

A Shape-Preserving Approach to Image Resizing

Guo-Xin Zhang¹ Ming-Ming Cheng¹ Shi-Min Hu¹ Ralph R. Martin²

¹Tsinghua National Laboratory for Information Science and Technology
Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
²School of Computer Science, Cardiff University, UK

Abstract

We present a novel image resizing method which attempts to ensure that important local regions undergo a geometric similarity transformation, and at the same time, to preserve image edge structure. To accomplish this, we define handles to describe both local regions and image edges, and assign a weight for each handle based on an importance map for the source image. Inspired by conformal energy, which is widely used in geometry processing, we construct a novel quadratic distortion energy to measure the shape distortion for each handle. The resizing result is obtained by minimizing the weighted sum of the quadratic distortion energies of all handles. Compared to previous methods, our method allows distortion to be diffused better in all directions, and important image edges are well-preserved. The method is efficient, and offers a closed form solution.

Categories and Subject Descriptors (according to ACM CCS): I.4 [Computing methodologies]: Image Processing and Computer Vision—Applications

1. Introduction

Current display devices have widely differing resolutions, so content aware image resizing is important. Such a tool is also highly desired by designers who need to create alternative images of different sizes. A good resizing method should allow deformation of images to arbitrary size while, simultaneously, prominent objects should be kept geometrically similar to their original shapes and any distortion should be optimally diffused into less important regions.

When resizing an image, simple methods such as scaling and cropping have obvious limitations. Simple image scaling distorts the entire image if the input and output aspect ratios differ too much. Cropping simply discards some image parts and is inappropriate when multiple objects of interest are far away from each other.

Seam carving [AS07, RSA08, SA09] is a recent method which can efficiently inhomogeneously resize images. It works by greedily removing (or inserting) one-dimensional seams passing through less important regions. As noticed by [WTS08], the discrete nature of seam carving may result in noticeably jagged edges of image objects after resizing. Further reduction of image size after *all* less-important regions have been removed leads to severe artifacts. Non-

uniform image wrapping [GSCO06, WGCo07] and random walks [ZHM08] give continuous solutions for content aware image resizing. However, all these methods only propagate distortion along the resizing direction. For images where any important regions are bigger than the final target size in the resizing direction, such methods inevitably generate severe distortion.

More recently Wang et al. [WTS08] proposed an optimized scale-and-stretch warping method for image resizing using a quad-mesh. Instead of forcibly leaving the sizes of certain regions unchanged, they attempt to ensure that important quads have homogeneous scaling, while minimizing bending of grid lines. This method distributes the distortion in all spatial directions. Compared to previous image warping methods, it better utilizes the available homogeneous regions to absorb the distortion. However, as shown in Figure 1, their method tends to avoid rotating quads, resulting in nearly horizontal or vertical grid lines. This leads to near-uniform scaling in any given row or column, and distortion is not well diffused. Another drawback is the absence of large-scale feature preservation. For a prominent object occupying many quads, locally, each quad has an acceptable homogeneous scaling, but over the whole object, non-trivial distortion may occur due to different scaling factors of quads—

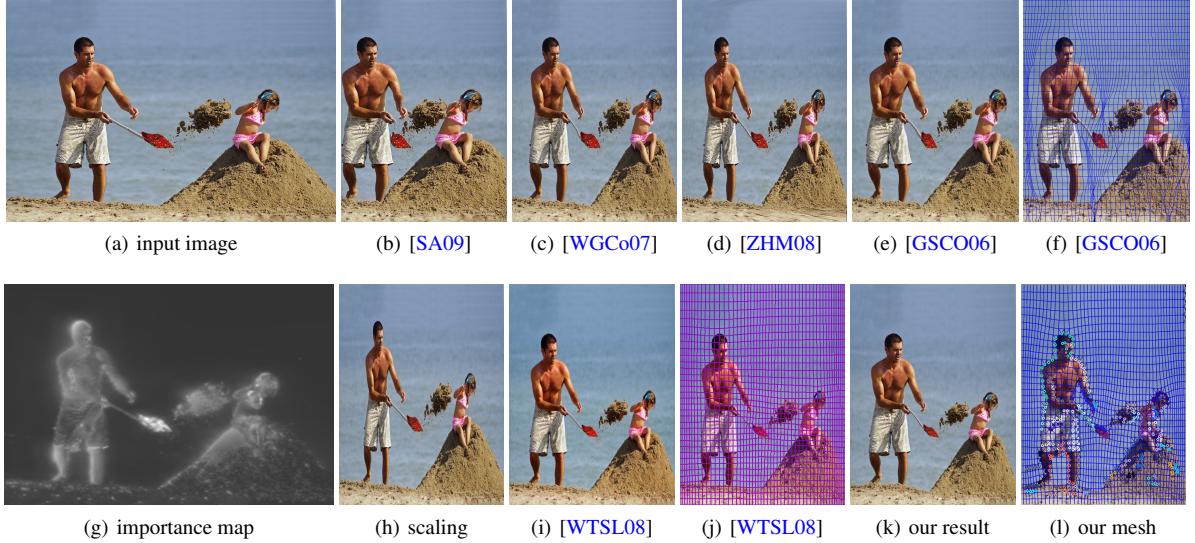


Figure 1: Comparison of different resizing methods. In (b)-(e), significant distortion occurs since unidirectional propagation of distortion cannot cope with images where the important information to be retained (here, from the man to the girl) is wider than the destination size. Our method better preserves prominent objects than [WTSLO8] and optimally propagates distortion in all directions. In (l), highlighted points with the same color represent a B-Handle—see Section 3.2.2 (we use the same notation in all figures).

even if a penalty is applied for differential scaling of neighboring quads.

We present a novel method which attempts to ensure that the new shapes of prominent objects are geometrically similar to their original shapes both locally and globally. Following [WTSLO8], we use a grid mesh and optimally diffuse distortion into less important regions in all directions. But unlike [WTSLO8], our method allows quads to undergo a similarity transformation, and preserves important edge features. As a result, distortion is better diffused, and large prominent objects are better preserved by edge similarity constraints.

