are set to 1 and 0.5 respectively. The handle shape and boundary positional constraints are both set as hard constraints. For interactive image recomposition, the additional handle positional constraints are treated as hard constraints. Thus, we can formulate recomposition as the problem of minimizing Eq. (3) with a set of equality constraints, for which the solution can be obtained in real-time by solving a sparse linear system. On the other hand, for image retargeting, we derived two sets of inequality constraints to prevent objects from being cropped off and overlapping each other. We utilize the CVX Matlab toolbox [4] to find the solution to the quadratic programming problem.

6 Results

Results in this paper were generated on a laptop with Intel Core2 Duo CPU 2.53GHz and 4GB memory. Inpainting of the background image is performed in the pre-processing stage. It takes less than 10s to inpaint the area covered by the person in the leftmost image of size 664x1024 in Figure 5 using both the CPU-based [19] and GPU accelerated inpainting [11] methods. Excluding the time taken for inpainting, our tearable image warping algorithm produces a retargeted result in about 2s to 4s for an image of resolution 1024x768. This is a significant speed up compared to scene carving [10], which takes almost 27 mins on the same computer to retarget the same image to 50% of its original size. The speed of our algorithm depends mainly on the number of the triangle meshes per image, which we have kept almost constant even for different image sizes. In contrast, the speed of scene carving decreases with increasing image size and increasing difference between input and output image sizes. Upon obtaining the retargeted result, our system allows the user to interactively modify the composition of the image.

We compare our tearable warping method with a traditional warping method based on [7] and with scene carving [10]. For fair comparison, we use the same manually segmented objects as ROIs for all methods. For the comparisons with traditional warping, we set the scale factor of all objects to $\max(s_x, s_y)$. To minimize object distortion for traditional warping, we apply hard constraint to preserve the salient triangles representing the objects. However, in extreme retargeting cases where no solution can be obtained for traditional warping, we relax the salient triangles preservation as soft constraint. For comparison with scene carving that allows objects to be cropped, we relax the object boundary constraint of tearable warping to allow objects to be cropped.

Fig. 5 (Top row) Input images with object handles shown in red, (second row) results of tearable image warping, (third row) results of traditional warping [7], and (bottom row) results of scene carving [10]. Red rectangles highlight background distortion and blue rectangles highlight object distortion.

6.1 Image Retargeting

Figure 5 compares the retargeted results of tearable image warping with traditional image warping [7] and with scene carving [10]. These examples are considered somewhat extreme retargeting because of their image contents and the target aspect ratios. To achieve the target aspect ratios, severe distortions to background, caused by over-compression often occur in the traditional warping method, as illustrated in its results in Figure 5 (third row). Results of traditional warping in column two and three of Figure 5 also exhibit obvious distortion to the main objects, as highlighted by the blue rectangles. In comparison, tearable image warping consistently distorts the background less in cases of extreme retargeting, as shown in its results in Figure 5 (second row). The retargeted triangle meshes shown in Figure 6 give some insights to how tearable warping can reduce the over-compression problem. In traditional warping, all triangles representing the object need to be preserved uniformly, leaving little room for distributing the compression. In tearable image warping, only edges representing the object handles need to be preserved, therefore compression can be distributed more evenly throughout the image, including areas "behind" the objects.

On the other hand, scene carving has little distortion in images with large homogenous regions, but for structurally complex images, severe distortions may occur, as shown in its results in the last column of Figure 5. Noticeably, scene carving may result in cropping, thus potentially destroying the global context. In contrast, tearable warping can preserve the global context much better. A unique feature of tearable warping is the effect of objects shifting with respect to their immediate background. This effect can be observed in many of our results in Figure 5. Interestingly,

