Pacific Graphics (2011)
Bing-Yu Chen, Jan Kautz, Tong-Yee Lee, and Ming C. Lin (Editors)

Short Papers

Feature-Aware Mesh for Image Retargeting

Yuzhen Niu[1], Feng Liu[1], and Yu-chi Lai[2]

[1]Department of Computer Science, Portland State University
[2]Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology

Abstract

Image retargeting adapts an input image to displays other than originally intended while minimizing visual distortion. Most warping-based methods formulate retargeting as a spatially-varying warping problem over a uniform mesh and encourage each mesh grid to undergo a conformal transformation to minimize visual distortion. However, the conformal transformation constraint on each individual mesh grid is often insufficient to preserve global image structures. In this paper, we present a feature-aware mesh based retargeting method. Our idea is to warp each visually salient image feature as a whole, thus preserving global image structures. Our method divides an image into a non-uniform mesh such that each salient image feature or object is enclosed only in one mesh element. Our method warps each mesh element with an affine transformation. We then formulate image retargeting as a spatially-varying affine warping problem. To further minimize the distortion, we encourage each affine transformation to be close to more distortion-free ones, such as conformal transformation. The spatially-varying affine warping problem is formulated as a quadratic energy minimization problem, which can be solved efficiently. Our experiments demonstrate that our method can effectively preserve global image structures.

Categories and Subject Descriptors (according to ACM CCS): I.4.9 [Computer Graphics]: Image Processing and Computer Vision—Applications

1. Introduction

Image retargeting adapts an image to displays other than originally intended. A significant amount of effort has been devoted to this problem recently [RGSS10]. Existing methods aim to fit an input image to a new size while minimizing visually objectionable distortion. While successful solutions have been developed to avoid local distortion, it is still challenging to preserve global image structures.

In this paper, we present a new warping-based retargeting method to address this problem. Existing warping-based methods first divide an input image into a uniform grid mesh (c.f. [GSCO06,WGCO07,WTSL08,ZCHM09]). Then they formulate image retargeting as a mesh warping problem. These methods encourage each mesh grid to undergo a conformal (similarity) transformation to avoid objectionable distortion. While these methods can successfully achieve a smooth transformation, they sometimes cannot preserve global image features. The conformal transformation constraint on an individual grid cannot guarantee the conformal transformation of a whole object that contains multiple grids. Moreover, existing methods use the saliency value of each grid to guide the warping. Each grid inside the same object often takes a different saliency value, which will also lead to inconsistent warping inside the object.

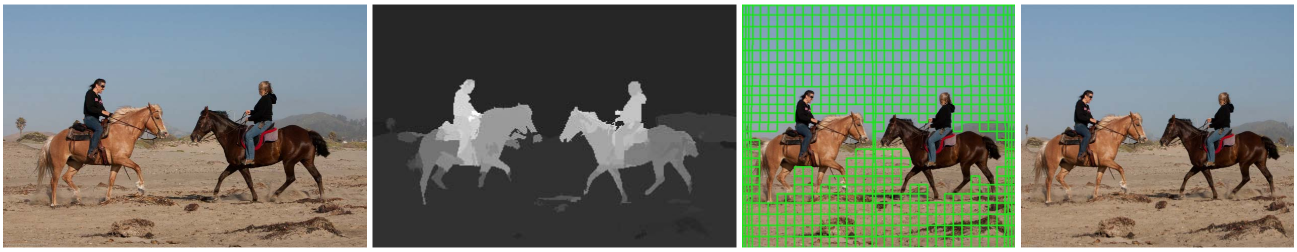
Our solution is to transform each salient image feature, like an object or a line, as a whole to avoid objectionable distortion. Our method first builds an adaptive mesh with non-rectangular mesh elements from an input image. Each salient image feature is enclosed by a single mesh element. Then, our method warps each mesh element with an affine transformation. In this way, we formulate image retargeting as a spatially varying affine warping problem, with the parameters of affine transformations as unknowns. To further minimize the distortion, we encourage each affine transformation to be close to more distortion-free ones, such as conformal transformation. Each energy term in our warping problem is at most quadratic, so our retargeting problem can be solved efficiently using a standard sparse linear solver. The main advantage of our method is its capability in preserving global image features.

