Image Retargeting Using Multi-Map Constrained Region Warping

Tongwei Ren
State Key Lab for Novel Software Technology, Nanjing University
Nanjing 210093, China
rentw@graphics.nju.edu.cn

Yan Liu
Department of Computing, The Hong Kong Polytechnic University
Hong Kong, China
csyliu@comp.polyu.edu.hk

Gangshan Wu
State Key Lab for Novel Software Technology, Nanjing University
Nanjing 210093, China
gswu@graphics.nju.edu.cn

ABSTRACT
Image retargeting aims to adapt images to various screens with small sizes and arbitrary aspect ratios. In this paper, we propose a novel image retargeting approach based on region warping, which emphasizes the image parts with important content while reducing the visual distortion over the whole image. First, the original image is decomposed into homogeneous regions and further represented by curve-edge trapezoid meshes. Then, two kinds of energy maps, importance map and sensitivity map, are calculated by visual attention model and weighted gradient map respectively. With mesh representation and energy map constraints, image retargeting is formulated to a constrained optimization problem of mesh vertexes relocation. Finally, the target image is generated by separately warping the regions based on the deduced optimal solution. The experiments on different images demonstrate the effective and efficiency of our algorithm.

Categories and Subject Descriptors
I.4.9 [**Image Processing and Computer Vision**]: Applications

General Terms
Algorithms, Human Factors

Keywords
Image retargeting, region warping, mesh representation, multiple energy maps

1. INTRODUCTION
The wide applications of multimedia technology lead to the significant need of convenient image display using portable devices, such as mobile phone and PDA [1]. However, the resolution and aspect ratio of the existing image may not match the target display screen. The simplest solution is to scale the image uniformly. But when the target screen is much smaller or very different in aspect ratio [2], uniform scaling may cause serious loss or distortion of important details.

To address this problem, image retargeting technique is proposed, which scales different image parts separately based on their content importance. According to processing level, current image retargeting methods can be roughly classified into three categories: object level retargeting, pixel level retargeting and region level retargeting. Object level retargeting methods, including cropping [5] and segmentation-based recomposition [3], separate image content into objects and background and only emphasize the important objects. Obviously, their effects largely depend on the performance of object segmentation/detection. Pixel level retargeting methods, such as seam carving [6] and non-homogeneous warping [4], treat each pixel as a processing unit and try to retain more important pixels using constrained optimization algorithms. However, they suffer from image structure distortion and high computation cost [8][9]. Region level retargeting methods divide the image into a set of regions and scale these regions separately. Compared with two other categories of methods, region level retargeting has much lower requirement of image segmentation than object level retargeting and preserves the image structure better than pixel level retargeting. However, in the existing region level retargeting methods, fisheye-view warping [7] requires manual interaction and may cause serious distortion, scale-and-stretch [8] focuses on distortion reduction but completely ignores important content emphasis, and mesh parametrization [9] requires a preliminary either-or judgment to detect the meshes possibly including object parts and rigidly retains these detected meshes in retargeting.

In this paper, we propose a novel multi-map constrained region warping approach, which optimizes important content emphasis and visual distortion reduction simultaneously. The core of our approach is representing the original image with curve-edge trapezoid meshes and formulating image retargeting as an optimization problem of mesh vertexes relocation. Specially, we utilize multiple energy maps to constrain different aspects of retargeting performance in optimization, which obtains better performance than using a single energy map.

2. RETARGETING USING MULTI-MAP CONSTRAINED REGION WARPING
Figure 1 illustrates an overview of our approach. First, the original image (Figure 1(a)) is represented with curve-edge trapezoid meshes based on region decomposition (Figure 1(b)). Then, two kinds of energy maps, importance map and sensitivity map (Figure 1(c)), are calculated to constrain important content emphasis and visual distortion reduction respectively. Under the energy map constraints, the regions are warped by optimally relocating the mesh vertexes (Figure 1(d)). Finally, the target image is generated from the warping result (Figure 1(e)).

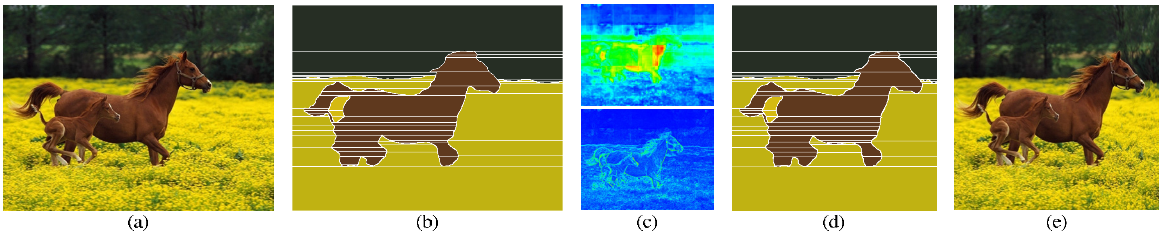

Figure 1. An overview of the proposed approach. (a) Original image. (b) Mesh representation of original image. (b) Importance energy map (top) and sensitivity energy map (bottom). (d) Mesh representation of target image. (e) Target image.

