FAST IMAGE RETARGETING VIA AXIS-ALIGNED IMPORTANCE SCALING

Sunghwan Choi, Bumsub Ham, and Kwanghoon Sohn

School of Electrical and Electronic Engineering, Yonsei University, Seoul, Korea

ABSTRACT

In this paper, we propose an image retargeting method that resizes an image by the axis-aligned importance scaling. The proposed method operates on the axis-aligned deformation space, where the mesh structure is parameterized in the 1D vector, i.e., quads on the same column (or row) share the single parameter. The unknown variables are thus dependent on the grid resolution only, allowing a fast and simple implementation on a moderate CPU. The optimal parameters are inferred in an iterative manner: these parameters are updated by scaling initial ones according to the deformation error. It is measured at each iteration by aggregating the transition cost for the deformed quad. Experimental results show that the proposed method preserves visually salient features without foldover artifacts better than competing methods. In addition, the optimal parameters can be calculated within 0.02 ms on a single-core CPU.

Index Terms— Image retargeting, importance scaling, axis-aligned deformation.

1. INTRODUCTION

Recently, image retargeting has been actively developed in order to meet the varying format of diverse display devices. It allows an image to be resized into arbitrary aspect ratio, while preserving visually prominent features. The underlying basic idea is that the image content is managed nonlinearly based on the visual importance. For example, a region with visually prominent features is retained, while a less-important region is uniformly modified for satisfying the target resolution. To capture visual importance of the image, automatic saliency detection algorithms such as [1] have been exploited as external information.

From the perspective of a deformation unit, image retargeting methods can be grouped into either pixel-wise or the mesh-wise methods. Pixel-wise methods resize the image by removing less-important pixels. The seam carving method [2] greedily removes optimal seams, computed by the dynamic programming, that pass through less important regions. The shift-map image editing method [3] infers optimal graph labels for the target resolution using the graph-cuts. These labels can also be computed by the importance filtering [4], where the shift gradient is integrated using a weighted filter. Unfortunately, the results of pixel-wise methods usually show spatial discontinuity after deformation due to their discrete nature, causing noticeable jags in structural objects. Mesh-wise methods [5], [6] offer a continuous solution to image resizing. A grid mesh is placed onto the image and then a new geometry for the target resolution is estimated. The quad faces covering important image regions remain intact at the expense of larger distortion to the other quads. These methods, however, suffer from foldover artifacts, i.e., self-intersections. Although this problem can be resolved by enforcing the nonlinear constraint in the energy functional, it requires complex and time consuming optimization.

We present a fast and foldover-free image retargeting method that iteratively estimates the new geometry for the target resolution by an importance-based local scaling. The retargeting problem is posed as finding a local scaling factor for every quad on the same column (or row) such that these axis-aligned quads, i.e., the 1D parameterized mesh structure, share the same scaling parameter. The amount of scaling for each local region is determined based on visual saliency information. That is, the quad with a high saliency remains the same, and *vice versa*. The optimal parameters are inferred in an iterative manner by aggregating the transition cost for deformed quad, followed by scaling each axis-aligned quad in proportional to the importance-weighted aggregated cost. Since the 1D parameterized mesh structure is leveraged, the unknown variables are dependent on the grid resolution only, allowing a fast and simple implementation on a moderate CPU. The efficiency stems from the axis-aligned iterative local scaling that completely resolves the foldover artifact without any cumbersome nonlinear constraints. The key aspect of the proposed method is that the distortion from the image resizing is smoothly distributed along the resizing direction. It enforces the continuity of the respecting structures within the image such as straight lines or arches after deformation.

The rest of this paper is organized as follows. Section 2 describes the proposed retargeting method. Section 3 demonstrates the performance of the proposed method by comparing with state-of-the-art methods. Finally, we conclude the paper in Section 4.