In the remainder of this paper, we first give a brief overview on image retargeting in Section 2. We then describe how we build and warp an adaptive mesh in Section 3.

© The Eurographics Association 2011.

DOI: 10.2312/PE/PG/PG2011short/083-088

delivered by EUROGRAPHICS DIGITAL LIBRARY
www.eg.org diglib.eg.org



(a) Input (b) Region saliency (c) Our result with and without mesh

Figure 1: *Algorithm overview. Given an input image (a) and its saliency and region information (b), our method builds a non-uniform mesh where each salient object is included into a single mesh element ((c) left) and warps each mesh element using an affine transformation. In this way, each salient object is warped as a whole and is free from objectionable visual distortion. Our method creates the final result (c) by solving retargeting as a spatially-varying affine transformation problem, with the parameters of affine transformation for each mesh element as unknowns.*

We report our experiment results in Section 4. We conclude our paper in Section 5.

2. Related Work

Image and video retargeting has been a hot topic these years and a range of methods have been developed. In this section, we give a brief overview on image retargeting.

Early image retargeting methods focus on the problem of adapting an input image to small displays. These methods identify an important region in the image, crop, and scale the important region to fit the target displays [CXF*03, SLBJ03, SAD*06]. These cropping methods can best preserve the selected content; however, they lose information around image boundaries totally.

Re-composition-based methods rearrange image content to fit the target display size. Setlur *et al.* [STR*05] and Cheng *et al.* [CWW07] segment an image into regions, scale regions differently, and re-composite them into a final image. These methods can achieve good results; however, their performance depends on image segmentation. While our method also has a segmentation step, our requirement of segmentation results is much looser and our result is less affected by the segmentation results. Recent recomposition methods exploit repetitiveness or redundancy in an image to rearrange the image content [CBAF08, PKVP09, WWF*10]. These methods sometimes distort global image structures. Ding *et al.* [DXY11] address this problem by filtering the importance map and the shift map. Alternatively, seam carving methods iteratively carve or insert unnoticeable seams to meet the target size [AS07].

Warping-based methods formulate image retargeting as a mesh warping problem. These methods minimize objectionable visual distortion by enforcing conformal (similarity) transformation as well as other shape-preserving constraints (c.f. [GSCO06, WGCO07, WTSL08, ZCHM09]). Our method is particularly related to Guo *et al.* [GLS*09] and Jin *et al.* [JLW10]. These methods build a mesh that aligns well with salient image features. These methods distribute less distortion in visually salient regions than those less salient ones. These methods achieve good results; however, they sometimes cannot preserve salient global image structures since a global feature is divided into multiple mesh elements (grids or triangles). Our method is built upon these warping-based methods. Our method better preserves the global image structures by building an adaptive mesh and solving a spatially-varying affine warping problem.

Recently, multiple retargeting operators are combined together to achieve good retargeting results [RSA09, DZPZ09]. These image retargeting methods have also been extended to handle with videos [NLLG10, RSA08, GKHE10, KLHG09, WFS*09, LG06, ZHM08] and stereoscopic images and videos [LHW*10, CLC11].

3. Feature-Aware Mesh based Retargeting

Our method divides an input image into a non-uniform mesh. We denote a mesh element as $\mathbf{F}$ that consists of a number of vertices $\mathbf{v}^i$. (For a uniform mesh, $\mathbf{F}$ is a rectangular grid.) We build a mesh such that a salient feature, such as an object or a line, is enclosed by a single mesh element. We call this non-uniform mesh a *feature-aware mesh*. We allow each mesh element to undergo an affine transformation. In this way, our method formulates image retargeting as a mesh warping problem where the unknowns are the affine transformation parameters. We constrain the affine transformations such that the global image structures are preserved within this spatially-varying affine warping framework. In the following subsections, we first describe how we build a feature-

© The Eurographics Association 2011.

aware mesh (Section 3.1) and how we solve image retargeting as a spatially-varying warping problem (Section 3.2).

