2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)

IMAGE RETARGETING WITH PROTECTION OF OBJECT ARRANGEMENT

Kazu Mishiba and Takeshi Yoshitome

Department of Electrical and Electronic Engineering, Tottori University, Tottori, Japan

ABSTRACT

The relative arrangement, such as relative positions and orientations among objects, can play an important role in expressing the situation such as sports games and race scenes, etc. In this paper, we propose a retargeting method that allows maintaining the relative arrangement. Our proposed retargeting method is based on a warping method which finds an optimal transformation by solving an energy minimization problem. To achieve protection of object arrangement, we introduce an energy that enforces all the objects and the relative positions among these objects to be transformed by the same transformation in the retargeting process. Experimental results demonstrate that our proposed method maintains the relative arrangement while protecting important regions.

***Index Terms*—** content-aware image retargeting, warping method, image resize, object arrangement

1. INTRODUCTION

As the diversity of electronic display devices increases (e.g., smart phones, notebooks, tablets, and HDTV), it is increasingly important that visual media display appropriately on these devices. Because the same image needs to be displayed with different resolutions on different devices, image resizing plays an important role in appropriate display. The classical resizing methods of homogeneous scaling and cropping have the drawbacks of distorting the entire image and discarding important parts of the image, respectively. Content-aware image retargeting [1, 2] has recently been developed to overcome these limitations. Several retargeting methods have been proposed, including seam carving [3, 4] and image warping [5, 6]. The key concept of these techniques is to resize images to an arbitrary resolution while protecting visually important regions from distortion. Seam carving changes the size of a source image by iteratively carving out or inserting paths of pixels, which is called *seam*. A seam is selected to pass through less important regions to keep important regions unchanged. Warping methods place a mesh onto an image and then deform the mesh by solving an optimization problem. Many methods use a regular grid mesh and few methods [6] use an irregular triangle mesh.

Recently developed methods have focused only on protection of the shape of visually important regions. Therefore, conventional methods often shrink or enlarge unimportant regions among objects, resulting in the change of the distance among these objects. Fig. 1 (b) is an example of the reduction of the image width using a conventional method. Although the retargeting result may be satisfactory to a viewer, it can be unsatisfactory to a photographer if one considers that the distance between the main objects (horses) has a significant meaning in the image. The distance among objects can play an important role in expressing the situation such as sports games and race scenes like Fig. 1 (a), etc. Thus, the change of a relative arrangement, such as relative positions and orientations among objects, can distort the intended content. In this paper, we propose a retargeting method that allows maintaining the relative arrangement among objects.

(a) (b) (c)

Fig. 1. The relative distance between the horses on the source image (a) is changed in the retargeting image (b) using seam carving [4]. The retargeting image (c) using our method maintains the relative distance.

2. PROPOSED METHOD

Our proposed method aims to achieve protection of the relative arrangement among objects on an image in retargeting. In this paper, the relative arrangement between object ω_i and object ω_j is defined by two parameters $(\theta_{i,j}, r_{i,j})$: $\theta_{i,j}$ is the

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

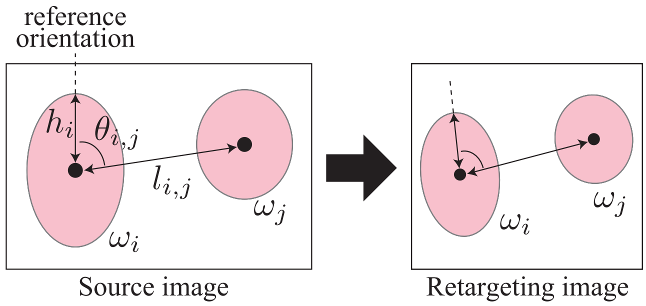

Fig. 2. Example of the protection of object arrangement.

relative direction between the reference orientation of ω_i and the direction of ω_j from ω_i, and $r_{i,j}$ is the ratio of the distance $l_{i,j}$ between ω_i and ω_j relative to the size h_i of ω_i, i.e., $r_{i,j} = h_i/l_{i,j}$ (see Fig. 2).

Our proposed method introduces other two conditions to produce more satisfactory results. The first condition is that the shape of important regions is protected. The second condition is that extreme deformation is avoided.