2.1 Mesh Representation

Mesh representation is used to control image structure in region level retargeting methods. The existing uniform grid mesh [8] and content-aware triangle mesh [9] can only approximately represent region boundaries with small segments and generate redundant mesh vertexes in homogeneous regions. It causes the difficulty in region boundary preservation and high computation cost.

To address these problems, we propose a curve-edge trapezoid mesh representation which makes use of the homogeneous region boundaries in mesh generation. Here, we only discuss the mesh representation for horizontal resizing. The mesh representation for vertical resizing can be generated similarly. First, we decompose the original image into homogeneous regions (Figure 2(a)) by Mean-Shift algorithm [10]. To each region, we detect the region boundary and decompose it into edge segments, in which the *y*-coordinates of points change monotonically. As shown in Figure 2(b), we select the endpoints and the local most left/right points of all edge segments as the key points (red and yellow points). To each key point inside image (red points), two auxiliary points with the same *y*-coordinate are found on the nearest edge segment(s) (blue points) or image boundary/boundaries (green points). And the key points on image boundary (yellow points) only have zero or one auxiliary point. With the key/auxiliary points and region/image boundaries, the original image can be represented with curve-edge trapezoid meshes (Figure 2(c)).

One advantage of our method is the computation cost of retargeting can be greatly reduced when the image structure is simple. In Figure 2(b), only 10 mesh vertexes (red points) need to be relocated. And the number of key point can be further reduced by region boundary smoothing. Another advantage is the distortion of region boundaries can be better controlled by key point relocation than representing the boundaries with small segments. More details will be discussed in section 2.3.

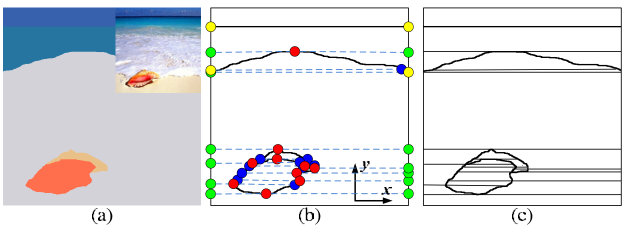

Figure 2. Curve-edge trapezoid mesh representation. (a) Region decomposition. (b) Mesh generation. (c) Mesh representation.

2.2 Multiple Energy Maps Computation

Different aspects of retargeting performance are usually influenced by different factors. However, all the existing methods only utilize a single energy map to constrain one aspect of performance in retargeting, which may cause some problems in the other unconstrained aspects. For example, seam carving only uses gradient in energy map to reduce visual distortion, and retains important objects depending on their high gradient boundaries. This strategy may lead to the serious distortion when objects' boundaries are broken [4]. A simple solution is combining all the factors into one energy map. Unfortunately, the generated energy map confuses the effects of different factors and may not lead to a better result. For example, when combine content importance with gradient in energy map, seam carving retains the important objects but breaks the strong edges in background [9].

In this paper, we propose a novel strategy that utilizes multiple energy maps to constrain different aspects respectively. Two kinds of energy maps, importance map and sensitive map, are generated from the saliency and face detection based visual attention model [5] and weighted gradient map [8] respectively. The former energy map is used for important content emphasis, and the later is for visual distortion reduction. Note here, different gradient maps are used for distortion reduction in horizontal resizing and vertical resizing. Of course, more energy maps can be adopted to constrain other aspects in retargeting. Figure 3 shows the results generated with different energy map(s). We can find the result using multiple energy maps achieves a better performance.

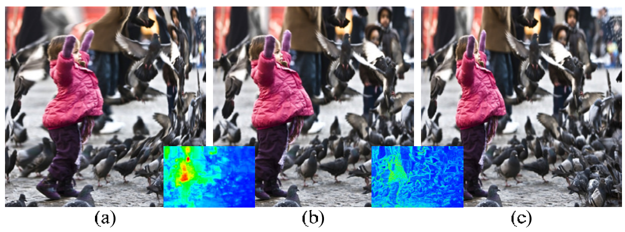

Figure 3. Results using different energy map(s). (a) Only using importance energy map. (b) Using importance energy map and sensitivity map. (c) Only using sensitivity map.