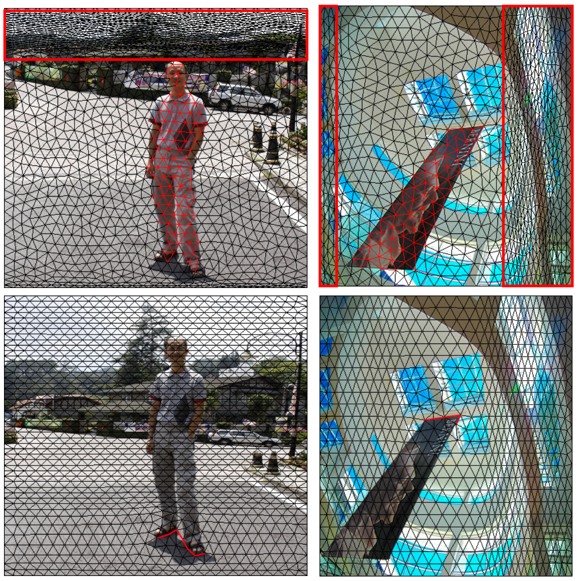

Fig. 6 (Top row) Resulting triangle meshes of traditional image warping; (bottom row) of tearable image warping.

this change of object-background relationship often produces a natural and semantics-preserving effect analogous to a shift of the viewpoint. This feature potentially is a powerful technique for aesthetics enhancement of images.

In terms of scene consistency, object distortion can never occur in tearable warping and scene carving because objects are not involved in the warping and seam carving process respectively. Furthermore, tearable warping and scene carving allow objects to overlap while maintaining the correct depth order, which gives more flexibility to perform extreme image retargeting without object distortion, as illustrated in the examples in Figures 7 and 8.