3.1. Feature-Aware Mesh

Instead of dividing an input image into a uniform grid mesh, we build a non-uniform mesh such that a salient feature, such as an object or a line, is inside a single mesh element. Our method first detects salient features and then builds a feature-aware mesh.

Our method combines techniques from saliency detection, image segmentation, and line detection to detect salient features. Specifically, our method employs a hough-transformation based method to detect lines [DH72]. We consider a line to be salient if it is at least half the image width long. Our method detects a salient object in two steps. We first use a graph-based method to detect image saliency [HKP06]. Then, we use the mean-shift image segmentation method to group salient pixels into salient regions [CM02]. In future, we can also use a salient region/object detection method, such as Cheng *et al*. [CZM*11], to segment salient objects directly.

Our method then divides an input image into a non-uniform mesh based on the saliency information. A straightforward method is to delineate each salient feature exactly around its boundary. However, this will make our method very sensitive to the salient feature detection result. Instead, our method uses a more conservative method. We first divide an input image into a uniform grid mesh. Then, we merge the grids that cover a salient feature into a single salient mesh element. We only merge κ percentage of the total image saliency into all the salient mesh elements. In this way, our method tends to include more content (possibly some less salient regions) instead of missing some part of salient content. We show an example of this feature-aware mesh in Figure 1, where our method divides the input image into a non-uniform mesh with two non-rectangular salient mesh element together with a large number of regular rectangular grids.

3.2. Spatially-varying Affine Warping

Our method aims to warp each salient feature consistently. Previous warping-based retargeting methods encourage each of the multiple grids that cover the same feature to undergo a conformal transformation. However, each of these grids can undergo a different conformal transformation, which often damages the global image structure. To address this problem, our method tries to include a salient image feature in one mesh element and accordingly warps each (possibly non-rectangular) mesh element with a single affine transformation. In this way, a salient feature can be warped consistently. We then formulate image retargeting as a spatially-varying affine warping problem, where the variables are the affine transformation parameters. We choose an affine transformation because it provides a reasonable amount of degrees of freedom for 2D transformation and makes defining quadratic energy terms easy, compared to a homography.

We denote the affine transformation matrix of a mesh element $\mathbf{F}^i$ as $\mathbf{A}^i$. We denote row k of $\mathbf{A}^i$ as $\mathbf{A}^i_k$ and a matrix element as $\mathbf{A}^i_{k,j}$. We denote the width and height of the input image as w, h and those of the output as $\hat{w}, \hat{h}$. We now describe how we define energy terms in the spatially-varying affine warping problem to produce a retargeting result.

Image boundary constraint. Our method enforces the target boundary constraint on vertices that are on the original image boundaries as follows.

$$\begin{cases} \mathbf{A}_1\mathbf{v} = 0, & \forall \mathbf{v} \in \text{the left boundary} \\ \mathbf{A}_1\mathbf{v} = \hat{w}, & \forall \mathbf{v} \in \text{the right boundary} \\ \mathbf{A}_2\mathbf{v} = 0, & \forall \mathbf{v} \in \text{the top boundary} \\ \mathbf{A}_2\mathbf{v} = \hat{h} & \forall \mathbf{v} \in \text{the bottom boundary} \end{cases} \tag{1}$$

where $\mathbf{A}$ is the affine transformation matrix that applies to the mesh element that contains vertex $\mathbf{v}$. If $\mathbf{v}$ is shared by multiple mesh elements, the above constraint is repeated using the corresponding affine transformation matrices.

Shared vertex constraint. A vertex that is shared by multiple mesh elements is transformed by multiple affine transformations. Our method encourages that these affine transformations move the common vertex to the same position.

$$E^s_v = \sum_{\mathbf{v}\in\mathbf{F}^i \& \mathbf{v}\in\mathbf{F}^j} \|(\mathbf{A}^i - \mathbf{A}^j)\mathbf{v}\|^2 \tag{2}$$

where $\mathbf{F}^i$ and $\mathbf{F}^j$ are any two mesh elements that share vertex $\mathbf{v}$. $\mathbf{A}^i$ and $\mathbf{A}^j$ are the corresponding affine transformations. E^s_v is computed over all the shared vertices. We use this as a soft constraint to increase the flexibility of our warping method. This flexibility together with the following shape-preserving constraints allows to concentrate more distortion to visually non-salient regions. Mathematically, this constraint alone will often make the system nearly overdetermined, which prohibits it from being a hard constraint. On the other hand, we find that the actual violation against this constraint is usually small, so we can post-process the warping result to meet this constraint exactly, as described later.