2.3 Constrained Region Warping

With the curve-edge trapezoid mesh representation and multiple energy maps constraints, we retarget the image by optimally warping the meshes. Two aspects, mesh area change and mesh shape distortion, are considered to realize important content emphasis and visual distortion reduction respectively.

Figure 4 shows an example of horizontal mesh warping. In following discussion, we represent the original coordinates of each point p_{ij} with (x_{ij}, y_{ij}) and the new coordinates after warping with (x'_{ij}, y'_{ij}), here $x_{ij} = i$ and $y_{ij} = j$. For only consider horizontal resizing, we assume the *y*-coordinate of each point is

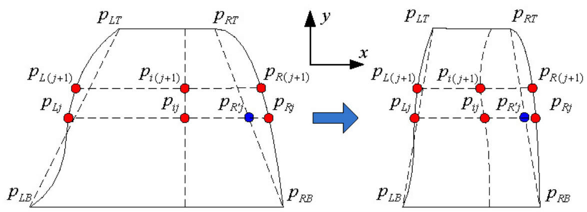

Figure 4. Horizontal warping of curve-edge trapezoid mesh.

unchanged in warping. To each point on the left/right edge, such as p_{Rj}, we first find a point $p_{R'j}$ with the same y-coordinate value same to p_{Rj} on the line segment between the two endpoints of the edge. Then, we constrain the x-coordinate difference between p_{Rj} and $p_{R'j}$ should be in proportion to the x-coordinate difference between p_{RT} and p_{RB} in warping. If $x_{RT} = x_{RB}$, we set $x'_{Rj} = x'_{RT} = x'_{RB}$. This constraint ensures the x-coordinates of the points on the same edge segment change consistently in warping. For $y_{Rj} = y_{R'j}$ and $p_{R'j}$ is on the line segment between p_{RT} and p_{RB}, the new x-coordinate of p_{Rj} can be represented as follows:

$$x'_{Rj} = \frac{x_{Rj} - x_{RB}}{x_{RT} - x_{RB}} x'_{RT} + \frac{x_{RT} - x_{Rj}}{x_{RT} - x_{RB}} x'_{RB}. \tag{1}$$

Similarly, the new x-coordinate of p_{Lj} can be represented as a linear combination of x'_{LT} and x'_{LB}. Then, the mesh area after warping can be calculated as follows:

$$S' = \sum_{j=y_{RB}}^{y_{RT}} (x'_{Rj} - x'_{Lj}). \tag{2}$$

We represent the mesh area changes in warping by a penalty function as follows, and realize important content emphasis by minimizing the penalty function:

$$f_C = \sum_k e_k^I (S'_k - S_k)^2, \tag{3}$$

where S_k and S'_k are the areas of the kth mesh before and after warping; e_i^I is the mean value of importance energy of each pixel within the kth mesh. With this optimization, the meshes with more important content are encouraged to have less area changes.

Besides mesh area change, we also consider the mesh shape distortion for reducing visual distortion. Assume the x-coordinates of the points in a row within a mesh are changed uniformly in warping, the new x-coordinate of each point can be represented as a linear combination of the new x-coordinates of the mesh vertexes. We define the mesh shape distortion as the sum of the horizontal distance change of each point within the mesh to its upper point, and pay more attention to mesh contour distortion:

$$D = \sum_{j=y_{RB}}^{y_{RT}} \sum_{i=x_{Lj}}^{x_{Rj}} e_{ij}^{HS} e_{i(j+1)}^{HS} \left(x'_{ij} - x'_{i(j+1)} \right)^2 + D_L + D_R, \tag{4}$$

where e_{ij}^{HS} is the sensitivity energy values of p_{ij} for horizontal resizing; D_L and D_R are the penalties for the distortion of the left edge and the right edge respectively, which are calculated as:

$$D_\square = \sum_{j=y_{\square B}}^{y_{\square T}} e_{\square j}^{HS} \left(\left(x'_{\square T} - x'_{\square B} \right) - \left(x_{\square T} - x_{\square B} \right) \right)^2, \tag{5}$$

where $\square$ denotes "L" or "R". Then we define the penalty function of mesh distortion as follows, and realize the visual distortion reduction by minimizing the penalty function:

$$f_D = \sum_k D_k, \tag{6}$$

where D_k is the shape distortion of the kth mesh.