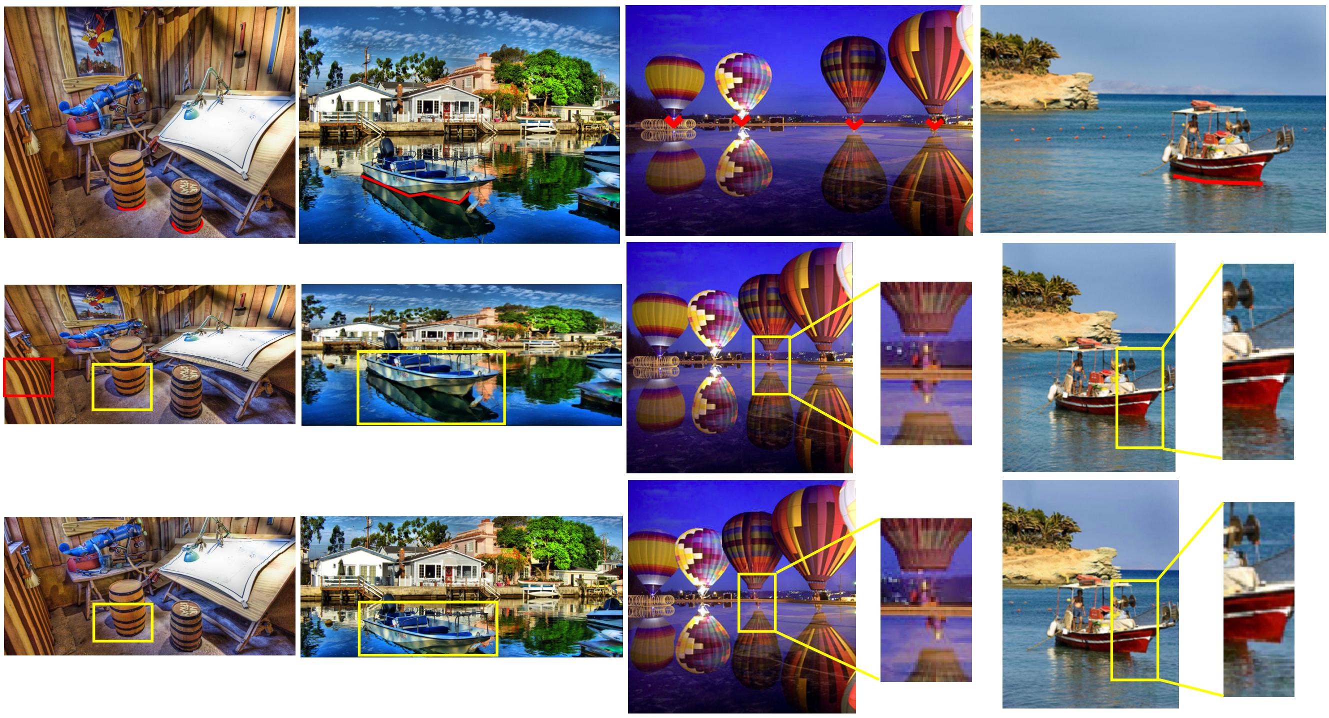

Fig. 7 (Top row) Input images, with object handles shown in red, (middle row) results of tearable warping, and (bottom row) results of scene carving. Yellow rectangles highlight and compare the results of both methods in maintaining consistent physical contact between objects and their environments. The red rectangle highlights feature distortion.

Fig. 8 Results with object occlusion. (Top) Input images with object handles shown in red, (bottom left) results of tearable warping, and (bottom right) results of scene carving.

In addition to its ability to better preserve global context, another key advantage of tearable warping over scene carving is its capability to maintain consistent connectedness of objects with their environments. Figure 7 shows examples in which scene carving fails in this aspect, particularly when shadows, reflections or ripples exist, leading to unpleasant image artifacts. In contrast, tearable warping is able to maintain consistent physical contacts.

Tearable image warping maintains consistent semantic connectedness of an object by keeping the object handle rigid. To keep an object handle rigid, the handle must be defined by at least three non-collinear points. However, in cases where there are many cluttered objects, keeping all the object handles rigid will restrict overlapping of objects and thus forbid extreme retargeting. In such case, the handle can be relaxed by specifying only two points. With only two points, the handle will not be preserved rigidly but the relative positions between objects will still be maintained and more overlapping is thus allowed in extreme retargeting. This is demonstrated in the middle example in Figure 8 and the penguin example in Figure 9. Relaxed handles can also be used for cases where objects are not in physical contact with their environment, so that their relative positions could still be preserved. The dancers example in Figure 8 is one such cases (the handles are on the floor).

A potential problem with both scene carving and tearable warping is that, at extreme retargeting ratio, unpleasant hole artifacts can happen due to the dis-occlusion of objects, as shown in Figure 9(c) and 9(d). Despite the use of an energy function to minimize holes in scene carving, this hole problem is still unavoidable in extreme retargeting. For tearable image warping, this problem can be avoided by creatively specifying the object handles. For example, by changing the handle of the penguin blocking the occluded penguin to a rigid handle, no more hole (or inpainting artifact) is visible in the retargeted image as demonstrated in Figure 9(e).

Another creative use of handles is to select only part of a physical object for protection and use the handle to ensure that the selected part of the object is later combined seamlessly with the non-selected part in the retargeted image. This can be particularly useful for images where the entire object fills most of the image frame like the example of the Eiffel Tower in Figure 10. Scene carving does not have this flexibility as it does not guarantee the semantic connectedness between different parts of the object.