Shape-preserving constraint. An affine transformation can still introduce objectionable distortion. To further minimize the visual distortion, our method encourages salient mesh elements to undergo a conformal transformation. Since our method works on affine transformations directly, this conformal transformation constraint can be easily implemented by constraining the transformation matrix elements directly.

$$E^p_F = w_F(\mathbf{A}_{11} - \mathbf{A}_{22})^2 + w_F(\mathbf{A}_{12} + \mathbf{A}_{21})^2 \tag{3}$$

where $\mathbf{A}_{i,j}$ is element i, j of the affine transformation $\mathbf{A}$ for the mesh element $\mathbf{F}$. w_F is the salient value of $\mathbf{F}$. Our method

© The Eurographics Association 2011.

(a) Input (b) Our result with uniform mesh (c) Our result with feature-aware mesh

Figure 2: *For the input image (a), our method creates a result with a uniform mesh (b) and a result with a feature-aware mesh (c). This example shows that the feature-aware mesh enables our method to warp an important object more consistently than a uniform mesh.*

encourages visually salient regions to undergo more restrictive conformal transformation than those less salient ones. E_F^p is computed over all the mesh elements.

When an image is resized to a small display, it is often desirable that salient regions are given more space than the others. Our method emphasizes salient regions by encouraging them to undergo a more restrictive translation transformation than those less salient ones.

$$E_F^t = w_F((\mathbf{A}_{11}-1)^2 + (\mathbf{A}_{22}-1)^2 + \mathbf{A}_{12}^2 + \mathbf{A}_{21}^2) \quad (4)$$

We combine all the energy terms described above and obtain the following quadratic programming problem.

$$E = \lambda^s \sum_{\mathbf{v}\in\mathcal{S}} E_v^s + \lambda^p \sum_{\mathbf{F}\in\mathcal{M}} E_F^p + \lambda^t \sum_{\mathbf{F}\in\mathcal{M}} E_F^t \quad (5)$$

s.t. the boundary constraint in Equation 1

where $\mathcal{S}$ is the set of all the shared vertices among mesh elements and $\mathcal{M}$ is the set of all the mesh elements. λ^s, λ^p, and λ^t are weighting parameters. Since the boundary constraint in Equation 1 is an equality hard constraint, the above quadratic programming problem is a quadratic minimization problem, and we solve this quadratic minimization problem using a standard sparse linear solver provided by Intel MKL. The speed of our method mainly depends on the number of mesh elements (cells) in our system. For an image with 1000 mesh elements, this optimization problem can be solved in realtime on a desktop machine with Intel Core2 3GHz CPU and 4G memory.

After we solve Equation 5 for the affine transformation for each mesh element, we compute the output vertex positions. Since a vertex is often shared by multiple mesh elements, we compute the final vertex position $\hat{\mathbf{v}}$ as a weighted average of positions computed using each relevant affine transformation.

$$\hat{\mathbf{v}} = \frac{\sum_{\mathbf{v}\in\mathbf{F}^i} w_F^i \mathbf{A}^i \mathbf{v}}{\sum_{\mathbf{v}\in\mathbf{F}^i} w_F^i} \quad (6)$$

where w_F^i and $\mathbf{A}^i$ are the saliency value and the affine transformation matrix of the mesh element $\mathbf{F}^i$. We find from our experiment that the difference between the positions of a vertex warped by different affine transformations is typically very small due to the shared vertex constraint. We find that this weighted average method works well. Once we obtain the final vertex positions, we render the retargeting result using texture mapping.

4. Experiments

We tested our method on a collection of images from *Flickr* as well as the *RetargetMe* benchmark provided by Rubinstein *et al.* [RGSS10]. We show some representative results in Figure 2 and Figure 3. We submitted all our results for the *RetargetMe* benchmark as well as comparisons to the state of the art methods as supplemental materials.