We first associate an image with the vertices of a grid mesh, and additional points sampled from the image edges (see Section 3.2.1). Next, all of these points are automatically grouped into local or larger-scale control point sets, which we call *handles* (see Figure 2). A handle can describe a local region (a mesh quad) or an extended feature (an image edge) in the original image. Using a novel definition for distortion energy for handles (see Section 3), *similarity constraints* are used to preserve the geometric similarity of the handles in an energy minimization process: we are able to keep important handles similar to their original shapes during resizing. Thus, shapes of important quads and image edges are well-preserved; in turn, the global shapes of prominent objects are thus also well-preserved.

Our main contributions are:

- a novel image resizing method which both preserves im-

portant local regions and image edges, and better diffuses distortion.

- an efficient closed-form linear solution for the quadratic distortion energy minimization problem needed to achieve this.

Furthermore, we also show there is a close relationship between our distortion energy, and the conformal energy used for least squares conformal parameterization in geometry processing.

2. Related work

The growing requirement for image resizing has brought this topic to current attention. Standard uniform scaling can be performed in real-time using linear, or higher-order, interpolation. However, this simple method does not consider image content and is highly distorting when the aspect ratio changes significantly. To overcome this problem, recent research has considered content-aware resizing [LXMZ03, FXZM03, STR^{*}05, TJS07, AS07, WGCo07, WTSLO8, WHZ^{*}08, SCSI08, HFRQ09]. A fundamental idea is to control image resizing using an *importance map* obtained from the image, ensuring that important regions undergo smaller changes than less-important ones.

Cropping has been widely investigated [SLBJ03, CXF^{*}03, LG05, LG06, SAD^{*}06, TJS07]. Such methods use a visual saliency model to determine important areas in an image or video and use them to select an optimal rectangle

from an input image or video frame. A clear drawback of cropping methods is that important regions may lie at opposite edges of the image—a common problem in real images.

Seam carving [AS07, RSA08, SA09] can efficiently resize an image or video by removing or duplicating seams in unimportant areas. A seam in an image is a contiguous one-dimensional path of pixels from top-to-bottom, or side-to-side, each having locally minimal importance. Although such methods can produce compelling results on many images, the discrete nature of these algorithms often leads to clear artifacts in objects with a well-defined structure. This problem is especially noticeable when too few unimportant seams can be found to reach the target size.

Non-uniform warping [LG05, GSCO06, WGCo07] and *random walk* [ZHM08] methods offer continuous solutions to image resizing. Liu et al. [LG05] use a non-linear image warping function to emphasize the important aspects of the image while retaining the surrounding context. However, image features outside regions of interest may be badly distorted by this method. Gal et al. [GSCO06] warps an image in a way that locally constrains the deformations to be similarity transformations. Wolf et al. [WGCo07] merge less important pixels in the reduction direction. However, these and random walk solutions [ZHM08] only propagate distortion along the resizing direction, as do seam carving approaches.

Scale-and-stretch [WTS08] offers a continuous solution to image resizing and finds a homogeneous scaling for prominent objects by global optimization. This allows distortion to be diffused in all directions, and has the advantage of distributing distortion over the whole image, irrespective of the resizing direction. This gives it advantages over traditional algorithms such as feature aware texturing [GSCO06], non-homogeneous content-driven video-retargeting [WGCo07] and seam carving [AS07, RSA08].

Recently, concepts from image resizing have also been applied to content-aware *shape resizing* and focus-and-context visualization of 3D models; typical works are [KSSCO08] and [WLT08].

Conformal parameterization techniques are widely used in geometry processing. Lévy et al. [LPRM02] described a least-squares approach for conformal parameterization by minimizing a quadratic conformal energy. Gu et al. [GY03] also introduced a novel method for global conformal parameterization. For a full review of parameterization techniques, including conformal parameterization, the reader is referred to [FH05]. Later, we show the close relationship between our method and least-squares conformal parameterization.

3. Shape-preserving image resizing

To take into account image content, we propose use of *similarity constraints* for resizing images, in order to preserve

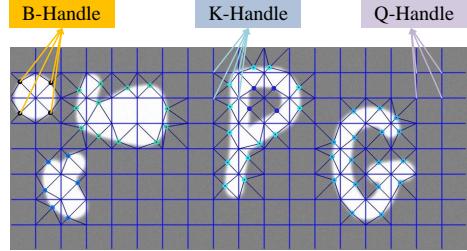


Figure 2: Handles.

the shapes of important regions and image edges. Ideally, each image region should undergo a homogeneous scaling. This is clearly impossible if the new image size has a different aspect ratio. To achieve arbitrary image resizing, we diffuse distortion throughout the image, allowing distortion to be greater in less important regions, while more important regions and image edges are less deformed.

3.1. Overview

We cover the image with a grid mesh $M = (V_1, F)$ with vertices V_1 and quad faces F , where $V_1 = \{v_i\}$ denotes initial vertex positions. To better capture significant objects, we also add additional edge points on image edges, denoted by $V_2 = \{v_i\}$. We write $V_+ = V_1 \cup V_2$ for the set containing all these points, which we call the *control points*. Our image resizing algorithm attempts to find a deformed mesh geometry V'_+ under similarity transformation constraints on the original control points. The output image is obtained from the final positions of the mesh vertices V'_1 using cubic interpolation [Wol90]. We group the control points into particular sets which we call *handles*, denoted by $\mathcal{P} = \{P_i\}$. Each handle is a subset of the control points: $P_i \subset V_+, \forall P_i \in \mathcal{P}$, and every control point belongs to one or more handles. We also require that each handle has at least 3 distinct control points.

See Figure 2 for a conceptual view of handles; we explain how to choose the additional edge points, and allocation of control points to different kinds of handles, later. Intuitively, a handle is a point set whose shape we wish should remain unchanged after deformation; more important handles are given larger weights. We denote by $\mathcal{P}' = \{P'_i\}$ the deformed handles corresponding to \mathcal{P} ; each P'_i is a subset of V'_+ . Each handle P_i is assigned a quadratic energy term, its *distortion energy* $\mathcal{E}(P'_i, P_i)$, which measures the dissimilarity between the deformed handle P'_i and its original shape P_i .