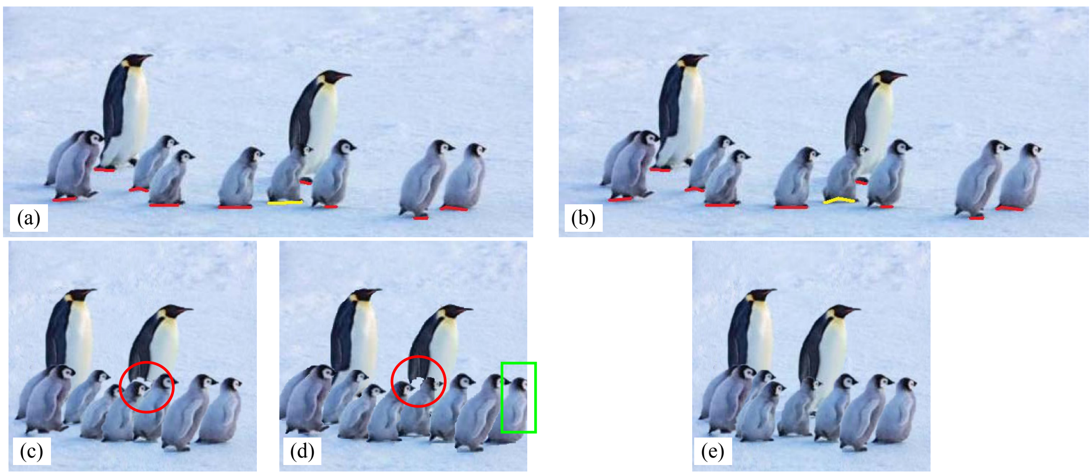

Fig. 9 (a)–(b) Input images with object handles, except that the yellow handle is non-rigid in (a) and rigid in (b). (c)–(d) Results of tearable warping (using handles in (a)) and scene carving, where both show hole artifacts (in red circles). (e) Results of tearable warping (using handles in (b)), where the yellow rigid handle has prevented the hole from showing up in the retargeted image. The green rectangle highlights unpleasant cropping of object in scene carving.

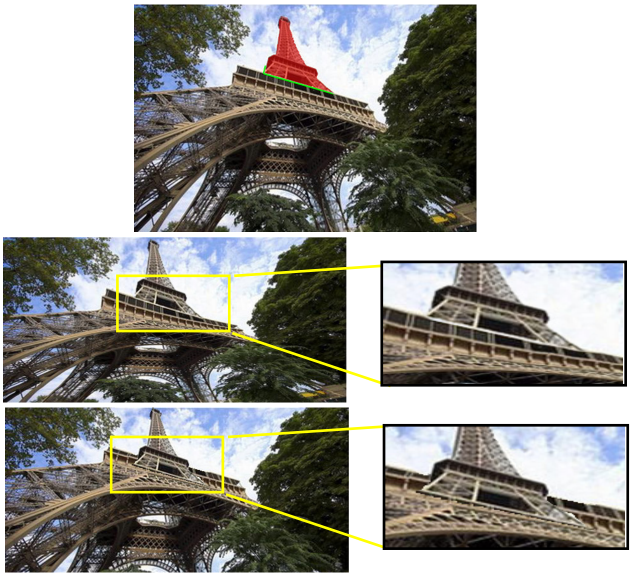

Fig. 10 Creative use of object handles. (Top row) Input image with only the top part of Eiffel Tower selected for protection (handle is shown in green), (middle row) result of tearable warping, and (bottom row) result of scene carving.

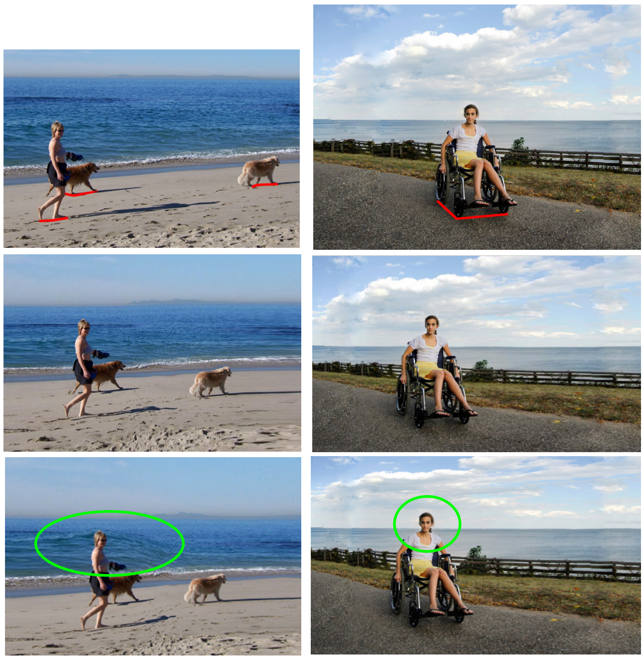

Fig. 11 Recomposition results. (Top row) Input Images, (middle row) results of tearable image warping, and (bottom row) results of traditional image warping.

