

Study on the Image Retargeting by Using Semantic Concepts

Lei Wang and ChunLei Chen

School of Computer Engineering Weifang University Weifang 261061 China
wanglpqq@gmail.com, chunlei.chen@wfu.edu.cn

Abstract

Image retargeting aims to change the resolution and aspect ratio of an image to fit it into different devices. The key and most challenging issue for this task is how to balance the tradeoff between preserving the important contents and minimizing the visual distortions. In this paper, we present an effective image retargeting method called concepts-based mesh parameterization. Specifically, we first construct a structure-preserved mesh image. Based on this mesh representation, the image retargeting problem is formulated into a constrained image mesh parameterization problem, finding a homomorphous target mesh with the desired device size. To emphasize different concepts and minimize visual distortion, different concepts are first detected from the image, and are associated into the image mesh. Then the problem is transformed into solving a linear equations system. The target image is rendered by texture mapping. Experiments demonstrates the effectiveness of our concept-based algorithm.

Keywords: *Sematic Concept, Image Retargeting, Texture Mapping, Saliency Map*

1. Introduction

Due to the fast growing diversity of display devices, such as cellular phones and PDAs, there comes the need that images must be adjusted across various setting like resolution or aspect ratios. The problem of adapting an image to various target screens is defined as image retargeting. A common solution to image retargeting is to uniformly rescale the original image according to the target screen size. This naive scaling is problematic (see Figure 1). The important objects in image maybe become too small to be recognized. Besides, this generally changes the overall image structure and often leads to visual distortion in the target image. For example, a straight line may become badly curved in the target image, if different parts of it happen to have different importance. Such structural distortion often causes the target image look hardly natural. Image cropping method like [1, 2] could avoid the distortion issue, but lose the image content information. When there exist multiple important objects that are far away from each other, cropping inevitably loses some of them in order to keep a necessary resolution of the selected object. Therefore careful and proper treatment is required to minimize such distortion while preserving the important contents. How to balance this tradeoff is the key and most challenging issue for content-aware retargeting.

To handle the distortion problem of image retargeting, many solutions have been proposed. A comprehensive introduction of the recent development in this area is presented in [3], where the existing methods are categorized into two types. The first is discrete methods, including seam carving [4, 5] and shift-maps [6, 7]. This type of methods try to remove or copy unimportant pixels while keeping the important ones rigid. The seam carving method achieves image resizing by iteratively carving less noticeable seams [4]. Although it tries to remove less noticeable seams, it still produces artifacts like breaking objects since image structures have not been explicitly considered. The second is continuous methods, including feature aware texture mapping [8, 9], scale-and-stretch [10], and energy-based deformation [11]. These methods try to compute a non-uniform

warping function from the original to the target image, which is designed to retain the important contents and warp the unimportant regions. Compared to the Seam Carving Method, as variables degrees of freedom decrease a lot, continuous methods reach higher computational efficiency. Distortion is also controlled and distributed better by operating on triangle scale. In order to emphasize important content, saliency map and other image cues (like edges, face) are merged to generate an important map. Accordingly the mesh is encouraged to morph in such a way that the size of an important mesh cell (triangle) is reduced less than non-salient ones. Besides, they view deformation in an energy minimization perspective and solve the problem iteratively, there still remains room for computational efficiency improvement.