The total distortion energy is formulated as:

$$E = \sum_{P_i \in \mathcal{P}} \omega_i \mathcal{E}(P'_i, P_i) \quad (1)$$

The new positions V'_+ are obtained by minimizing the quadratic energy E with appropriate boundary conditions.

We now proceed in Section 3.2 to discuss the choice of additional points and handles for an input image. We then

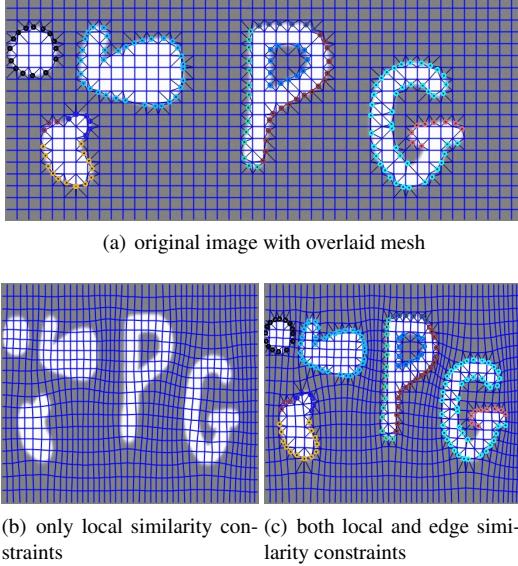


Figure 3: Illustration for edge similarity constraints: With only local similarity constraints (b) image edges are not well preserved. Adding edge similarity constraints (c) helps to overcome this artifact.

give a detailed definition of distortion energy in Section 3.3 and its relation to conformal energy from geometry processing in Section 3.4. Weights and boundary conditions are discussed in Section 3.5.

3.2. Choice of Additional Control Points and Handles

3.2.1. Additional control points V_2

We use the method in [Ste98] to detect image edges as series of linked edge points in the original image. Edges are considered to be image features which should retain a similar shape under deformation. For each mesh quad that strictly contains one or more edge points, we add that edge point nearest to the quad center to V_2 : thus, each quad face can contain at most one point in V_2 .

3.2.2. Handles \mathcal{P}

As shown in Figure 2, we define three types of handles:

- **Q-Handle** A *Q-Handle* comprises the four corners of a quad face which does not contain an additional control point. A Q-Handle is added to \mathcal{P} for every such quad face.
- **B-Handle** A *B-Handle* is a subset of V_+ , each element belonging to the same image edge. For each image edge containing at least 3 points in V_+ , we add a B-Handle to \mathcal{P} containing all of its points in V_+ . See Figure 2 for an example. Highlighted control points with the same color represent a B-Handle.
- **K-Handle** A *K-Handle* is a set containing the four corners and the inner control point of a quad face which contains an additional control point. We add a K-Handle to \mathcal{P}

for every such quad face. Intuitively, such K-Handles are used to link the similarity constraints for Q-Handles and B-Handles.

In this paper, we call the distortion energy terms for Q-Handles and K-Handles the *local similarity constraints*, and call energy terms for B-Handles the *edge similarity constraints*.

3.3. Distortion energy

Given a handle $P = \{p_i\}$ with m distinct control points, its deformed position is $P' = \{p'_i\}$; note that $m \geq 3$. Its *distortion energy* is defined to be

$$\mathcal{E}(P', P) = \min_{s \in \mathcal{S}} \sum_{i=1}^m |s(p_i) - p'_i|^2 \quad (2)$$

where \mathcal{S} is the set of similarity transformations s in \mathbb{R}^2 having the general form:

$$s(p) = \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \quad p = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2, \quad (3)$$

where c, d, t_x, t_y are four real parameters determining a unique similarity transformation.

Intuitively, the energy measures the shape distortion of P' relative to P , while ignoring translation and rotation. It is the minimal possible square distance between P' and P , if P may be replaced by any similarity transformation of itself.

As we allow P' to vary, $\mathcal{E}(P', P)$ has a quadratic dependence on P' , and can be alternatively written (see Appendix B for details):

$$\mathcal{E}(P', P) = |C_P b_{P'}|^2 \quad (4)$$

where

$$C_P = A_P (A_P^T A_P)^{-1} A_P^T - I, \quad (5)$$

and

$$A_P = \begin{bmatrix} x_1 & -y_1 & 1 & 0 \\ y_1 & x_1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_m & -y_m & 1 & 0 \\ y_m & x_m & 0 & 1 \end{bmatrix}, \quad b_{P'} = \begin{bmatrix} x'_1 \\ y'_1 \\ \vdots \\ x'_m \\ y'_m \end{bmatrix}^T. \quad (6)$$

For each handle P , we calculate the corresponding C_P , and then find the minimal solution of E directly. The C_P matrix is identical for all Q-Handles (because C_P is invariant if P undergoes a similarity transformation, and all Q-Handles have the same shape), allowing us to more efficiently set up the equation system; note that Q-Handles are usually the most common type.

3.4. Relation to Conformal Parametrization

There is a close relationship between our distortion energy and least squares conformal parametrization, which is a



Figure 4: Comparison of meshes produced by optimal scale-and-stretch [WTS08] and our method when utilizing local similarity constraints only. Even without our extra edge similarity constraints, our method diffuses distortion more uniformly to less important regions.

well-studied topic in geometry processing. The latter seeks a mapping from a triangular mesh in R^3 to a region in R^2 which preserves angles as well as possible. The positions of the vertices in R^2 are found by minimizing a quadratic total conformal energy function. For a triangle T and its image T' , the quadratic conformal energy is denoted by $\mathcal{C}(T', T)$. Again, this quadratic energy measures the shape difference between the deformed and original triangles. Total conformal energy is then defined by

$$\mathcal{C}(\mathcal{T}) = \sum_{T \in \mathcal{T}} \mathcal{C}(T', T) \quad (7)$$

where \mathcal{T} is the set of all triangles. The conformal energy is based on a set of quadratic constraints each of which restricts a triangle under a similarity transformation. For a detailed description of conformal energy, see [LPRM02, FH05].