6.2 Image Recomposition

The result of traditional warping in the left column of Figure 11 shows severe distortion above the lady's head. No similar distortion is detected in the corresponding result of tearable warping. In addition, the inability to allow changes to object-background relationship has limited the flexibility of traditional warping for image aesthetics enhancement. For example, the unpleasant effect of having the horizon cutting through the lady's neck in Figure 11 cannot be changed by traditional warping. With tearable warping, users can move objects or background to avoid merger or to effect a view change to make the subject more visually dominant. Figure 12 shows more examples of recomposition using tearable warping.

7 Discussion

The tearable image warping approach is not limited to triangle-based warping but can be applied to any other warping methods, such as quad-based and pixel-based image warping methods, as long as there is a way to preserve the object handles. In addition, tearable image warping can be used for completely tearable or completely non-tearable object boundary. Completely tearable objects, such as birds flying in the sky, have *no object handle* and its resulting tearable warping is equivalent to pure cut-and-paste. Vice versa, completely non-tearable objects, such as windows, can be defined with *full object handle* and its resulting tearable warping is equivalent to pure image warping. In short, tearable image warping is a unified approach that can smoothly transition from pure cut-and-paste to pure image warping.

One drawback of the tearable image warping approach is that it requires inpainting, which is still an open problem in computer vision. Fortunately, this drawback is not that critical to tearable image warping, as compared to pure cut-and-paste approaches. We find that good inpainting is seldom required for tearable image warping, especially for image retargeting. Artifacts of inpainting often occur near the object handle or the object's center, and these areas are most likely still covered by the object in the retargeted or recomposed image. Furthermore, the holes are usually compressed with the background image in the retargeting

Fig. 12 More results of tearable image warping. (Top row) Input Images, (middle row) retargeted images, and (bottom row) recomposed results of retargeted images.

process, making it even less likely to show up. As illustrated in Figure 13, although the inpainted background image is far from perfect, none of these artifacts are actually exposed in the result. However, for rare cases where inpainting artifacts is visible, we allow users to interactively touch up the inpainting artifacts.

In our algorithm, we did not apply any feature preservation to preserve prominent lines and curves because we found that our results seldom have significant feature distortion, due to the fact that warping is spread out more uniformly throughout the image. However, for images with very complex structural details like the art room example in column one of Figure 7, distortion is hard to avoid in extreme retargeting. In these cases, we can add an optional line preservation constraint to ensure that straight lines are not distorted.

Fig. 13 (Left) Retargeted and recomposed background layer. Red outline shows the inpainted hole, which is compressed in the retargeting process. (Right) Result with the object re-inserted. The red areas show that only small part of the inpainted hole is visible in the final image.

8 Conclusion

We have introduced tearable image warping, a new approach that unifies image warping and cut-and-paste techniques, for content-aware image retargeting and recomposition. The key concept of tearable warping to allow an object to be partially detached from its original background makes several noteworthy contributions. Firstly, it significantly reduces distortion inherent to traditional warping in extreme image retargeting. Secondly, it can achieve better scene consistency by simultaneously protecting objects, ensuring correct depth order of objects and maintaining consistent semantics connectedness between objects and their environment. Lastly, it allows the object-background relationship to be changed, which is a powerful feature for aesthetics enhancement.

Acknowledgments

This work is supported by the Singapore MOE Academic Research Fund (Project Number: T1 251RES0804).