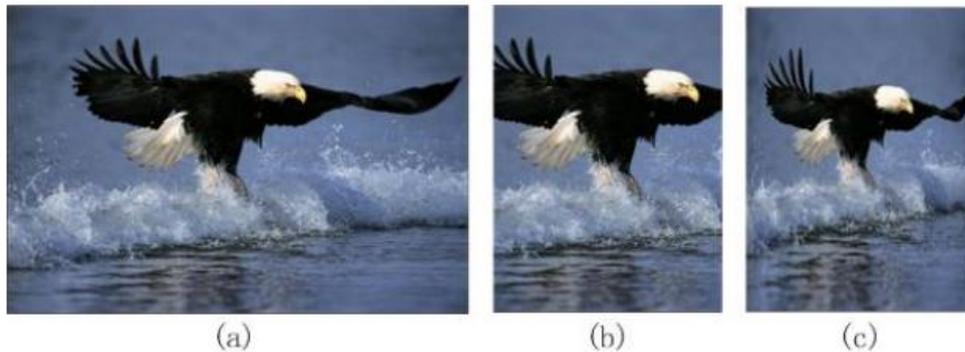


Figure 1. An Example of Image Retargeting. From Left to Right: Input Image (a), Cropping (b) and Resized using Scaling(c)

Effective image retargeting should emphasize important content while retaining surrounding context with minimal visual distortion. The challenge is to minimize the visual distortion. In this paper, we propose a concept-based mesh parameterization. Similar to other continuous methods, we define image retargeting as a mesh parameterization problem. The image is represented by a mesh that is consistent with the underlying image structures, and then adapt images via mesh transformation. This mesh representation enables us to easily preserve image structures during retargeting. Also adapting images via transforming mesh ensures smooth image transformation. Compared to previous work, our method has two main contributions: 1) Instead of encoding preservation of salient objects and image structure as constraints during mesh parameterization, the concepts of image is adopted to define the importance of the mesh. Different concepts will add different constraints on the mesh reshape optimize. 2) Without lose of retargeting quality, we transform the mesh parameterization problem into a linear system only once, which is more efficient than existing methods that use iterative solver. The rest of this paper is organized as follows. A brief overview of related work is given in section II. In section III, we provide our concepts-based retargeting algorithm in detail. Experiment results are shown in section IV, followed by conclusion in the last section.

2. Related Work

Almost all image browsers on small-screen devices provide scaling functionality. A large image is uniformly down sampled to fit the target screens. Important objects in the image are often small and difficult to identify. To account for different aspect ratios, either the distortion caused by aspect ratio change is introduced or black letter box is used to fill in the blank space, thus wasting the precious space. Numerous algorithms have been proposed for media retargeting across various settings such as aspect ratios. To maintain the overall structural consistency and better present the important contents, content-aware methods has become the main stream for media retargeting. These methods

usually rely on image analysis, such as attention model extraction and object detection, to identify the important region.

Seam carving [4] compute the importance map from the image saliency map, and detect seams that with less important contents. They try to resizes an image by reducing or adding one seam at each iteration. Each seam consists of a continuous chain of the least important pixel from each row or column so that the carving operation would not alter the important contents. Ren, et al., [13] use an edge map to define important content information. The edge map is built by combining artificial edges connecting different salient objects, boundaries from canny operator and mean shift segmentation. Wang et al. [14] defines each pixels importance as the result of multiplying an image saliency map and a gradient magnitude map. Setlur, *et al.*, [15] proposed to segment the prominent object which is most likely the salient object and to re-compose them onto resized background. [8] Asks user to specify important regions and try to compute a continuous warping function from the original image to the target. The warping is non-uniform in such a way that the important contents receive little changes while the unimportant areas suffer the most distortion. Eye tracking systems are also used to infer the important region [16].

The existing methods normally use image edge information or image saliency directly to retain important pixels. They have achieved great success respectively. However, the cues they used are all low-level information. In this paper, we investigate how high level cues affect the image retargeting and present a concept-based image retargeting algorithm.

3. Image Retargeting

Image retargeting tries to convert a source image to a target image adapted to display in a smaller size and different aspect ratio device. Similar to [12], we also formulate image retargeting as a mesh parameterization problem that emphasizes the important content while retaining the main structure (like boundaries, straight lines). Different from [12], instead of using salient objects to define the important map, we adopt semantic concepts map to construct structure constraints. The same structure can have different importance's when it appear in different semantic concept. For example, a straight line in car must have big affection on the final distortion than that in the grass. The outline of our algorithm is illustrated in Figure 2. To emphasize different semantic area, we first calculate a semantic concept map of the source image (Subsection III-A). Then we build a controlling mesh of the input image that is consistent with the underlying image structures, and associate semantic information with the source mesh (Subsection III-B). In order to improve visual quality of the target image, we add as set of structure constraints in different semantic area (Subsection III-C). Afterwards, the target mesh is by solving a linear system. Retargeting result is finally rendered using the standard texture mapping algorithm.