We show in Appendix A that for each triangle T , $\mathcal{C}(T', T) = \alpha_T \mathcal{E}(T', T)$, where α_T is a constant depending on T , and furthermore α_T is invariant if T undergoes a similarity transformation. Thus, the distortion energy (defined for arbitrary 2D point sets) is closely related to the classical conformal energy (defined for triangles).

3.5. Weights and Boundary Conditions

More important quads should be distorted less, so they are given larger weights. We first compute an importance value for each pixel, as explained shortly. The weight ω_i for a Q-Handle or a K-Handle is then defined as the average of the

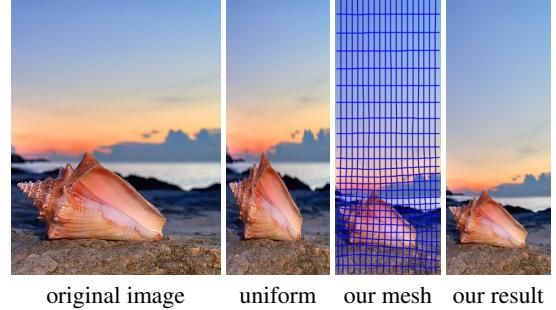


Figure 5: Resizing this image with uniform scaling leads to significant distortion. Our method attempts to subject important regions to a similarity transformation to preserve image features. The shell is scaled in a more meaningful way and distortion is distributed over less important regions to achieve the destination image size. As the shell is wider than the destination image, any resizing technique operating only in the resizing direction would cause significant distortion. Our method optimally diffuses distortion in all directions.

importance values of those pixels in the corresponding quad. The weight of a B-Handle P_B is defined as $\alpha \omega_{P_B}$, where ω_{P_B} is the average weights of pixels in all quads which contain at least one point in P_B ; α is a constant number determining the importance of the edge similarity constraints and is often set to 1.

To determine the importance value for each pixel, an importance map is computed, as is done in most content aware image resizing algorithms. Importance maps can be defined in a bottom up manner to emphasize strong or unusual stimuli [IKN98, AS07, WTS08, HZ07, HZ07, GMZ08, ZS06], or a top down manner to emphasize specific scene elements such as human faces [VJ04] or regions of interest [WRL*04]. Integration of both top-down and bottom-up importance has also been used [WGCo07]. In our experiments, we have used an equally-weighted linear combination of two bottom-up measures proposed in [LSZ*07]: the multi-scale contrast and color spatial-distribution. More sophisticated measures or hand-tuned importance maps could potentially further improve the quality of our results. We normalize the importance values to lie in the range $[\epsilon, 1]$, where ϵ is a small positive constant to ensure computational stability (we typically set $\epsilon = 0.2$).

We must also enforce boundary conditions when computing the optimal solution. We do so by making sure that mesh vertices lying around the outer boundary of the image remain there after resizing. To resize the image from $w \times h$ to $w' \times h'$, the boundary condition is

$$\begin{cases} v'_{i,y} = 0 & \text{if } v_{i,y} = 0 \\ v'_{i,y} = h' & \text{if } v_{i,y} = h \\ v'_{i,x} = 0 & \text{if } v_{i,x} = 0 \\ v'_{i,x} = w' & \text{if } v_{i,x} = w \end{cases}. \quad (8)$$

Method	[SA09]	[WTSL08]	[WGCo07]	[ZHM08]	Our
Time	1~2s	(0.034+0.002n)s	5.6s	0.21s	0.024s
Size	400×500	1024×754	1024×754	1024×754	1024×754

Table 1: Speed comparison.