2.1. Problem Formulation

Our proposed retargeting method is based on a warping method. In our warping method, a mesh is projected onto the image and then deformed to a target size to satisfy desired conditions as much as possible by solving an energy minimization problem. An initial mesh $\mathbf{M} = \{\mathbf{V}, \mathbf{F}\}$ is represented as a set containing two components. $\mathbf{V} = \{\boldsymbol{v}_i\}_{i=1}^{\alpha}$ is a set of vertex positions where α is the number of vertices and $\boldsymbol{v}_i = [v_{i,x}, v_{i,y}]^T$ is a vector consisting of x and y coordinate values. $\mathbf{F} = \{f_i\}_{i=1}^{\beta}$ is a set of faces where β is the number of faces. Our method can use both triangle and quadrilateral mesh. Let us denote an element after retargeting by adding a prime mark, e.g., $\boldsymbol{v}'_i$ indicates a vertex position after retargeting. The transformation of an image of size $W \times H$ to $W' \times H'$ is a problem of finding optimal vertex positions $\mathbf{V}'_{opt}$ by solving an energy minimization problem.

The energy minimization problem consists of energy functions E and boundary conditions. E increases less as the deformed mesh better satisfies a desired condition. The optimal mesh transformation is expressed as

$$\mathbf{V}'_{opt} = \arg\min_{\mathbf{V}'} \sum_{z \in Z} \lambda_z E_z \tag{1}$$

subject to boundary conditions:

$$\begin{cases} v'_{i,x} = 0 & \text{if} \quad v_{i,x} = 0, \\ v'_{i,x} = W' & \text{if} \quad v_{i,x} = W, \\ v'_{i,y} = 0 & \text{if} \quad v_{i,y} = 0, \\ v'_{i,y} = H' & \text{if} \quad v_{i,y} = H. \end{cases} \tag{2}$$

Here Z is a set of desired conditions, E_z is an energy function for condition z and λ_z is a weight parameter for E_z. Using a larger λ_z enhances the corresponding condition.

2.2. Energy Definition

We use the following three conditions to be satisfied in the retargeting process:

- maintaining the relative arrangement among objects,
- protection of important regions,
- avoiding extreme deformation.

This subsection describes the following three energies corresponding to the above conditions: *arrangement energy*, *rigid energy*, and *smoothing energy*.

Arrangement energy. First, we describe which transformation maintains the relative arrangement among objects. Let us take the case of two objects, ω_i and ω_j, as an example. Let us represent the position of ω_i by the center of gravity $\boldsymbol{g}_i^{\omega}$ and denote the relative position of ω_j relative to ω_i by $\boldsymbol{g}_{ij}^{\omega} = \boldsymbol{g}_j^{\omega} - \boldsymbol{g}_i^{\omega}$. When ω_i is transformed by an affine transformation T that rotates the reference orientation and scales the size of ω_i, the relative position $\boldsymbol{g}_{ij}^{\omega}$ should be transformed by T to maintain the relative arrangement ($\theta_{i,j}$, $r_{i,j}$). The same holds for ω_j. Thus, we can see that all the objects and all the relative positions among objects should be transformed by the same transformation T to maintain the relative arrangement among objects. Fig. 2 shows an example of the protection of relative arrangement. In addition, T should be a transformation that maintains the aspect ratios of objects to avoid distortion. Thus, a similarity transformation is used as T.

Next, we define the transformation and the center of gravity of an object to formulate the arrangement energy. They can be expressed using faces because an object is on faces. The transformation of an object can be approximated by the transformation of the faces where the object is on. The n edge vectors constituting the face f are deformed from $\{\boldsymbol{e}_i^f | i = 1, ..., n\}$ to $\{\boldsymbol{e}_i^{f'} | i = 1, ..., n\}$ in the retargeting process. The transformation of the face can be computed as the linear transformation of the mapping from before to after retargeting:

$$\boldsymbol{J}_f = \left[\boldsymbol{e}_1^{f'}, ..., \boldsymbol{e}_n^{f'}\right] \left[\boldsymbol{e}_1^{f}, ..., \boldsymbol{e}_n^{f}\right]^+ \tag{3}$$

where $[\cdot]^+$ is an operator that calculates $\boldsymbol{A}^+ = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T$ for matrix $\boldsymbol{A}$. This mapping is a linear approximation for a quadrilateral face. The center of gravity of an object can be approximated by the weighted average of the center of gravity of the faces where the object is on. The center of gravity of the face f can be calculated by

$$\boldsymbol{g}^f = \frac{1}{|\mathbf{V}(f)|} \sum_{\boldsymbol{v} \in \mathbf{V}(f)} \boldsymbol{v} \tag{4}$$

where $\mathbf{V}(f)$ is a set of the vertices constituting the face f and $|\mathbf{V}(f)|$ is the number of elements in $\mathbf{V}(f)$. The center of gravity of an object can be expressed as

$$\boldsymbol{g}^{\omega} = \frac{1}{\sum_{f \in \mathbf{F}(\omega)} a_f} \sum_{f \in \mathbf{F}(\omega)} a_f \boldsymbol{g}^f \tag{5}$$

where a_f is an area of f. The center of gravity of an object after retargeting is approximately calculated using a_f instead of a'_f.