3.1. Multi-class Segmentation

Unlike object recognition methods that aim to find a particular object, human also have the ability to partition or segment an entire image into distinct recognizable regions. The similar task in computer vision is called multi-class image segmentation. In our task, our goal is to emphasize different semantic area in an entire image, thus need to classify all pixels in an image which could benefit from multi-class image segmentation. Most multi-class segmentation methods take into account local (pixel or region) appearance, which ignores the visual consistent of neighbor areas, generating many small wrong classified areas. Gould, *et al.*, [17] achieves this goal by constructing a conditional Markov random field (CRF) over the image the encodes local and pair wise probabilistic preferences. Finding the most probable segmentation is then equivalent to optimizing the energy defined by this CRF energy function. Besides, they also incorporate contextual

information that captures spatial relationships between classes in the framework. Two-stage framework is proposed to leverage the relative location of the different object classes. For each pixel is first predicted using a boosted classifier trained on standard appearances-based features. Then, the prediction is combined with recomputed relative location maps to form a relative location feature. This feature is then incorporated into a unified model that combines appearance and relative location information to make a final prediction. The proposed method removes the small wrong-classified area (see Figure 3), and achieves the results that are above state-of-the-art. In our method, the multi-class segmentation from [17] is adopted.

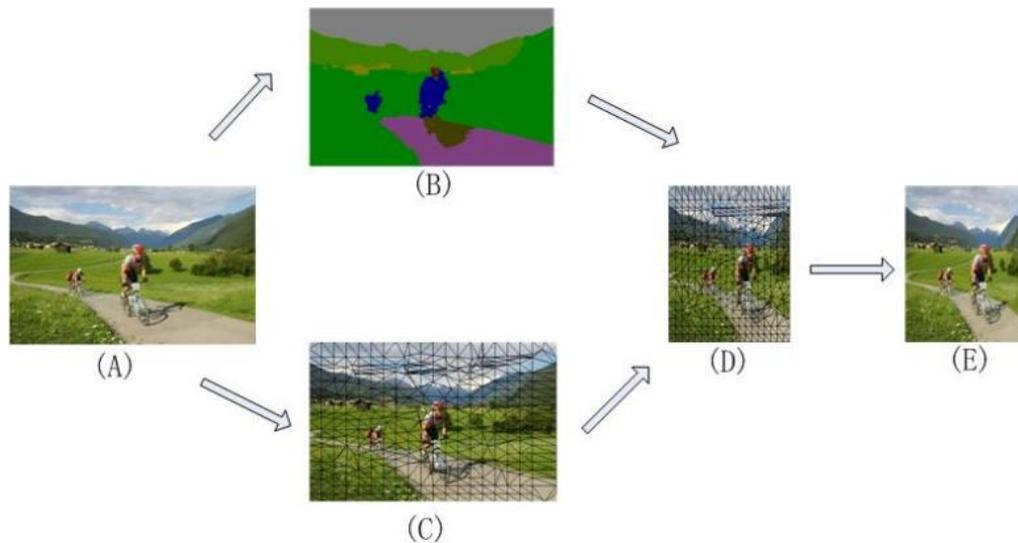


Figure 2. Algorithm Overview. Our Algorithm First Builds Semantic Map (b) and a Feature Relevant Controlling Mesh (c) from the Input (a). The Source Mesh is then Associated with Concepts Information, and the Target Mesh with Desired Resolution is Produced by Solving a Mesh Parameterization Problem (d). Standard Texture Mapping is Finally used to Render Target Image (e)