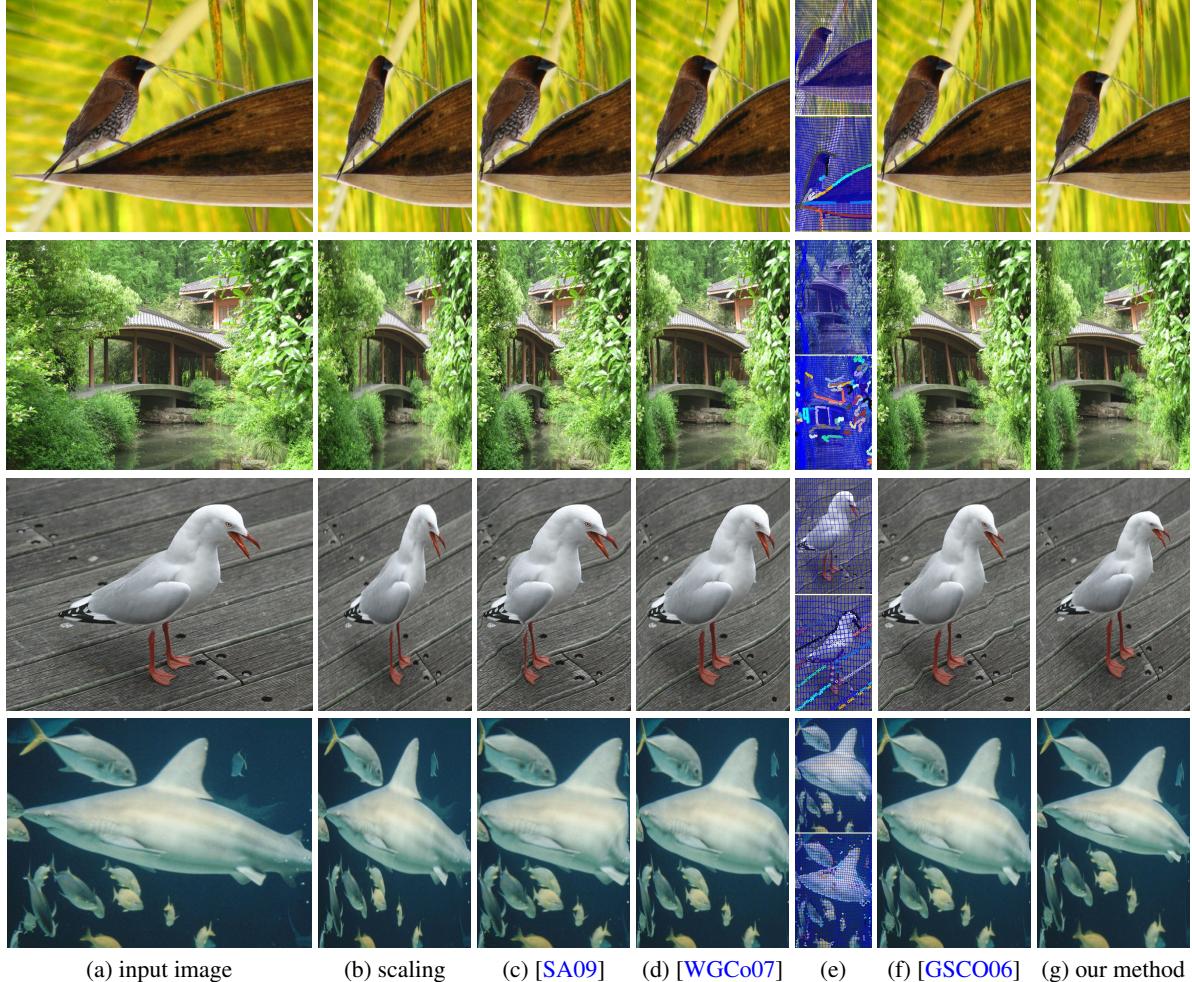


Figure 6: Further comparisons. The meshes produced by [GSCO06] (upper) and our method (lower) are illustrated in (e).

4. Results and Discussions

We have implemented our algorithm on a computer with a 2.33GHz Pentium Duo CPU and 4GB RAM.

Our method is very efficient: the most computationally expensive part is solving a small sparse linear system which depends on the size of mesh; a finer mesh produces better results but takes longer. To make a fair comparison with [WTSL08] and [GSCO06], we use grids of the same resolution for the same picture (the quad size is typically 20×20 , and produces sufficiently good results). Solving the linear system for a 1024×754 image typically takes 0.024s.

We demonstrate the effectiveness of our algorithm with several challenging examples (see Figures 1 and 3–7).

4.1. Comparison

We mostly compare our work with two related state-of-the-art methods of differing classes: seam carving [SA09] for discrete methods and [WTSL08] for continuous methods. For the sake of fairness, we use the same importance maps mentioned in Section 3.5 for all the methods. We also make comparisons with other continuous methods such as [GSCO06, WGCo07, ZHM08].

The seam carving method is effective when dealing with

images containing uniform or highly textured regions. In addition, it allows higher flexibility of operations on pixels, and further applications such as object removal can also be performed. However, the discrete nature of this algorithm makes it less effective when dealing with objects with clear structures. Figures 1, 6 and 7 show noticeable staircase effects in large scale objects. In situations when an important object is wider than the destination image (e.g. Figures 1, 4 and 5), scaling in only the resizing direction is insufficient.

Perhaps more relevant is [WTS08] which gives a continuous solution and makes use of less important regions in all directions. As noted earlier, its main problem is a tendency to produce nearly horizontal or vertical grid lines, causing nearly uniform scaling factors in any given row or column, and it lacks global feature preservation. Our method allows similarity transformations, and the distortion is better diffused, with global edges better preserved—see Figures 1, 4 and 7.

In parameterization, *triangle folding* can be a problem. This occurs when a triangle has the opposite orientation, locally, before and after deformation. Least squares solution of the energy equation in conformal parameterization cannot prevent triangle folding [FH05]. However, since a similarity transformation preserves orientation, any inconsistent orientation will cause a large increase in distortion energy. A similar observation applies to our method. Such issues rarely occur in practice.

Table 1 compares times taken by our method and other methods such as seam carving (SC) [AS07] and optimized scale-and-stretch (OSS) [WTS08]. Both our method and OSS use 20×20 grids and take a significantly shorter time than SC. The cost of OSS depends on the number of iterations n . Our method is quicker than OSS as it uses a closed form linear solution. SC, [WGCo07] and [ZHM08] are much slower as they work on pixels, or alternatively, a much denser grid. [ZHM08] is much faster than [WGCo07] as they use a multigrid solver. The times in the table do not include preprocessing steps such as edge detection or importance map computation, and the proposed importance map computation takes about 0.1s for a 1024×754 picture.

4.2. Limitations

As the constraints in our method are all soft constraints, our method cannot guarantee to strictly preserve edges. The seagull example in Figure 6 shows such a case; some lines are not well preserved. However, this is partly because of deficiencies in edge detection. Details of the detected edges can be seen in the mesh of the seagull example. Many of the edges are partly occluded by the seagull and are thus not treated as a single line which should be preserved. Nevertheless, our approach generally works well and preserves edges to a large extent; Figure 3 shows the effects of the edge similarity constraints.

Since handles are allowed to rotate, our method may cause a prominent object to rotate slightly. However, it will not rotate seriously in practice as this would cause large distortion in its vicinity. Nevertheless, even small artifacts can be noticeable as humans can be very sensitive to certain rotations, e.g. vertical or horizontal structural lines. The second example in Figure 6 shows such a case: some columns rotate slightly. The best solution may be to detect such structures and penalize their rotation.

As shown by the seagull image in Figure 6, our method diffuses distortion to less important regions. Artifacts may be more or less visible in such regions depending on content. An ideal solution for image resizing may be to separately process foreground and background layers, perhaps scaling the foreground while cropping the background. However, such an approach relies on other difficult problems: foreground / background segmentation, and image completion.

We also find that our results are sensitive to the quality of the importance map—see Figure 8. In this case, the importance map is rather poor, causing our method to produce worse results than [WTS08], which tends to produce horizontal or vertical grid lines, making the result more like uniform scaling in this example. In contrast, our method allows the grid to rotate; this flexibility may distort objects if the importance map is poor.

5. Conclusions and future work

We have given a novel image resizing method which attempts to retain the shapes of local regions and significant image edges. Distortion energy is used to minimize the distortion of *handles*: mesh quads and image edges. Compared to previous continuous image resizing methods, distortion is better diffused into less important regions in all directions, and the shapes of prominent objects are better preserved. The method has a fast closed form solution.

