RAPID IMAGE RETARGETING BASED ON CURVE-EDGE GRID REPRESENTATION

Tongwei Ren[1], *Yan Liu*[2], *and Gangshan Wu*[1]

[1]State Key Laboratory for Novel Software Technology, Nanjing University, China
[2]Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China

ABSTRACT

Image retargeting technique attracts more and more attention for convenient image display on mobile devices. However, current methods can't well balance the retargeting efficiency and effectiveness, which limits their applications on the mobile devices with low computing ability. In this paper, we propose a novel image retargeting approach by combining uniform sampling and structure-aware curve-edge grid representation. We first decompose the original image into curve-edge grids by dynamic programming, and then generate the target image by uniformly sampling the pixels within the grids. The simplicity of retargeting procedure and sampling strategy enables our approach to easily achieve good computational efficiency. Furthermore, the constraint of curve-edge grid representation ensures important content emphasis and image structure preservation in the target image. Experiments on different images demonstrate the effectiveness and efficiency of our approach.

***Index Terms*—** Image retargeting, curve-edge grid, structure-aware representation

1. INTRODUCTION

The wide applications of multimedia techniques lead to the significant requirements of image display on various mobile devices, such as cellular phones and PDAs [1]. However, the existing images usually have much higher resolution and fixed aspect ratios. They should be adapted to match the target screens on these mobile devices. This problem is commonly referred to *image retargeting*.

Effective image retargeting technique should retain the emphasized important content and sufficient context information in the target image while reducing the visual distortion for good viewing experience [2]. To achieve this objective, current image retargeting methods mainly resize the image by using constrained optimization. Seam carving [1] calculates the energy of each pixel and iteratively removes the least energy seams by dynamic programming. Non-homogenous warping [3] relocates the pixels by solving a large sparse linear system. Scale-and-stretch method [4] and constrained region warping [2] decompose the image into square meshes or trapezoid meshes and relocate the mesh vertices by quadratic programming. Multi-operator method [5] utilizes multiple operators in retargeting and finds the best operator sequence by maximizing the similarity between the original image and the target image. These methods present prominent effectiveness in content-aware image resizing. But they are time consuming for solving the optimization problems. It seriously limits their applications on the mobile devices with low computing ability.

To address the problem of high computational cost, a simple solution is to uniformly scale the image to the target size, but it may cause important content distortion and details loss. Cropping [6] and fisheye-view warping [7] retain the region covering as many important objects as possible, and discard or non-linearly warp the rest part. They may cause much background information loss, and they are invalid when the important objects are sparsely distributed. Row/column removal method [8] removes the least important row (or column) and diffuses its importance to preserve image structure. But it uses a simple column (or row) as the processing unit, so it obtains poor performance when the image has complex structure.

In this paper, we propose a novel image retargeting approach by combining uniform sampling and structure-aware image representation. Discretely uniform scaling, the fastest method for image resizing, is utilized to achieve good computational efficiency. However, unconstrained uniform sampling will distort the image structure and cause important details loss. So we propose a curve-edge grid representation of the original image (Fig. 1(d)) that is consistent to the underlying structure of image content and importance energy map (Fig. 1(b)), and constrain pixel sampling within the grids. Specifically, we propose a new image decomposition method in curve-edge grids generation by using dynamic programming on a connected graph (Fig. 1(c)), whose result can avoid the artifacts in the following sampling. By assigning different importance to the grids and retaining the grid boundaries in retargeting, the target image (Fig. 1(e)) can well emphasize the important content and preserve image structure.

2. CUREVE-EDGE GRID REPRESENTATION

Current rapid image retargeting methods adopt the rigid pixel sampling strategies, such as warping the image based on the pre-defined templates [7] or removing a column/row of pixels [8]. These strategies can't well fit image structure, and prevent them to achieve effective performance in retargeting.