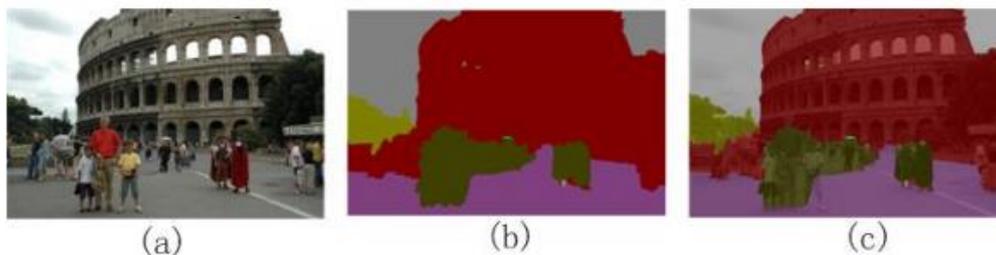


Figure 3. Multi-class Segmentation Result from [17]

3.2. Mesh Representation

The key issue of generating a mesh from an image is to generate a mesh that is consistent with the input image structures. Similar to [17], we first detect feature points from the input image is, and then use Delaunay Triangulation [18] algorithm to generate the mesh as follows:

- 1) The input image boundary is evenly discretized, and all the points are used as part of the feature points.
- 2) Edge detector like canny is used for extracting some feature points.
- 3) Grid based points are added to make the mesh nearly uniform density.
- 4) Constrained Delaunay Triangulation algorithm [18] is used to generate a feature consistent mesh (Figure 4)

In order to treat image regions discriminatively, the source mesh is associated with the semantic concept map. Each categories in the semantic concepts is assigned an important weight manually. With these associated concept information, we can perform mesh parameterization in such a way that certain important areas shrink less than those with less importance.

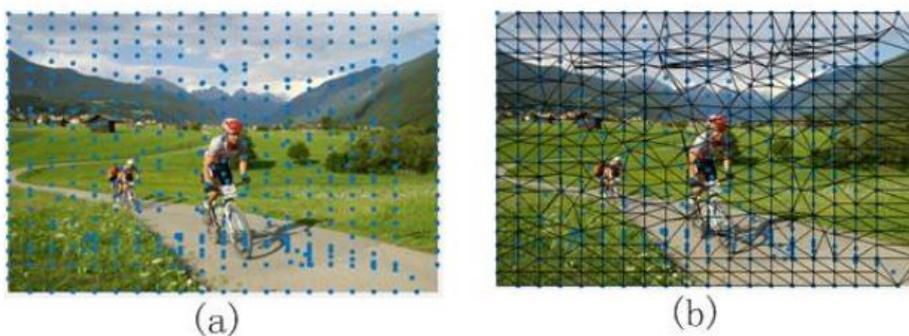


Figure 4. Mesh Generation. (a) Input Image with Feature Points Evenly Distributed. (b) Delaunay Triangulation Result

3.3. Structure Constraints

Image retargeting should preserve important areas that are immune from even slight distortion. Furthermore, images structures in the different areas, which show in terms of strong edges, are also important visual features. They should be retained as-rigid-as possible. To achieve these goals, several constraints including boundary constraint and strong edge constraint are defined.

1) Boundary constraint: In the process of image retargeting, the boundary points on the source image should also on the boundary of the target image. For each corner points on the up and bottom boundaries, the v - of the boundary points of the target mesh are fixes as the corresponding target image's up and bottom boundary respectively. Their u -coordinates are to be solved by energy function. Similarly, for those corner points on the left and right boundary, their u -coordinates are fixed while solving for their v -coordinates.

Suppose that the target image's up, bottom, left, and right boundary points are Q_U , Q_B , Q_L and Q_R respectively. Correspondingly, target image's up and bottom y -coordinates are y_U and y_B . The left and right x -coordinates are x_L and y_R . The above boundary constraints can be expressed as:

$$F_B = \sum |u_U - y_U| + \sum |u_B - y_B| + \sum |u_L - x_L| + \sum |u_R - x_R| \quad (1)$$

2) Strong Edge constraint: Another important visual features are strong edges, and they are vital clues for understanding image content. They should be maintained as-rigid-as possible. We detect those strong edge segments using Hough transform.