The most important extension of this work would be to resizing of video. Although our least-squares solution has advantages in continuity, continuity between video frames needs further consideration. Another extension would be the use of more sophisticated importance detection algorithms.

Acknowledgements

We thank all the anonymous reviewers for their valuable comments. We are also grateful to Yu-Shuen Wang for his help with the results of [WTS08]. This work was supported by the National Basic Research Project of China (Project Number 2006CB303106), the Natural Science Foundation of China (Project Number U0735001) and the National High Technology Research and Development Program of China (Project Number 2009AA01Z327).

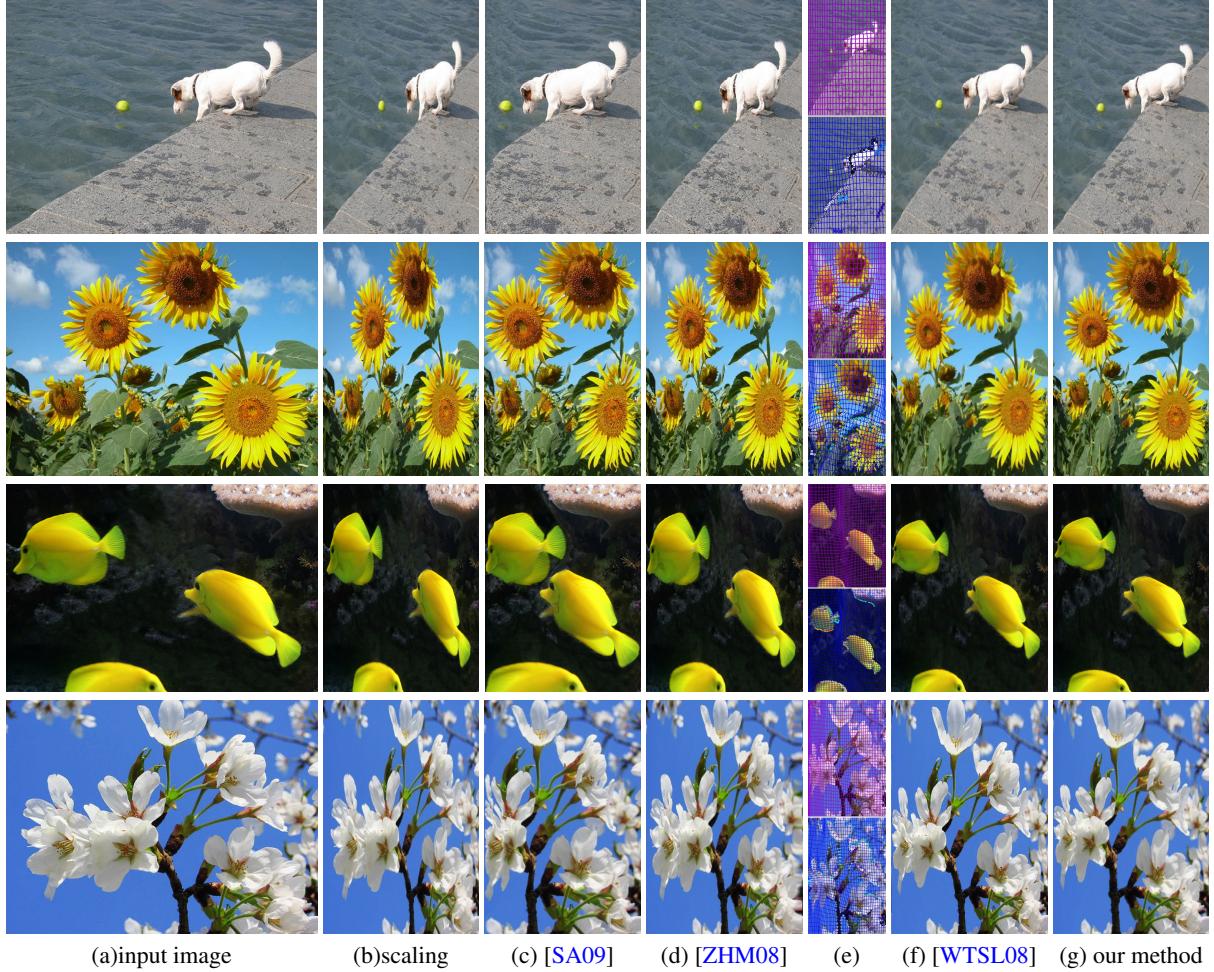


Figure 7: Further comparisons. The meshes produced by [WTSLO8] (upper) and our method (lower) are illustrated in (e).

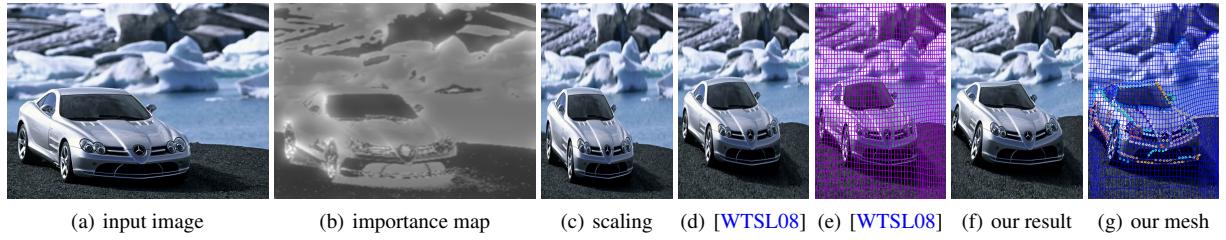


Figure 8: A less successful example.

References

- [AS07] AVIDAN S., SHAMIR A.: Seam carving for content-aware image resizing. *ACM Trans. Graph.* 26, 3 (2007), 10.
- [CXF*03] CHEN L.-Q., XIE X., FAN X., MA W.-Y.,

ZHANG H., ZHOU H.-Q.: A visual attention model for adapting images on small displays. *Multimedia Syst.* 9, 4 (2003), 353–364.