In retargeting, we wish to optimize the important content emphasis and visual distortion reduction simultaneously, and define the objective function as:

$$\min f_C + f_D. \tag{7}$$

It is a quadric programming problem of the x-coordinates of the mesh vertexes. As shown in Figure 2(b), in the optimization, the mesh vertexes on image boundary (green and yellow points) are constrained to be on the target image boundary, and the inner auxiliary points on region boundaries (blue points) are represented by the corresponding key points (red points). Moreover, the horizontal relative relationship between two mesh vertexes on an edge is required to be retained in warping, which can avoid the pseudo-solution of the optimization problem.

For our method treats the horizontal resizing and vertical resizing separately, the resizing order should be considered in retargeting. We examine our approach by experiments and find it robust to resizing order (Figure 5). We also consider resizing direction in retargeting, for shrinking the image in one direction and stretching it in another direction may lead to the same aspect ratio but totally different performance [8]. We assume the retaining ratio of image content is in proportion to its importance energy and calculate a scaling ratio $\rho = \min\left\{ \max_j \sum_i e_{ij}^I / W', \max_i \sum_j e_{ij}^I / H' \right\}$. Here, e_{ij}^I is the importance energy of p_{ij}, W' and H' are the required width and height of the result. After a preliminary uniform scaling with ratio ρ, the scaled image is shrunken or stretched in each direction as necessary. This strategy can effectively avoid the serious distortion of important content.

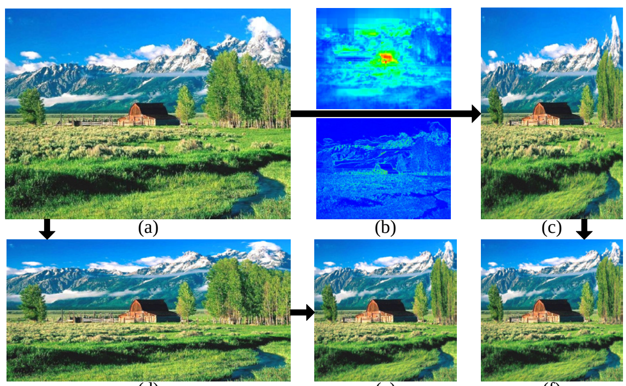

Figure 5. Results with different resizing orders. (a) Original image. (b) Importance map (top) and sensitivity map for horizontal resizing (bottom). (c) and (f) Results of horizontal-vertical resizing. (d) and (e) Results of vertical-horizontal resizing.

3. EXPERIMENTS

To illuminate our performance, we implement our approach and test it in several aspects. Figure 5 shows the target images generated from the same original image but with different resizing orders. Experiment shows our approach is robust to the order of horizontal and vertical resizing and suitable for retargeting an image to various aspect ratios.

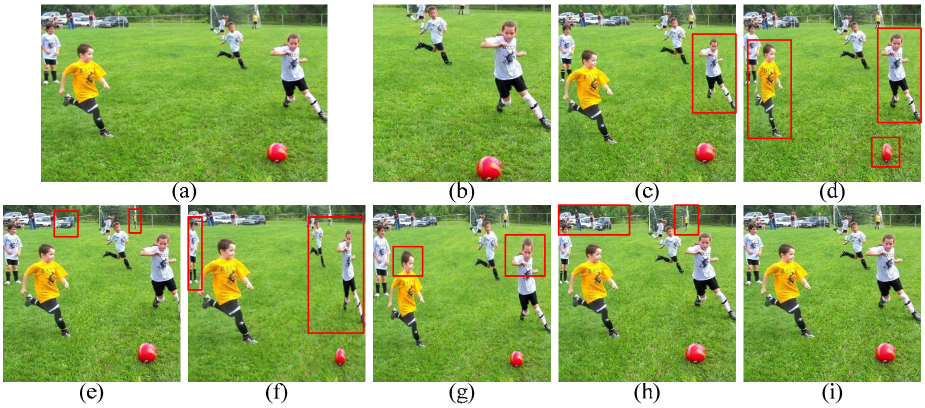

Figure 6. Example of comparison with the existing typical methods. (a) Original image. (b) Result of cropping. (c) Result of recomposition. (d) Result of improved seam carving. (e) Result of non-homogeneous warping. (f) Result of fisheye-view warping. (g) Result of scale-and-stretch. (h) Result of mesh parametrization. (i) Our result.