After receiving segments from Hough transform. The segments are further filtered out by requiring that the length of each segment should exceeds a given threshold. Furthermore, only one segment is reserved if several segments are close enough.

If the points lie in the lines of the source image, we require that the same points in the target image should still pass through a line (linearity constraint). Let's consider the retargeted lines as:

$$RL_k = \bigcup Q_{k,i} | i = 1, \dots, m_k. \quad (2)$$

Then the retagged line RL_k could be defined by the expression $y = a_k * x + b_k$. Here a_k and b_k could be easily represented by two points in RL_k . And other points should satisfies the line function mentioned above.

However, not all the line segments could apply the linearity constraint, as there are also some other constraints, and it is also not necessary that all the line segments should obtain the rule (like the lines in the sky). So we define the following soft energy constraint and encourage meet of the above linearity constraint as much as possible. Besides, in different semantic areas, a weight is assigned to force them to meet the linearity constraint.

$$E_{se} = \sum_{c=1}^C (\omega_c * \sum_{i=1}^{m_k} (v_{k,i} - a_k * U_{k,i} + b_k)^2) \quad (3)$$

3) Neighborhood constraint: Another important constraint is neighborhood constraint. The neighborhood constraint means if the two points are neighbor, their shift distance in target images should not change too much. Assume that a point p_i and its neighbor point p_j . The shift distances in the targeted image are Δp_i and Δp_j . Then the constraint could be defined as:

$$E_n = \sum_{c=1}^C (\omega_c * \sum_{j \in \text{neig}(i)} (\Delta p_i - \Delta p_j)^2) \quad (4)$$

Also, in different semantic areas, the weight is assigned. For certain important areas, a higher weight is assigned to restrict the constraint, and for other areas, a lower weight is used to loose the constraint. All the constraints are combined together, and the image retargeting problem is formulated as solving the following problem:

$$\min(E_{se} + E_n) \text{ s.t. } F_B = 0. \quad (5)$$

As each item of the energy function is square errors, the energy function could be transformed into a conditioned base linear problem $Ax = b$, s.t. $Cx = 0$, which could be solved in an easily way.

After we get the target 2D triangular mesh, we just need to map the input image onto the mesh to get the final retargeted image. The input image acts as the texture image and the locations of points before deformation are the texture coordinates. The standard texture mapping method is trivial especially when such operation is accelerated by a common PC or cellular phone graphics chip. Our method runs automatically and reaches high computational efficiency as we just need to solve a linear system only once.

4. Experiments

We validate our method on a variety of images. Besides, we also compare our method with the state-of-the-art methods. Each image is retargeted to half of its original width with its height kept unchanged. The methods we compare with include Seam Carving [4], Quadratic programming [19], and Scale-and Stretch [10]. Some representative results are shown in Figure 5.

Seam carving method tries to remove the seam which is defined as the less important line from top to bottom. While selecting seam, the structure of the image is not considered, thus the structure shape cannot be retained. From Figure 5, the first image shows that our method prevents road shrink, which is better than seam carving method. Scale-and-Stretch method scales the salient object in different scale, keeping the salient region ratio well. However, if there is hard divider in background, it tries to change the divider ratio significantly (see first image in Figure (5)). Quadratic programming method adds salient objects constraint in framework to resize image. However, some humans in the background also be ignored as they are seen as background, which is not necessary. In our method, the semantic areas like humans, and bicycles are very well kept as well as shadows on the ground. The rest of pictures shows that our method keeps semantic areas (like humans, fish) from obvious. In general, compared to the existing approaches, our method can keep integrality of the image content, prevent significant objects (not only salient objects) from distortion and distribute distortion across the image better.