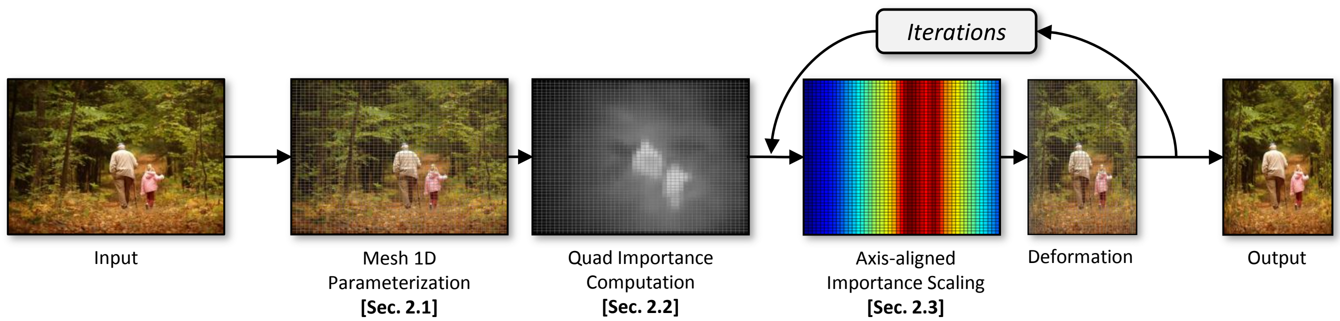


Fig. 1. Algorithm overview.

2. PROPOSED METHOD

For ease of algorithm exposition, we only discuss the resizing in horizontal direction in this paper. Application on the other cases is straightforward. The objective is that the image I' of the target width W' and height H' is reconstructed from the initial image of width W and height H. For this, the retargeting problem is posed as controlling a quad spacing in the space of axis-aligned deformations, similar to [6]. The proposed method consists of two steps, the quad importance computation and the axis-aligned importance scaling (AIS) process as shown in Fig. 1. Optimal local width parameters are computed via the AIS for obtaining the new mesh geometry, where each quad is adaptively scaled in an iterative manner. Therefore, the distortion is spread according to the significance of each quad during iteration.

2.1. Mesh 1-D Parameterization

An image I can be represented as $M \times N$ quads overlaid onto the image. Previous retargeting methods [5] typically parameterized the deformations in the 2D, leading to optimization problems with $M \times N$ unknowns. In contrast, in the space of axis-aligned deformations, the deformation is parameterized in the 1D, where every quad on the same column (or row) shares the same parameter for the width (or the height). The number of variables is thus linear in the size of the image boundary, i.e., the unknown variables are only $O(M+N)$. Furthermore, controlling the amount of deformation is simple, since it merely involves the changes of quad spacing. Hence, the overall manipulation of mesh structure always can be done in a feasible way.

To build a mesh structure, a uniform grid is overlaid over the image with N columns and M rows. The image I can then be represented as a mesh $G = (V_w, V_h)$ with a set of widths of the columns V_w and a set of heights of the rows V_h, where $V_w = \{w_1, w_2, \ldots, w_N\}$ and $V_h = \{h_1, h_2, \ldots, h_M\}$. $w_i \in \mathbb{R}$ and $h_i \in \mathbb{R}$ denote the unknown width and height of the columns and the rows, respectively. From the definition of a mesh structure, the image is partitioned into axis-aligned quads. Accordingly, the deformation is defined by the 1D vector of unknowns $V = (V_w, V_h)^T \in \mathbb{R}^{M+N}$. Our goal is then to infer the unknowns of V_w (V_h for resizing height) for the desired deformed mesh geometry.

2.2. Quad Importance

Our strategy to retarget an image is that the geometric structure should be preserved in salient regions, and *vice versa*. To this end, the amount of deformation is guided by a quad importance map that characterizes the visual attractiveness of each quad, i.e., each quad is assigned its importance value by averaging per pixel saliency values within the quad. We adapt the automatic saliency detection algorithm [1] to capture the amount of saliency at each pixel described as the saliency map. It ranges from 0 to 1, and smaller values denote less importance regions.