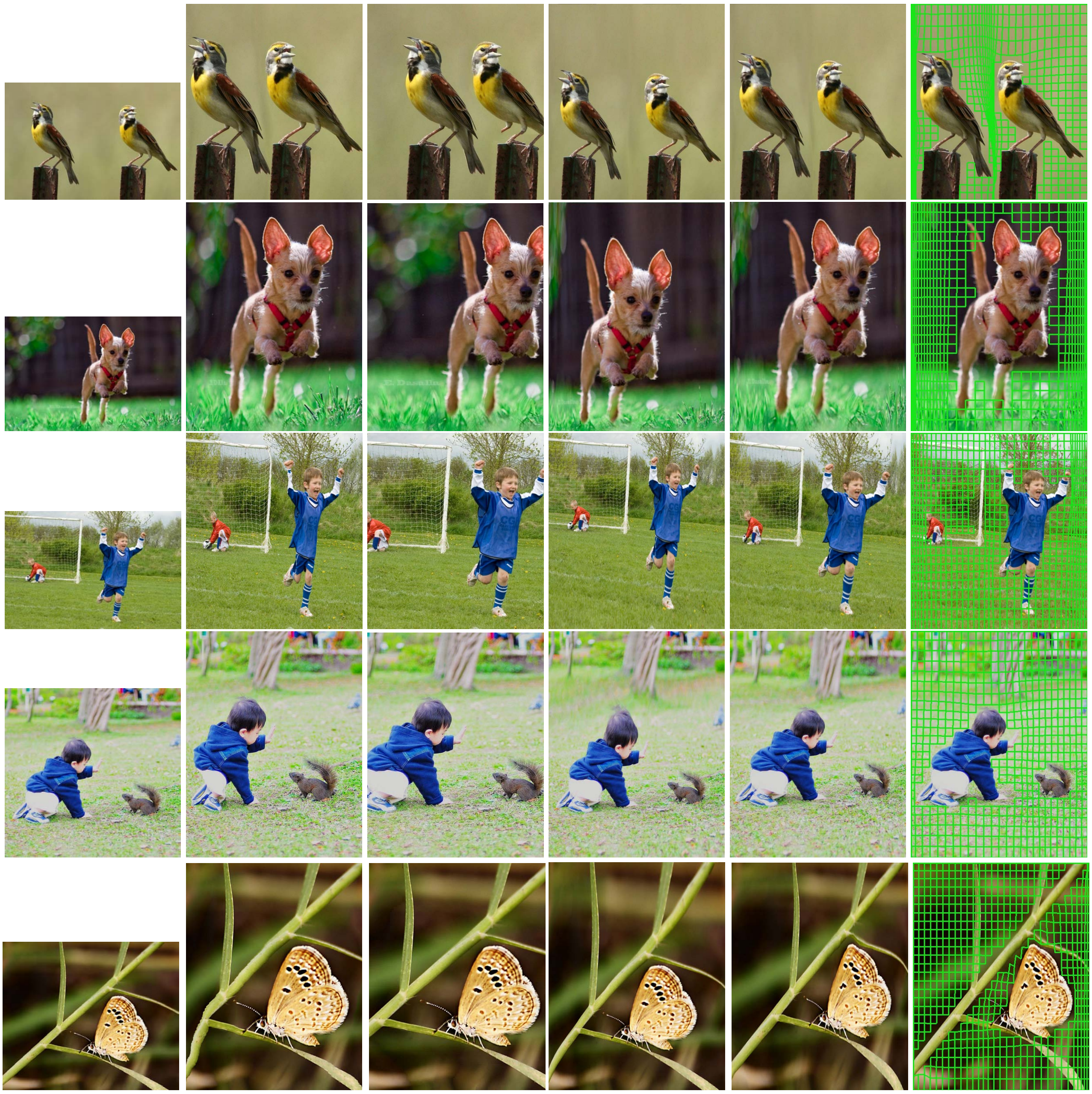
To first verify that the feature-aware mesh can help our method better preserve global image structures, we show an example in Figure 2. For this example, our method creates retargeting results with a uniform mesh and a feature-aware mesh respectively. In the result with the uniform mesh (b), the beaks of the Curlew birds are apparently stretched while their bodies almost remain the same aspect ratio as the input. In contrast, the feature-aware mesh warps the birds consistently, as shown in (c).