To address this problem, we decompose the image and carry out uniform sampling within the generated regions. Compared with the rectangle region [6] or image column/row [8], these regions can represent image structure better. There exist many image segmentation methods providing effective solutions of decomposing an image into homogenous regions, such as mean-shift [9]. But uniformly sampling the pixels on

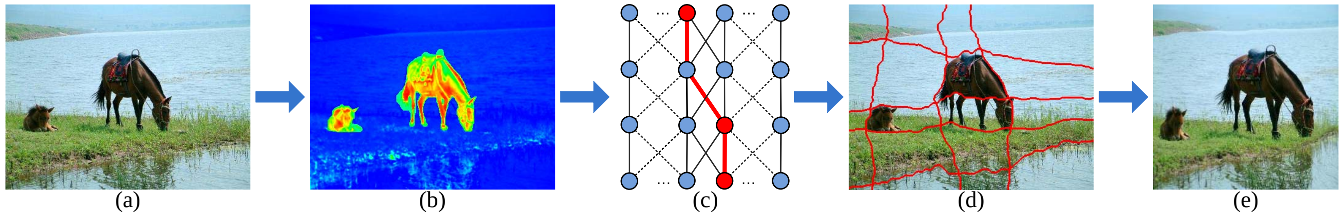


Fig. 1. An overview of the proposed approach. (a) Original image. (b) Importance energy map. (c) Carving graph. (d) Curve-edge grid representation. (e) Target image.

these regions may destroy the region boundaries (Fig. 4(d)) or cause "cut" effects between the adjacent regions (Fig. 4(e)). To avoid the artifacts, the generated region boundaries in our approach should cross through the image vertically (or horizontally). Hence, we propose a new image decomposition method to satisfy the requirement. In our method, the vertical decomposition and horizontal decomposition are processed separately. Here, we only discuss the vertical decomposition, and the horizontal decomposition can be dealt with similarly.

2.1. Carving graph construction

As shown in Fig. 2, we treat each pixel joint point (red points in Fig 2(a)) as a vertex in graph and allow 8-connected edges in vertical direction. In this way, the image in size of $W \times H$ can be represented as a connected graph with $(W-1)\times(H+1)$ vertices, which is named "carving graph" (Fig 2(b)).

To each edge between vertices $v_{x,y}$ and $v_{x',y+1}$ in carving graph (x' is x or $x \pm 1$), we assign a weight to it to denote the difference between pixels on its two sides:

$$w\left(v_{x,y}, v_{x',y+1}\right) = \frac{\left\|\frac{\partial}{\partial x} I[x,y]\right\| \cdot \left|\frac{\partial}{\partial x} E[x,y]\right|}{l\left(v_{x,y}, v_{x',y+1}\right)}, \quad (1)$$

where $w(,)$ and $l(,)$ are the weight and the length of the edge; $\left\|\frac{\partial}{\partial x} I[x,y]\right\|$ and $\left|\frac{\partial}{\partial x} E[x,y]\right|$ are the distances between pixel on the left or crossed by the edge (marked with blue in Fig. 2(a)) and its right pixel in Lab color space and importance energy map respectively. In importance energy map calculation, we utilize saliency region detection [10] and face detection [11]. Some manual interactions may be required when the image is too complex to automatically calculate the energy maps.

2.2. Curve-edge grid generation

In carving graph, a connected path with high weight sum from top to bottom is in corresponding to a satisfied region boundary with the obvious difference between the regions on its two sides. So we formulate image decomposition to the problem of selecting the paths with high edge weight sums in carving graph.

Fig. 3 shows an example of path selection. To each pixel in the bottom row, we find the paths with the highest edge weight sum by dynamic programming (Fig 3(b)):

$$\begin{aligned} M[x,y] = \max\{ & M[x-1,y-1] + w(v_{x-1,y-1}, v_{x,y}), \\ & M[x,y-1] + w(v_{x,y-1}, v_{x,y}), \\ & M[x+1,y-1] + w(v_{x+1,y-1}, v_{x,y})\} \end{aligned} \quad (2)$$

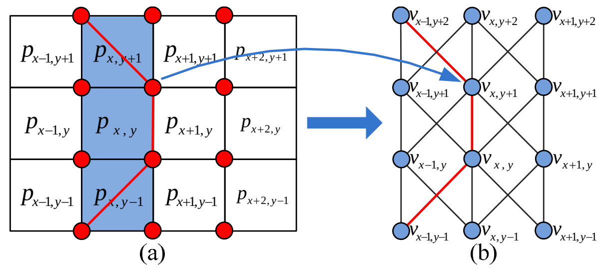


Fig. 2. Carving graph construction. (a) Image pixels. (b) Carving graph.