2.3. Iterative AIS

2.3.1. Deformation Error

The changes of aspect ratio for the quad should be penalized, since it is preferred that the original aspect ratio of salient regions is preserved after deformation. Let us define the transition cost for the deformed quad at t iteration as follows:

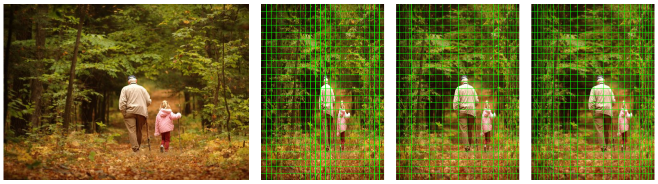
$$d^t(i) = (\frac{w_i^t N}{W} - 1)^2, \tag{1}$$

where w_i^t represents the inferred width of column i at t-th iteration. The transition cost penalizes all local deformations.

For measuring the deformation error of axis-aligned quads at t iteration, the transition cost is aggregated along the column as follows:

$$R_j^t = e^{-\frac{d^t(j)}{\sigma}} \sum\nolimits_{i \in M} S_{i,j}, \tag{2}$$

where $S_{i,j}$ represents the quad importance value locating at (i, j) in the grid and σ controls the sensitivity of the transition cost. The deformation error increases when axis-aligned



(a) Original (b) Initial (c) Iter. 1 (d) Iter. 5

Fig. 2. The iterative inference process. (a) an original image, (b) initial grid, (c) result at iteration 1, and (d) result at iteration 5.

quads have salient information. With this, the retargeting problem is solved by iteratively scaling the quad spacing in proportional to the deformation error value, such that the quad having a high saliency is likely to preserve its original aspect ratio.

2.3.2. *Iterative AIS*

To find the desired deformation, an iterative process is leveraged in the proposed AIS scheme in which mesh parameters V_w^t at time t is updated from an initial guess V_w^0. Initial guess V_w^0 is set to uniformly scaled parameters, i.e., $\frac{W'}{N}$, as shown in Fig. 2(b).

Once the deformation error at the current iteration t is calculated, the optimal mesh parameters $V_w^t = \{w_i^t | 1 \leq i \leq N\}$ can be simply computed as follows:

$$w_i^t = W' \frac{R_i^t + \epsilon_i^t}{\sum\limits_{k \in N} R_k^t}, \tag{3}$$

where ϵ_i^t enforces the deformation stability at i-th column that enforces the smoothness of the resulting deformation. Let us define the index set $\Psi^t = \{i | R_i^t \leq \tau M\}$, where τ is a threshold value that controls the grid smoothness. The deformation stability term ϵ_i^t is defined as follows:

$$\varepsilon_i^t = \begin{cases} \lambda_i & \text{if } i \in \Psi^t, \\ -\frac{1}{|\Psi^t|} \sum_{j \in \Psi^t} \lambda_j & \text{otherwise}, \end{cases} \tag{4}$$

where $\lambda_i = \tau M - R_i^t$ denotes the amount of error compensation for preventing the abrupt deformation at i-th column, and $|\Psi^t|$ is the cardinality of the set Ψ^t. The deformation stability term as in (4) prevents the low saliency region from being deformed abruptly. This term resolves the limitation of the conventional retargeting methods: the performance of conventional retargeting methods largely depend on the quality of the saliency map. That is, it enforces robustness to the retargeting process by evenly distributing the deformation across the image. Fig. 3 demonstrates this property. It shows the behavior of the proposed method according to the deformation stability term, i.e., varying threshold τ. It also shows the results of the inter-row coherence filtering (ICF) method [8]. The ICF method directly computes the scaling factors for each pixel from the saliency map. Therefore, as shown in Fig. 3(a), the region at the right wing of an eagle is distorted since saliency values at that region abruptly fall off to zero value. In our method, however, the amount of deformation is smoothly propagated across the image as τ increases.