Overall, our experiments show that our method, as well as many of the state of the art methods, can all produce reasonable results. The performance of our method is at least comparable to the existing methods. Sometimes our method is better in preserving global image structures. For example, the third row in Figure 3 shows that our method simultaneously maintains the straightness of the goal and prevents the boys from being distorted. The fourth row in this figure shows that our method protects the boy from being distorted. The last row in the same figure shows that our method can keep the grass stem straight better than some other methods.

5. Conclusion

In this paper, we presented a feature-aware mesh based image retargeting method. Our method first divides an input

© The Eurographics Association 2011.



(a) Input (b) [RSA08] (c) [PKVP09] (d) [WTSL08] (e) Our result

Figure 3: *Comparisons to the state of the art methods.*

image into a non-uniform mesh so that each salient object or feature, like a line, is included in a single mesh element. By allowing each mesh element to undergo only an affine or more distortion-free 2D transformation, our method can warp each salient object consistently. Our method eventually formulates image retargeting as a spatially-varying affine warping problem, which can be efficiently solved as a quadratic energy minimization problem. Our experiment shows that our method can produce at least comparable results to the state of the art methods. Our method sometimes can better preserve global image structures. In the future, we would like to extend our method to videos.

© The Eurographics Association 2011.

Acknowledgements

We would like to thank reviewers for their constructive suggestions. The Flickr images are used under a Creative Commons license from Flickr users: mikebaird, cyclewidowpatti, pure_awesome, 42737468@N02, matthewfch, and yogendra174.

References

[AS07] AVIDAN S., SHAMIR A.: Seam carving for content-aware image resizing. *ACM Trans. Graph. 26*, 3 (2007), 267–276. 2

[CBAF08] CHO T. S., BUTMAN M., AVIDAN S., FREEMAN W. T.: The patch transform and its applications to image editing. In *CVPR '08: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2008), pp. 1–8. 2

[CLC11] CHANG C.-H., LIANG C.-K., CHUANG Y.-Y.: Content-aware display adaptation and interactive editing for stereoscopic images. *IEEE Transactions on Multimedia* (2011). 2

[CM02] COMANICU D., MEER P.: Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Machine Intell. 24*, 5 (2002), 603–619. 3

[CWW07] CHENG W.-H., WANG C.-W., WU J.-L.: Video adaptation for small display based on content recomposition. *IEEE Transactions on Circuits and Systems for Video Technology 17*, 1 (2007). 2

[CXF*03] CHEN L.-Q., XIE X., FAN X., MA W.-Y., ZHANG H.-J., ZHOU H.-Q.: A visual attention model for adapting images on small displays. *Multimedia Systems 9*, 4 (2003), 353–364. 2

[CZM*11] CHENG M.-M., ZHANG G.-X., MITRA N. J., HUANG X., HU S.-M.: Global contrast based salient region detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2011), pp. 409–416. 3

[DH72] DUDA R. O., HART P. E.: Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM 15* (January 1972), 11–15. 3

[DXY11] DING Y., XIAO J., YU J.: Importance filtering for image retargeting. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2011). 2

[DZPZ09] DONG W., ZHOU N., PAUL J.-C., ZHANG X.: Optimized image resizing using seam carving and scaling. *ACM Trans. Graph. 28*, 5 (2009), 1–10. 2

[GKHE10] GRUNDMANN M., KWATRA V., HAN M., ESSA I.: Discontinuous seam-carving for video retargeting. In *CVPR '10: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2010), pp. 569 – 576. 2

[GLS*09] GUO Y., LIU F., SHI J., ZHOU Z., GLEICHER M.: Image retargeting using mesh parametrization. *IEEE Trans. on Multimedia 2009 11*, 4 (2009), 856–867. 2