It is found that the connected paths with high edge weight sums are derived from several paths in the top of the image, which represent the local dominant content distinctions in the top half of image. So we select these clustered paths to represent image structure in decomposition. To each clustered path, we retain the path with the largest edge weight sum from all the paths started from it (red paths in Fig. 3(b)). Similarly, we carry out the procedure inversely from bottom to the top, and use the selected paths to represent the image structure of the bottom half (Fig. 3(c)). For the paths may be over-clustered when the image height is large, we penalize the diagonal path by setting the lengths of diagonal edges as twice of the vertical ones in Equation (1). For the path generations from different directions are not symmetrical, there may generate partly overlapped paths (Fig. 3(d)), which may cause the over-decomposition of image. So we combine these paths when the number of their overlap pixels is larger than the pre-defined threshold (using 0.5H in our experiments). In path combination, to each distinct part of the two paths, we make the selection according to its position: we select the part from the top-bottom path when it lies in the top half and from the bottom-top path when it lies in the bottom half. In this way, we obtain the vertical decomposition of image (Fig. 3(e)). Similarly, we can obtain the horizontal decomposition of image. And then, the image can be represented as the curve-edge grids according to the decomposition result (Fig. 3(f)).

3. CONSTRAINED UNIFORM SAMPLING

By the constraint of curve-edge representation, we resize the image by uniform sampling. In our approach, the horizontal resizing and the vertical resizing are independently constrained by the corresponding grid boundaries, so they can be carried out separately. Here, we only discuss the horizontal resizing.

According to the image decomposition procedure, only the grid boundaries generated in vertical image decomposition is used to constrain the horizontal resizing. Assume the image is

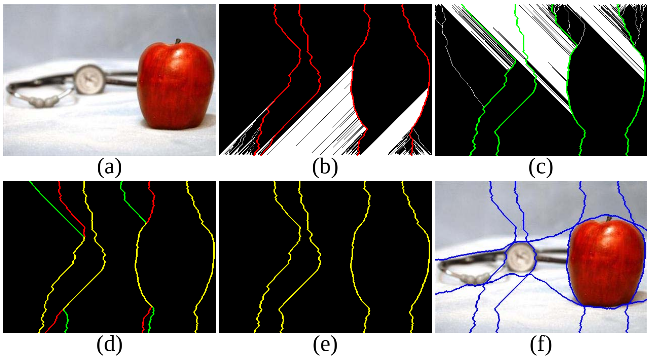


Fig. 3. Example of curve-edge grid representation. (a) Original image. (b) Result of top-bottom path selection. (c) Result of bottom-top path selection. (d)-(e) Result of path combination. (f) Curve-edge grid representation.

decomposed into $M \times N$ curve-edge grids, the grids in the same column $\{G_{i,\#}\}$ are treated as a processing unit, and the pixels within it are uniformly sampled in each row. Obviously, if we want to retain a unit boundary, the same number of pixels should be removed in each row on its left/right. It means the removed pixel number of each row should be equal within a unit when all the grid boundaries are retained.

To each processing unit u_i, the valid number of removed pixels in each row is limited by the minimal width of all the rows within it. However, only using the minimal width may lead to the excessive sampling in the less important processing units, which may cause serious context information loss and destroy image layout (Fig. 4(f)). So we calculate the valid number of removed pixels in each row of unit u_i as follows:

$$c(i) = \min\left\{ wid_{\min}(u_i), \frac{N(i) - N(i)\cdot \bar{E}(i)^{\alpha}}{H} \right\}, \tag{3}$$

where $c(i)$ is the valid number of removed pixels in each row of u_i; $wid_{\min}(u_i)$ is the minimal width of all rows in u_i; $N(i)$ is the total pixel number in u_i; $\bar{E}(i)$ is the average pixel importance energy in u_i, in the value range of [0,1]; α is a parameter to control the emphasis of important content, and larger α leads to more emphasis (using $\alpha = 1$ in our experiment).

We orderly distribute the number of required removed pixels to the units according to their average pixel importance energies, from low to high. When the number exceeds the sum of valid numbers of all the units, we iteratively remove the unit boundaries with the least importance and sample the pixels in the combined units. Here, the importance of a unit boundary is defined as the maximum value of the average pixel importance energy of the grids adjacent to it. When all the unit boundaries are removed, the image is scaled uniformly. Fig. 4(g) shows an example of our retargeting results. For the uniform sampling is constrained by curve-edge grid representation, the image structure is well preserved in target image.