(a) ICF [8] (b) $\tau = 0$ (c) $\tau = 0.05$ (d) $\tau = 0.1$

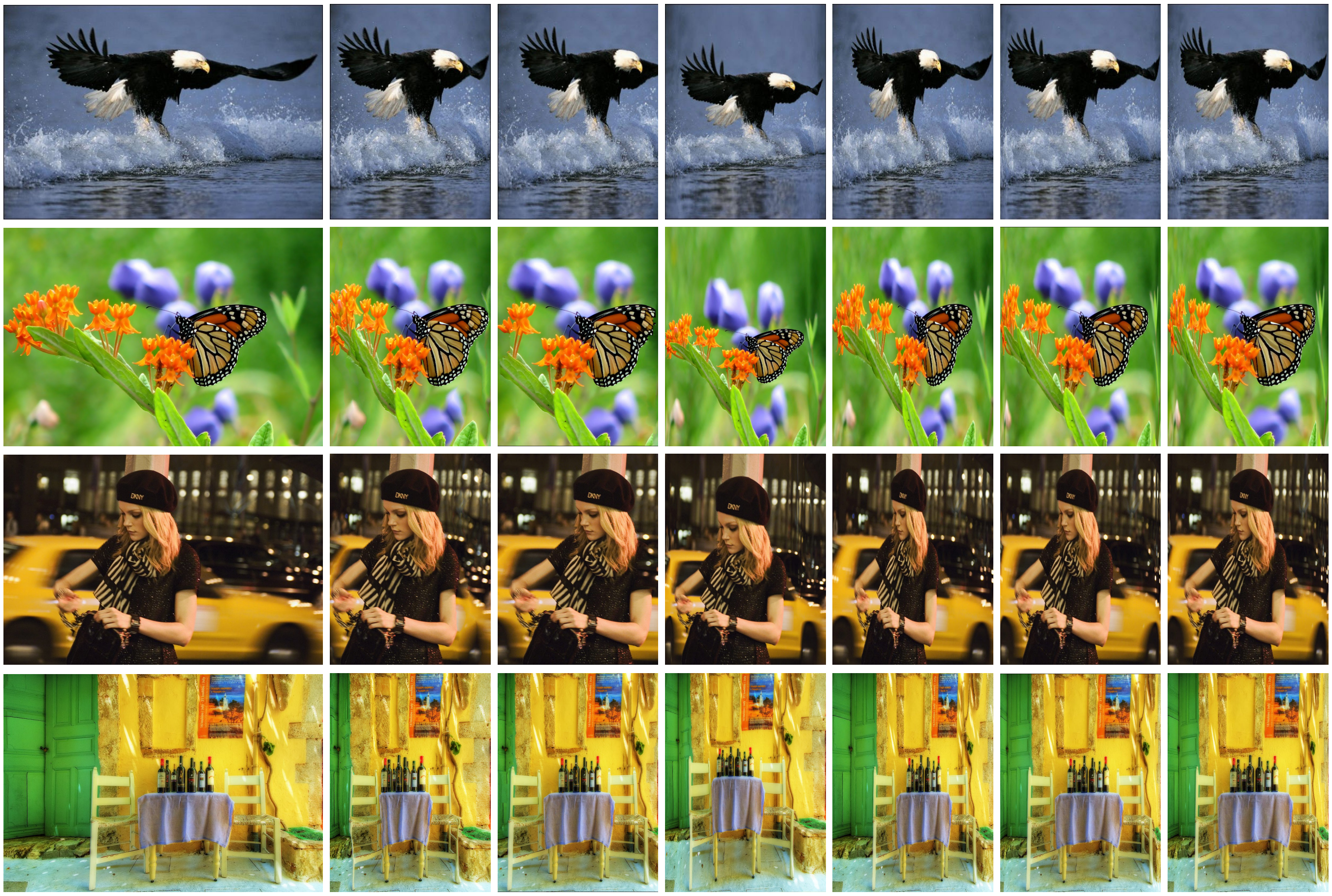
Fig. 3. The behavior of the proposed method according to the deformation stability term. The results of (a) ICF method [8], and (b)-(d) the proposed method with varying threshold τ.

The new mesh parameter V_w^t obtained from (3) is used to update the image I^{t+1} for the next iteration. In this way, the deformed mesh structure is refined incrementally, until there is no significant changes in the mesh parameters or to some fixed iteration number, as shown in Fig. 2. In our experiments, it has been observed that the optimal mesh parameters can be obtained within 5 iterations, achieving 0.02 ms for convergence. Thanks to the 1D parametrization of a mesh structure and efficient calculation for the new geometry with AIS, the proposed method typically takes a few milliseconds on the single-core CPU.