References

1. Avidan, S., Shamir, A.: Seam carving for content-aware image resizing. ACM Trans. on Graphics, 26(3), (2007).
2. Criminisi, A., Perez, P., Toyama, K.: Object removal by exemplar-based inpainting. IEEE Trans. on Image Processing, 13(9), 1200-121 (2004).
3. Dong, W., Zhou, N., Paul, J.-C., Zhang, X.: Optimized image resizing using seam carving and scaling. ACM Trans. on Graphics, 28(5), 1-10 (2009).
4. Grant, M., Boyd, S., Ye, Y.: CVX: Matlab software for disciplined convex programming. http://www.stanford.edu/~boyd/cvx.
5. Guo, Y., Liu, F., Shi, J., Zhou, Z,-H., and Gleicher, M.: Image retargeting using mesh parametrization. IEEE Trans. on Multimedia, 11(5), 856-867 (2009).
6. Gal, R., Sorkine, O., Cohen-Or, D.: Feature-aware texturing. Proceedings of Eurographics Symp. on Rendering, pp. 297-303 (2006).
7. Jin, Y., Liu, L., Wu, Q.: Non-homogeneous Scaling Optimization for Realtime Image Resizing. The Visual Computer, 26(6-8), 769-778 (2010).
8. Liu, L., Chen, R., Wolf, L., and Cohen-Or, D.: Optimizing photo composition. Computer Graphics Forum, 29(2), 469-478 (2010).
9. Liu, L., Jin, Y., Wu, Q.: Realtime aesthetic image retargeting. Proceedings of Eurographics Workshop on Computational Aesthetic in Graphics, Visualization, and Imaging, pp. 1-8 (2010).
10. Mansfield, A., Gehler, P., Van Gool, L., Rother, C.: Scene carving: Scene consistent image retargeting. Proceedings of European Conference on Computer Vision (ECCV), pp. 143-156 (2010).
11. Yousef, M., Hussien, K.F.: ParXII: Optimized, data-parallel exemplar-based image inpainting. Proceedings of International Conference on Computer Graphics and Interactive Techniques (ACM SIGGRAPH), Poster Proceedings. ISBN 978-1-4503-0971-4 (2011).
12. Nishiyama, M., Okabe, T., Sato, Y., Sato, I.: Sensation-based photo cropping. Proceedings of ACM International Conference on Multimedia, pp. 669-672 (2009).
13. Rother, C., Kolmogorov, V., Blake, A.: "GrabCut" — Interactive foreground extraction using iterated graph cuts. ACM Trans. on Graphics, 23(3), 309–314 (2004).
14. Rubinstein, M., Shamir, A., Avidan, S.: Improved seam carving for video retargeting. ACM Trans. on Graphics, 27(3), 1-9, (2008).
15. Rubinstein, M., Shamir, A., Avidan, S.: Multi-operator media retargeting. ACM Trans. on Graphics, 28(3), 1-11 (2009).
16. Santella, A., Agrawala, M., Decarlo, D., Salesin, D., Cohen, M.: Gaze-based interaction for semiautomatic photo cropping. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 771–780 (2006).
17. Suh, B., Ling, H., Bederson, B., Jacobs, D., W.: Automatic thumbnail cropping and its effectiveness. Proceedings of ACM Conference on User Interface and Software Technology (UIST), pp. 95-104 (2003).
18. Setlur, V., Takagi, S., Raskar, R., Gleicher, M., Gooch, B.: Automatic image retargeting. Proceedings of Mobile and Ubiquitous Multimedia (MUM), pp. 59–68, (2005).
19. Teorex Developing Ideas, Inpaint Software. http://www.theinpaint.com/.
20. Vaquero, D., Turk, M., Pulli, K., Tico, M., Gelfand, N.: A survey of image retargeting techniques. Proceedings of SPIE Applications of Digital Image Processing XXXIII (2010).
21. Wang, Y.-S., Tai, C.-L., Sorkine, O., Lee, T.-Y.: Optimized scale-and-stretch for image resizing. ACM Trans. on Graphics, 27(5), 1-8 (2008).
22. Zhang, M., Zhang, L., Sun, Y., Feng, L.: Ma, W.: Auto cropping for digital photographs. Proceedings of IEEE International Conference on Multimedia and Expo (ICME), pp. 438-441, (2005).