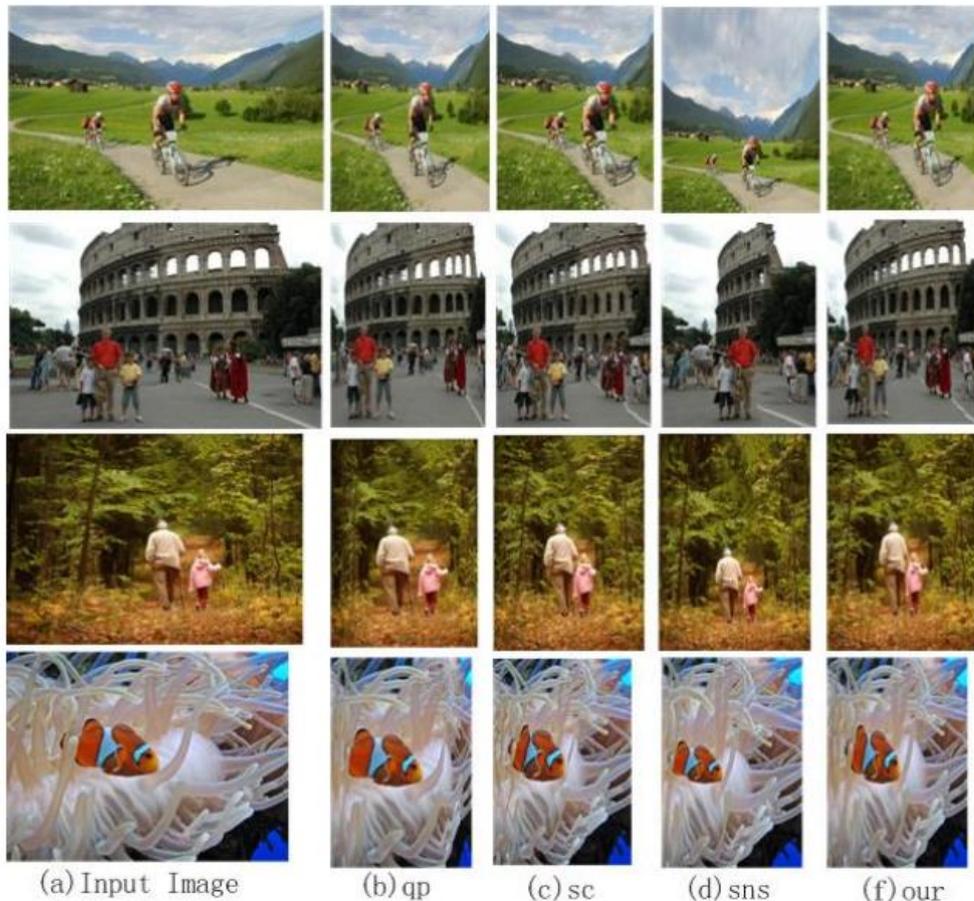


Figure 5. Visual Comparison of Retargeted Images using Quadratic Programming (qp)[19], Seam Carving (sc)[14], Scale-and-Stretch (SNS)[10] and our Method

5. Conclusions

In this paper, we propose an effective image retargeting method which fit an image to a new resolution and aspect ratio while persevering important semantic areas. Specifically, we first construct a structure-preserved mesh image. Based on this mesh representation, the image retargeting problem is formulated into a constrained image mesh parameterization problem, finding a homomorphous target mesh with the desired device size. To emphasize different concepts and minimize visual distortion, different concepts are first detected from the image, and are associated into the image mesh. To maintain the structure of the images, several constraints like boundary constraint, strong edge constraint and neighborhood constraint are introduced. Then the problem is transformed into solving a linear equations system. Finally, texture mapping is used to generate the retargeted image. Experiment results show that our method can generate high quality retargeting results that persevering important semantic areas. At current time, we manually assign weights for different semantic areas, which is ad hoc. In the future, we could learn these weights from large data. Also, we could incorporate more constraints like similar area constraint which enforces similar area having same scale ratio.