4. EXPERIMENTS

To illuminate our performance, we implement our approach with C++ and test it on a PC with Duo CPU 2.80GHz and 2GB RAM in simulation.

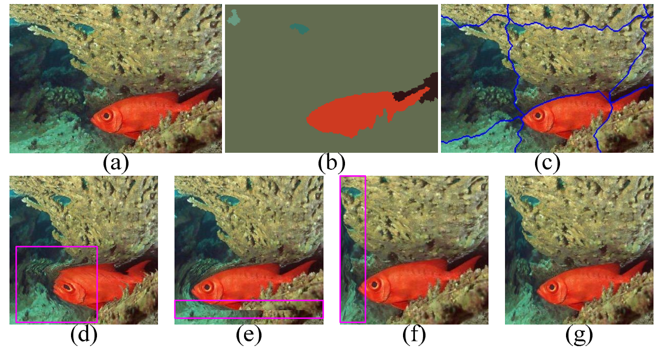


Fig. 4. Example of constrained uniform sampling. (a) Original image. (b) Segmentation result by mean-shift. (c) Curve-edge grid representation. (d)-(e) Results based on mean-shift segmentation. (f)-(g) Our results with different valid numbers.

Fig. 5 shows the target images generated from the same original image but with different sizes and aspect ratios. It shows our approach is robust to different target screens.

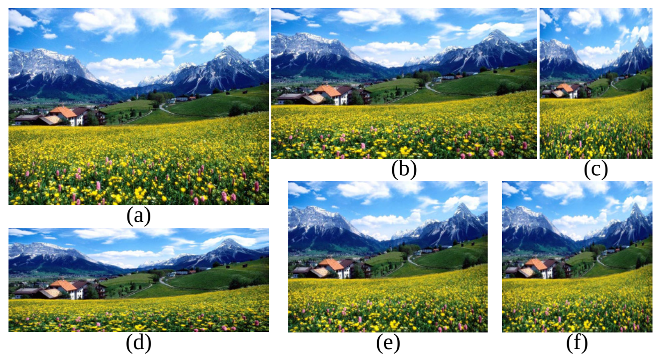


Fig. 5. Target images with different sizes and aspect ratios. (a) Original image. (b)-(f) Target images with aspect ratios in 16:9, 3:4, 2.5:1, 4:3, and 1:1, respectively.

Fig. 6 shows an example of comparison with the existing typical image retargeting methods, including uniform scaling (SL), cropping (CP) [6], fisheye-view warping (FEW) [7], row/column removal (R/C) [8], seam carving (SC) [1], non-homogenous warping (NHW) [3], scale-and-stretch (S&S) [4], constrained region warping (CRW) [2] and multi-operator (MO) [5]. The target image is required to be half in width of the original image. We can see that all the existing rapid image retargeting methods can't deal with the image well. Compared with them, our algorithm obtains a comparable result to the optimization based methods.

We further evaluate the effectiveness of our method by carrying out a user study in comparison with the above methods. Fifty images with various styles are used as the original images. And twelve users with different occupations and in age of 22 to 46 are invited to do pair-wise comparisons of the target images generated by our approach and other methods. In evaluation, each our result is required to compare with any other corres-ponding target image by three users, and the dominant judg-ment is treated as the final evaluation. Tab. 1 shows the result of user study. We can find that our approach is prominently better than other rapid methods in performance, and our effectiveness is comparable to the optimization based methods.

(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Fig. 6. Example of comparison with other methods. (a) Original image. (b) Results of SL. (c) Result of CP (d) Result of FEW. (e) Result of R/C. (f) Result of SC. (g) Result of NHW. (h) Result of S&S. (i) Result of CRW. (j) Result of MO. (k) Our result.

Tab. 1. Result of user study.