[FH05] FLOATER M. S., HORMANN K.: Surface parameterization: a tutorial and survey. In *Advances in Multiresolution for Geometric Modelling* (2005), Springer,

- pp. 157–186.
- [FXZM03] FAN X., XIE X., ZHOU H.-Q., MA W.-Y.: Looking into video frames on small displays. In *Proceedings of International Conference on Multimedia* (2003), pp. 247–250.
- [GMZ08] GUO C. L., MA Q., ZHANG L. M.: Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform. In *Proc. CVPR* (2008), pp. 1–8.
- [GSCO06] GAL R., SORKINE O., COHEN-OR D.: Feature-aware texturing. In *Proceedings of Eurographics Symposium on Rendering* (2006), pp. 297–303.
- [GY03] GU X., YAU S.-T.: Global conformal surface parameterization. In *Proceedings of the Eurographics/ACM SIGGRAPH symposium on Geometry processing* (2003), pp. 127–137.
- [HFRQ09] HUANG H., FU T., ROSIN P. L., QI C.: Real-time content-aware image resizing. *Science in China Series F: Information Sciences* 52, 2 (2009), 172–182.
- [HZ07] HOU X. D., ZHANG L. Q.: Saliency detection: A spectral residual approach. In *Proc. CVPR* (2007), pp. 1–8.
- [IKN98] ITTI L., KOCH C., NIEBUR E.: A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 20, 11 (1998), 1254–1259.
- [KSSCO08] KRAEVOY V., SHEFFER A., SHAMIR A., COHEN-OR D.: Non-homogeneous resizing of complex models. *ACM Trans. Graph.* 27, 5 (2008), 1–9.
- [LG05] LIU, FENG, GLEICHER, MICHAEL: 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.
- [LG06] LIU F., GLEICHER M.: Video retargeting: automating pan and scan. In *Proceedings of the 14th International Conference on Multimedia* (2006), ACM, pp. 241–250.
- [LPRM02] LÉVY B., PETITJEAN S., RAY N., MAILLOT J.: Least squares conformal maps for automatic texture atlas generation. *ACM Trans. Graph.* 21, 3 (2002), 362–371.
- [LSZ*07] LIU T., SUN J., ZHENG N.-N., TANG X., SHUM H.-Y.: Learning to detect a salient object. In *Proc. CVPR* (2007), pp. 1–8.
- [LXMZ03] LIU H., XIE X., MA W.-Y., ZHANG H.-J.: Automatic browsing of large pictures on mobile devices. In *Proceedings of the eleventh ACM international conference on Multimedia* (2003), ACM, pp. 148–155.
- [RSA08] RUBINSTEIN M., SHAMIR A., AVIDAN S.: Improved seam carving for video retargeting. *ACM Trans. Graph.* 27, 3 (2008), 1–9.
- [SA09] SHAMIR A., AVIDAN S.: Seam carving for media retargeting. *Commun. ACM* 52, 1 (2009), 77–85.
- [SAD*06] SANTELLA A., AGRAWALA M., DECARLO D., SALESIN D., COHEN M.: Gaze-based interaction for semi-automatic photo cropping. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (New York, NY, USA, 2006), ACM, pp. 771–780.
- [SCSI08] SIMAKOV D., CASPI Y., SHECHTMAN E., IRANI M.: Summarizing visual data using bidirectional similarity. In *Proc. CVPR* (2008), pp. 1–8.
- [SLBJ03] SUH B., LING H., BEDERSON B. B., JACOBS D. W.: Automatic thumbnail cropping and its effectiveness. In *Proceedings of the ACM Symposium on User Interface Software and Technology* (2003), ACM, pp. 95–104.
- [Ste98] STEGER C.: An unbiased detector of curvilinear structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 20, 2 (1998), 113–125.
- [STR*05] SETLUR V., TAKAGI S., RASKAR R., GLEICHER M., GOOCHE B.: Automatic image retargeting. In *Proceedings of International Conference on Mobile and Ubiquitous Multimedia* (2005), vol. 154, ACM, pp. 59–68.
- [TJS07] TAO C., JIA J., SUN H.: Active window oriented dynamic video retargeting. In *Proceedings of the Workshop on Dynamical Vision, ICCV* (2007).
- [VJ04] VIOLA P., JONES M. J.: Robust real-time face detection. *Int. J. Comput. Vision* 57, 2 (2004), 137–154.
- [WGC07] WOLF L., GUTTMANN M., COHEN-OR D.: Non-homogeneous content-driven video-retargeting. In *Proc. ICCV* (2007), pp. 1–6.
- [WHZ*08] WEI L.-Y., HAN J., ZHOU K., BAO H., GUO B., SHUM H.-Y.: Inverse texture synthesis. *ACM Trans. on Graph.* 27, 3 (2008), 52:1–52:9.
- [WLT08] WANG Y.-S., LEE T.-Y., TAI C.-L.: Focus+context visualization with distortion minimization. *IEEE Trans. Vis. Comput. Graph.* 14, 6 (2008), 1731–1738.
- [Wol90] WOLBERG G. (Ed.): *Digital Image Warping*. IEEE Computer Society Press, 1990.
- [WRL*04] WANG J., REINDERS M. J. T., LAGENDIJK R. L., LINDBERG J., KANKANHALLI M. S.: Video content representation on tiny devices. In *Proceedings Of IEEE Conference on Multimedia and Expo* (2004), IEEE, pp. 1711–1714.
- [WTSL08] WANG Y.-S., TAI C.-L., SORKINE O., LEE T.-Y.: Optimized scale-and-stretch for image resizing. *ACM Trans. Graph.* 27, 5 (2008), 1–8.
- [ZHM08] ZHANG Y.-F., HU S.-M., MARTIN R. R.: Shrinkability maps for content-aware video resizing. *Comput. Graph. Forum* 27, 7 (2008), 1797–1804.
- [ZS06] ZHAI Y., SHAH M.: Visual attention detection in video sequences using spatiotemporal cues. In *Proceedings of the 14th annual ACM international conference on Multimedia* (New York, NY, USA, 2006), ACM, pp. 815–824.

Appendix A: Relation between distortion energy and discrete conformal energy for triangles

Here we prove that our definition of distortion energy is closely related to conformal energy defined on triangles.