[GSCO06] GAL R., SORKINE O., COHEN-OR D.: Feature-aware texturing. In *EGSR '06: Proceedings of the 17th Eurographics Symposium on Rendering* (2006), pp. 297–303. 1, 2

[HKP06] HAREL J., KOCH C., PERONA P.: Graph-based visual saliency. In *Proceedings of Neural Information Processing Systems* (2006). 3

[JLW10] JIN Y., LIU L., WU Q.: Nonhomogeneous scaling optimization for realtime image resizing. *The Visual Computer 26*, 6-8 (2010), 769–778. 2

[KLHG09] KRÄHENBÜHL P., LANG M., HORNUNG A., GROSS M.: A system for retargeting of streaming video. *ACM Trans. Graph. 28*, 5 (2009), 1–10. 2

[LG06] LIU F., GLEICHER M.: Video retargeting: automating pan and scan. In *Multimedia '06: Proceedings of the 14th international conference on Multimedia* (2006), pp. 241–250. 2

[LHW*10] LANG M., HORNUNG A., WANG O., POULAKOS S., SMOLIC A., GROSS M.: Nonlinear disparity mapping for stereoscopic 3d. *ACM Trans. Graph. 29*, 4 (2010). 2

[NLLG10] NIU Y., LIU F., LI X., GLEICHER M.: Warp propagation for video resizing. In *CVPR '10: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2010), pp. 537–544. 2

[PKVP09] PRITCH Y., KAV-VENAKI E., PELEG S.: Shift-map image editing. In *ICCV '09: Proceedings of the Twelfth IEEE International Conference on Computer Vision* (2009), pp. 151–158. 2, 5

[RGSS10] RUBINSTEIN M., GUTIERREZ D., SORKINE O., SHAMIR A.: A comparative study of image retargeting. *ACM Trans. Graph. 29* (December 2010), 160:1–160:10. 1, 4

[RSA08] RUBINSTEIN M., SHAMIR A., AVIDAN S.: Improved seam carving for video retargeting. *ACM Trans. Graph. 27*, 3 (2008), 1–9. 2, 5

[RSA09] RUBINSTEIN M., SHAMIR A., AVIDAN S.: Multi-operator media retargeting. *ACM Trans. Graph. 28*, 3 (2009), 1–11. 2

[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* (2006), CHI '06, pp. 771–780. 2

[SLBJ03] SUH B., LING H., BEDERSON B. B., JACOBS D. W.: Automatic thumbnail cropping and its effectiveness. In *UIST '03: Proceedings of the 16th annual ACM symposium on User Interface Software and Technology* (2003), pp. 95–104. 2

[STR*05] SETLUR V., TAKAGI S., RASKAR R., GLEICHER M., GOOCH B.: Automatic image retargeting. In *MUM' 05: the 4th International Conference on Mobile and Ubiquitous Multimedia* (Dec. 2005), pp. 59–68. 2

[WFS*09] WANG Y.-S., FU H., SORKINE O., LEE T.-Y., SEIDEL H.-P.: Motion-aware temporal coherence for video resizing. *ACM Trans. Graph. 28*, 5 (2009). 2

[WGCO07] WOLF L., GUTTMANN M., COHEN-OR D.: Non-homogeneous content-driven video-retargeting. In *ICCV '07: Proceedings of the Eleventh IEEE International Conference on Computer Vision* (2007), pp. 1–6. 1, 2

[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, 2, 5

[WWF*10] WU H., WANG Y.-S., FENG K.-C., WONG T.-T., LEE T.-Y., HENG P.-A.: Resizing by symmetry-summarization. *ACM Trans. Graph. 29* (December 2010), 159:1–159:10. 2

[ZCHM09] ZHANG G.-X., CHENG M.-M., HU S.-M., MARTIN R. R.: A shape-preserving approach to image resizing. *Computer Graphics Forum 28*, 7 (2009), 1897–1906. 1, 2

[ZHM08] ZHANG Y.-F., HU S.-M., MARTIN R. R.: Shrinkability maps for content-aware video resizing. *Computer Graphics Forum 27*, 7 (2008), 1797–1804. 2

© The Eurographics Association 2011.