	Better	Similar	Worse
SL	39	10	1
CP	26	19	5
FEW	23	24	3
R/C	18	28	4
SC	5	39	6
NHW	9	38	3
S&S	9	34	7
CRW	4	37	9
MO	2	42	6

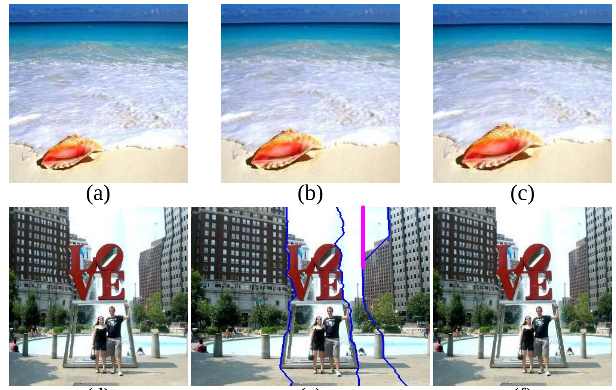
We also evaluate the efficiency of our approach. For the energy maps calculation has been evaluated to be efficient [10][11], we only discuss the time cost in image resizing. The time complexity of image resizing is $O(W_S H_S)$. When retarget the above fifty images in sizes from 800×600 to 900×600 to the target size 300×300, the average time cost is 0.023s in curve-edge grids generation and 0.016s in sampling. Moreover, the image requires just once decomposition in retargeting, and then the user can adjust the target size by only carrying out sampling. The performance of the simulation test shows our method is suitable to the devices with low computing ability.

We compare the efficiency of our approach with the existing methods. For the time cost is seriously influenced by implementation, we just present the time costs of some methods with exact implementations for comparison. In our experiment, the average time costs of uniform sampling, cropping and seam carving are 0.004s, 0.017s, and 2.981s, respectively.

Considering the users may have different preferences in retargeting, we provide user interaction approaches in implementation. Fig. 7(a)-(c) shows the target images with different emphasis degrees of important content by adjusting parameter α in Equation (3). User interaction is also allowed to improve the target image quality in our implementation (Fig. 7(d)-(f)).

5. CONCLUSION

This paper presents a novel image retargeting approach by combining uniform sampling and curve-edge grid representation. A dynamic programming based image decomposition method is proposed for curve-edge grid generation, and some constraints are used to improve the effect in sampling. With the low computational cost in uniform sampling and the powerful image structure representation of curve-edge grids, the proposed approach achieves good performance in both retargeting efficiency and effectiveness.



(a) (b) (c) (d) (e) (f)

Fig. 7. Examples of user interactions. (a)-(c) Target images with $\alpha = 1$, 1.5, 2 , respectively. (d) Automatic target image. (e) User interaction by adding edges (magenta is the modified edge). (f) Improved target image.

6. ACKNOWLEDGEMENTS

This work is supported by Grant 60721002, 60975043 of the National Natural Science Foundation of China and the Hong Kong General Research Fund PolyU 5204/09E.

7. REFERENCES

[1] M. Rubinstein, A. Shamir, and S. Avidan, “Improved Seam Carving for Video Retargeting,” *TOG*, 27(3):1-9, 2008.

[2] T. Ren, Y. Liu, and G. Wu, “Image Retargeting Using Multi-map Constrained Region Warping,” *MM*, China, 2009.

[3] L. Wolf, M. Guttmann, and D. Cohen-Or, “Non-homogeneous Content-driven Video-retargeting,” *ICCV*, Brazil, 2007.

[4] Y.S. Wang, C.L. Tai, O. Sorkine, and T.Y. Lee, “Optimized Scale-and-Stretch for Image Resizing,” *TOG*, 27(5), 1-8, 2008.

[5] M. Rubinstein, A. Shamir, and S. Avidan, “Multi-operator Media Retargeting,” *TOG*, 28(3):1-11, 2009.

[6] 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,” *MM Sys.*, 9(4):353-364, 2003.

[7] F. Liu and M. Gleicher, “Automatic Image Retargeting with Fisheye-View Warping,” *UIST*, USA, 2005.

[8] S. Cho, H. Choi, Y. Matsushita, and S. Lee, “Image Retargeting Using Importance Diffusion,” *ICIP*, Egypt, 2009.

[9] D. Comaniciu and P. Meer, “Mean Shift: A Robust Approach toward Feature Space Analysis,” *TPAMI*, 24(5):603–619, 2002.

[10] R. Achanta, S. Hemami, F. Estrada, and S. Süsstrunk, “Frequency-tuned Salient Region Detection,” *CVPR*, USA, 2009.

[11] P. Viola and M.J. Jones, “Robust Real-Time Face Detection,” *IJCV*, 57(2):137-154, 2004.