Let a triangle T have image triangle T' after deformation.

T may be represented as $T = \{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}$ in a local x - y plane with its normal along the z -axis. Let $Q = A(A^T A)^{-1} A^T$ where A has the form of A_P in Eqn. 6. Then the distortion energy can be described as

$$\begin{aligned}\mathcal{E} &= b^T (C^T C) b = b^T ((Q - I)^T (Q - I)) b \\ &= b^T (I - Q) b = b^T W b\end{aligned}\quad (9)$$

where $W = I - Q$, and C and b are C_P and b_P in Eqs. 5 and 6. To formulate the equation for Q , we write

$$\begin{aligned}L &= \begin{bmatrix} x_1 & -y_1 \\ y_1 & x_1 \\ x_2 & -y_2 \\ y_2 & x_2 \\ x_3 & -y_3 \\ y_3 & x_3 \end{bmatrix} O_{2 \times 6} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \\ O_{6 \times 6} &= [O_{2 \times 6}, O_{2 \times 6}, O_{2 \times 6}] \\ s_x &= x_1 + x_2 + x_3, s_y = y_1 + y_2 + y_3 \\ s_2 &= x_1^2 + x_2^2 + x_3^2 + y_1^2 + y_2^2 + y_3^2 \\ \Delta &= 3s_2 - s_x^2 - s_y^2 \\ K &= \begin{bmatrix} s_x & s_y \\ -s_y & s_x \end{bmatrix}\end{aligned}$$

whereupon

$$Q = \frac{1}{\Delta} \left(3LL^T + s_2 O_{6 \times 6} - LKO_{2 \times 6}^T - (LKO_{2 \times 6})^T \right) \quad (10)$$

Note that Q is invariant if T undergoes a similarity transformation. Without loss of generality, therefore, we may translate the triangle and scale it uniformly so that

$$s_x = 0, \quad s_y = 0, \quad s_2 = 3 \quad (11)$$

In this case, $K = 0$, $\Delta = 9$, and Equation 10 becomes

$$Q = \frac{1}{3} (LL^T + O_{6 \times 6}). \quad (12)$$

It is convenient to write Q in complex form, where

$$z_j = x_j + iy_j, \quad z'_j = x'_j + iy'_j, \quad j = 1, 2, 3. \quad (13)$$

The complex forms for L , b , Q and W are

$$\begin{aligned}L_z &= [z_1, z_2, z_3]^T, \quad b = [z'_1, z'_2, z'_3]^T, \\ Q_z &= \frac{1}{3} (L_z L_z^H + O_{z, 3 \times 3}), \\ W_z &= \frac{1}{3} (3I_3 - O_{z, 3 \times 3} - L_z L_z^H),\end{aligned}\quad (14)$$

where $O_{z, 3 \times 3}$ is a 3×3 matrix with all elements 1. The distortion energy is then

$$\mathcal{E} = b^T W b = b_z^H W_z b_z. \quad (15)$$

In detail W_z is:

$$W_z = \frac{1}{3} \begin{bmatrix} 2 - z_1 \bar{z}_1 & -1 - z_1 \bar{z}_2 & -1 - z_1 \bar{z}_3 \\ -1 - z_2 \bar{z}_1 & 2 - z_2 \bar{z}_2 & -1 - z_2 \bar{z}_3 \\ -1 - z_3 \bar{z}_1 & -1 - z_3 \bar{z}_2 & 2 - z_3 \bar{z}_3 \end{bmatrix} \quad (16)$$

Now, the conformal energy in [LPRM02] is defined as:

$$\mathcal{E}' = b_z^H W' b_z \quad (17)$$

where d_T is twice the area of the triangle, and

$$W' = \frac{1}{d_T} [z_3 - z_2, z_1 - z_3, z_2 - z_1]^H [z_3 - z_2, z_1 - z_3, z_2 - z_1]$$

Note that W' is also invariant under a similarity transformation. It is easy to verify that $W' = 9W_z/d_T$ if we follow Equation 11 and set

$$z_1 + z_2 + z_3 = 0, \quad |z_1|^2 + |z_2|^2 + |z_3|^2 = 3. \quad (18)$$

In general, (without the assumption in Equation 11), we have

$$W' = 3\gamma W_z/d_T = \alpha_T W_z \quad (19)$$

where

$$\gamma = |z_1 - z_c|^2 + |z_2 - z_c|^2 + |z_3 - z_c|^2, \quad z_c = \frac{z_1 + z_2 + z_3}{3}. \quad (20)$$

Here, $\alpha_T = 3\gamma/d_T$ is a constant which depends on the original triangle T ; note that α_T is invariant for triangles related by a similarity transformation.

Appendix B: Detailed derivation of Equation 4

We use the same notations here as in Section 3.3. For any $s \in \mathcal{S}$, denote

$$X_s = [c \quad d \quad t_x \quad t_y]^T. \quad (21)$$

We have for $i = 1, \dots, m$ that

$$s(p_i) - p'_i = \begin{bmatrix} x_i & -y_i & 1 & 0 \\ y_i & x_i & 0 & 1 \end{bmatrix} X_s - \begin{bmatrix} x'_i \\ y'_i \end{bmatrix}. \quad (22)$$

Then Equation 2 can be further represented as

$$\mathcal{E}(P', P) = \min_{s \in \mathcal{S}} |A_P X_s - b_{P'}|^2 \quad (23)$$

The optimal X_s can be found by solving $A_P X_s = b_{P'}$ in a least square sense: $X_s = (A_P^T A_P)^{-1} A_P^T b_{P'}$. Thus,

$$\mathcal{E}(P', P) = |(A_P (A_P^T A_P)^{-1} A_P^T - I) b_{P'}|^2 = |C_P b_{P'}|^2. \quad (24)$$

It is easy to verify that if P contains at least two distinct control points, $\text{rank}(A_P) = 4$. In such cases, $A_P^T A_P$ is non-singular, and C_P is well-defined. But when P contains only two distinct points, $C_P = 0$, which is why we require each handle to contain at least 3 distinct control points.