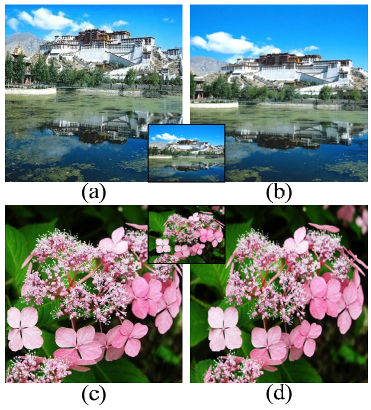

Figure 7. Examples of bad results. (a) and (c) Our results. (b) Result of cropping. (d) Result of uniform scaling.

To further evaluate the effectiveness of our method, we carry out a user study in comparison with the existing typical methods. Thirty images with various styles are used, in sizes from 800*600 to 960*600. Seven existing methods are selected for comparison and all target images are required in size of 300*300. Figure 6 shows an example of comparison. We can see that cropping [5] cannot preserve all important objects and recomposition [3] has the problem in object spatial relationship. Improved seam carving [6] and scale-and-stretch [8] destroy the structure of objects for they have low gradient. Non-homogenous warping [4] causes object self-intersection. Fisheye-view warping [7] distorts the unfocused objects. Mesh parametrization [9] distorts the region near the objects. Compared with them, our algorithm obtains a more comparable result. Table 1 shows the result of user study. Twenty-one participants in age of 20 to 45 are invited, including student, officers and company employees. In evaluation, each our result is required to compare with any other corresponding target image by three participants, and the dominant judgment is treated as the final evaluation. We can find that our method is better than most existing method and has a very similar performance to mesh parametrization. However, the efficiency of our method is better, for two methods both use constrained quadratic programming and our variable number is much smaller.

Table 1. Result of user study.

	Better	Similar	Worse
Cropping	25	2	3
Recomposition	22	6	2
Improved seam carving	17	10	3
Non-homogeneous warping	19	9	2
Fisheye warping	23	6	1
Scale-and-stretch	16	12	2
Mesh parametrization	5	23	2

In experiment, we also find some limitations of our approach. For example, while avoiding object shape distortion is more important than retaining background content, our method (Figure 7(a)) leads to a worse result than cropping (Figure 7(b)). And when keeping image layout is more important than emphasizing the object(s) and reducing visual distortion, the result of uniform scaling (Figure 7(d)) is more acceptable than ours (Figure 7(c)).

4. CONCLUSION

This paper presents a novel image retargeting method using multi-map constrained region warping. Important content emphasis and visual distortion reduction are simultaneously optimized under constraint of importance map and sensitivity map respectively. Moreover, the proposed curve-edge mesh representation usually requires much less vertexes to represent the image, and thus reduce the computation cost efficiently. In the future, our work will focus on improving high level constraints to overcome the limitation and extending our method to video.

5. ACKNOWLEDGMENTS

We would like to thank Yu-Shuen Wang and Yanwen Guo for providing the experiment results. This work is supported by Grant 60533080 of the National Natural Science Foundation of China and Grants A-PA9F of the Hong Kong RGC Direct Allocation.

6. REFERENCES

[1] H. Liu, S. Jiang, Q. Huang, C. Xu, and W. Gao. Region-based visual attention analysis with its application in image browsing on small displays. *ACM MM*, Germany, 2007.

[2] T. Ren, Y. Liu, and G. Wu. Image retargeting based on global energy optimization. *ICME*, USA, 2009.

[3] V. Setlur, S. Takagi, R. Raskar, M. Gleicher, and B. Gooch. Automatic image retargeting. *Int'l Conf. Mobile and Ubiquitous Multimedia*, New Zealand, 2005.

[4] L. Wolf, M. Guttmann, and D. Cohen-Or. Non-homogeneous content-driven video-retargeting. *ICCV*, Brazil, 2007.

[5] L.-Q. Chen, X. Xie, X. Fan, W.-Y. Ma, H.-J. Zhang, and H.-Q. Zhou. A visual attention model for adapting images on small displays. *Multimedia Systems*, 9(4):353-364, 2003.

[6] M. Rubinstein, A. Shamir, and S. Avidan. Improved seam carving for video retargeting. *SIGGRAPH*, USA, 2008.

[7] F. Liu and M. Gleicher. Automatic image retargeting with fisheye-view warping. *Symp. UIST*, USA, 2005.

[8] Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee. Optimized scale-and-stretch for image resizing. *TOG*, 27(5), 2008.

[9] Y. Guo, F. Liu, J. Shi, Z.-H. Zhou, and M. Gleicher. Image retargeting using mesh parametrization. *TMM*, 2009.

[10] D. Comaniciu, and P. Meer. Mean shift: a robust approach toward feature space analysis. *TPAMI*, 24(5):603-619, 2002.