3. EXPERIMENTAL RESULTS

The experiment was performed with the publicly available `RetargetMe` datasets [7]. We followed the scenarios of [7] and compared the performance of the proposed method with that of state-of-the-art methods in the context of spatial quality and the running time. The proposed method was implemented in C++ without code optimization and was simulated on a PC with Quad-core CPU 2.93GHz. All parameters were fixed during experiments: the grid resolution $M \times N$ was set to 25×25. τ and σ were set to 0.1 and 3.0, respectively. The AIS terminates at 5 iterations.

Fig. 4 shows the results from the proposed method and competing algorithms. It demonstrates that the proposed method preserves the salient structure well, while other methods show some inconsistent appearance at structured objects. The pixel-wise methods such as the seam-carving (SC) remove pixels on the same object inconsistently. The scale-and-stretch (SNS) method distorts the vertical components of the scene, as shown in the bottom of Fig. 4, even though the constraints for preventing this artifact were modeled in the optimization formulation. In contrast, the door structures are



(a) Original image (b) SC [2] (c) SM [3] (d) SNS [5] (e) IF [4] (f) AD [6] (g) Proposed

Fig. 4. The results of the proposed method and competing algorithms. (a) Original image, (b) seam carving [2], (c) shift map [3], (d) scale-and-stretch [5], (e) importance filter [4], (f) axis-aligned deformation [6], and (g) the proposed method. All images were resized to half width.

well preserved in our method since the deformation operates on the axis-aligned space, resulting in the changes of quad spacing only.

The running time of the proposed method and competing algorithms are shown in Table 1. The proposed method finishes the retargeting process including inference and rendering within 0.04 ms, when the image with the resolution 1024×700 is resized to half width. It can be observed that the proposed method is much faster than the original axis-aligned deformation (AD) method [6], and produces competitive results as shown in Fig. 4(f). Furthermore, the iterative AIS works on a single-core CPU without parallel processing, which allows the interactive image resizing with a simple implementation.

Table 1. Complexity comparisons

Method	Time (ms)
Seam Carving (SC) [2]	102000
Shift Map (SM) [3]	45000
Importance Filter (IF) [4]	80
Scale and Stretch (SNS) [5]	36
Axis-aligned Deformation (AD) [6]	4
Proposed method	0.04

4. CONCLUSION

In this paper, we have proposed a retargeting method via adaptively scaling the 1D mesh structure for meeting the target resolution. The computation of optimal deformation parameters was iteratively inferred by scaling the quad parameters with the proportional to the deformation error. The deformation stability term was introduced for preventing abrupt deformation at the low saliency region. Thus, the deformation error is evenly distributed across the image. In addition, the computational complexity of the proposed method is very low, enabling its practical uses in various applications for modest systems.

5. REFERENCES

[1] J. Harel, C. Koch, and P. Perona, “Graph-based visual saliency,” *Advances in Neural Information Processing Systems*, pp. 545-552, MIT Press, 2006.

[2] M. Rubinstein, A. Shamir, and S. Avidan, “Improved seam carving for video retargeting,” *ACM Trans. Graphics (TOG)*, vol. 27, no. 3, pp. 16, 2008.

[3] Y. Pritch, E. Kav-Venaki, and S. Peleg, “Shift-map image editing,” *in Proc. Int’l Conf. Comput. Vis. (ICCV)*, pp. 151-158, 2009.

[4] Y. Ding, J. Xiao, and J. Yu, “Importance filtering for image retargeting,” *in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 89-96, 2011.

[5] Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee, “Optimized scale-and-stretch for image resizing,” *ACM Trans. Graphics (TOG)*, vol. 27, no. 5, pp. 118, 2008.

[6] D. Ranozzo, O. Weber, and O. Sorkine, “Robust image retargeting via axis-aligned deformation,” *Computer Graphics Forum*, vol. 31, pp. 229-236, 2012.

[7] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir, “A comparative study of image retargeting,” *ACM Trans. Graphics (TOG)*, vol. 29, no. 6, pp. 160, 2010.

[8] K. Sun, B. Yan, and Y. Gao, “Low complexity content-aware image retargeting,” *in Proc. IEEE Int’l Conf. Image Process. (ICIP)*, 2012.