Acknowledgments

Our work is supported by the Science and Technology Development Program of Shandong (No. 2012G0020127), and the Science and Technology Development Program of Weifang (No. 2015GX008).

References

- [1] M. Nishiyama, T. Okabe, Y. Sato and I. Sato, "Sensation-based photo cropping," Proceedings of the 17th ACM international conference on Multimedia, (2009), pp. 669-672.
- [2] 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, vol. 9, no.4, (2003), pp.353-364.
- [3] M. Rubinstein, D. Gutierrez, O. Sorkine and A. Shamir, "A comparative study of image retargeting", ACM Transactions on Graphics, vol. 29, no. 6, Article No.160, (2010).
- [4] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing", ACM Transaction on Graphics, Article No. 10, vol. 26, no. 3, (2007).
- [5] M. Rubinstein, A. Shamir and S. Avidan, "Improved seam carving for video retargeting", ACM Transaction on Graphics, Article No.16., vol. 27, no. 3, (2008).
- [6] Y. Pritch, E. Kav-Venaki and S. Peleg, "Shift-map image editing", Proceedings of the 12th International Conference on Computer Vision, (2009), pp. 151-158.
- [7] Y. Hu and D. Rajan, "Hybrid shift map for video retargeting", Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition, (2010), pp. 577-584.
- [8] R. Gal, O. Sorkine and D. Cohen-Or, "Feature-aware texturing", Proc. 17th Eurographics Conf. on Rendering Techniques, (2006), pp. 297-303.
- [9] L. Wolf, M. Guttman and D. Cohen-Or, "Non-homogeneous content driven video retargeting", Proceedings IEEE 11th International Conference on Computer Vision, (2007), pp.1-6.
- [10] Y.-S. Wang, C.-L. Tai, O. Sorkine and T.-Y. Lee, "Optimized scale and stretch for image resizing", ACM Transaction on Graphics, vol. 27, no. 5, (2008), pp. 118.
- [11] Z. Karni, D. Freedman and C. Gotsman, "Energy-based image deformation", Proceedings of the Symposium on Geometry Processing, (2009), pp. 1257-1268.
- [12] Y. Guo, F. Liu, J. Shi, Z.-H. Zhou and M. Gleicher, "Image retargeting using mesh parameterization", IEEE Transactions on Multimedia, (2009), vol. 11, no. 5, pp. 856-867.
- [13] T. Ren, Y. Guo, G. Wu and F. Zhang, "Constrained sampling for image retargeting", 2008 IEEE International Conference on Multimedia and Expo, (2008), pp. 1397-1400.
- [14] P. Kaufmann, O. Wang, A. Sorkine-Hornung, O. Sorkine-Hornung, A.Smolic and M. Gross, "Finite element image warping", Computer Graphics Forum, vol. 32, no. 2, (2013), pp. 3139-2013.
- [15] V. Setlur, S. Takagi, R. Raskar, M. Gleicher and B. Gooch, "Automatic Image Retargeting", Proceedings of 4th International Conference on Mobile and Ubiquitous Multimedia, (2005), pp. 59-68.
- [16] A. Santella, M. Agrawala, D. DeCarlo, D. Salesin and M. Cohen, "Gaze-based interaction for semi-automatic photo cropping", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, (2006), pp. 771-780.
- [17] S. Gould., R. Jim, C. David, E. Gal and K. Daphne, "Multi-Class Segmentation with Relative Location Prior", International Journal of Computer Vision, (2008), vol. 80, no. 3, pp. 300-316.

- [18] R. Seidel, "Constrained delaunay triangulations and voronoi diagrams with obstacles", Technical Report 260, (1980), Inst. for Information Processing.
- [19] R. Chen, D. Freedman, Z. Karni, C. Gotsman and L. Liu, "Content aware image resizing by quadratic programming", Proceedings of Computer Vision and Pattern Recognition Workshops (CVPRW), (2010), pp. 1-8.