Finally, we define the arrangement energy E_a for maintaining the relative arrangement among objects. From above discussion, we can see that the following two conditions should be satisfied to maintain the relative arrangement. The first condition is that all the faces where the objects are on are transformed by the same transformation T. The second condition is that all the relative positions among objects are transformed by the same T. Thus, E_a can be expressed as the sum of two energies corresponding to these conditions.

The first condition is expressed as the energy

$$E_{a1} = \sum_{\omega \in \Omega} \sum_{f \in \mathbf{F}(\omega)} a_f \|\boldsymbol{T}_{\Omega} - \boldsymbol{J}_f\|_F^2 \tag{6}$$

where Ω is a set of all objects, $F(\omega)$ is a set of the faces where the object ω is on, $\|\cdot\|_F$ is the Frobenius norm of a matrix, and $\boldsymbol{T}_{\Omega}$ is a similarity transformation:

$$\boldsymbol{T}_{\Omega} = \begin{bmatrix} c & -d \\ d & c \end{bmatrix}. \tag{7}$$

E_{a1} increases more as the transformation of a face is further deformed from T_{Ω}.

The second condition is expressed as the energy

$$E_{a2} = \sum_{(\omega_i, \omega_j) \in \Phi_{\Omega}} l_{i,j} \|\boldsymbol{T}_{\Omega} \boldsymbol{g}_{ij}^{\omega} - \boldsymbol{g}_{ij}^{\omega\prime}\|_2^2 \tag{8}$$

where $\|\cdot\|_2$ is the ℓ^2-norm of a vector, $l_{i,j} = \|\boldsymbol{g}_{ij}^{\omega}\|_2$ and $\Phi_{\Omega} = \{(\omega_i, \omega_j) \in \Omega \times \Omega | \omega_i \neq \omega_j\}$ is a set of all the possible ordered pairs of all objects. E_{a2} increases more as the transformation of the relative position $\boldsymbol{g}_{ij}^{\omega}$ is further deformed from T_{Ω}.

Rigid energy. We introduce the rigid energy E_r to protect important regions. To achieve protection of important regions, higher importance faces are to be less transformed. This condition can be formulated as

$$E_r = \sum_{f \in \mathbf{F}} s_f \|\boldsymbol{J}_f - \boldsymbol{I}\|_F^2 \tag{9}$$

where $\boldsymbol{I}$ is an identity matrix and s_f is an importance of f that is calculated by summing the pixel importance on f. E_r increases more as a face is further deformed.

Smoothing energy. We introduce the smoothing energy E_s to avoid extreme deformation. To achieve smooth transformation, the transformation of a face should be as similar as possible to the transformation of its adjacent faces. This condition can be formulated as

$$E_s = \sum_{f_i \in \mathbf{F}} \sum_{f_j \in N(f_i)} (a_{f_i} + a_{f_j}) \|\boldsymbol{J}_{f_i} - \boldsymbol{J}_{f_j}\|_F^2 \tag{10}$$

where $N(f_i)$ is a set of adjacent faces that share an edge with f_i. E_s increases more with an increasing difference in the transformation of a face between its adjacent faces.

2.3. Implementation

Our proposed method first loads an image, and then objects subject to protection of the relative arrangement are manually selected by a user. Next, the pixel importance is computed using the saliency detection [7]. This method produces a favorable result with a relatively low computational complexity. Importance values are normalized in the range of $[0, 1]$. Finally, an energy minimization problem is solved to obtain an optimal mesh transformation. The energy minimization problem is formulated as follows:

$$\mathbf{V}'_{opt} = \arg\min_{\mathbf{V}'} \lambda_a E_a + \lambda_r E_r + \lambda_s E_s \tag{11}$$

subject to Eq. (2) where $E_a = E_{a1} + E_{a2}$. E_{a1} and E_{a2} can respectively be transformed into the following:

$$E_{a1} = \sum_{\omega \in \Omega} \sum_{f \in \mathbf{F}(\omega)} a_f \left\| \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \boldsymbol{t} - \mathrm{vec}(\boldsymbol{J}_f) \right\|_2^2, \tag{12}$$

$$E_{a2} = \sum_{(\omega_i, \omega_j) \in \Phi_{\Omega}} l_{i,j} \left\| \begin{bmatrix} g_x & -g_y \\ g_y & g_x \end{bmatrix} \boldsymbol{t} - \begin{bmatrix} g'_x \\ g'_y \end{bmatrix} \right\|_2^2 \tag{13}$$

where $\boldsymbol{t} = [c, d]^T$, $\boldsymbol{g}_{ij}^{\omega} = [g_x, g_y]^T$, and vec$(\cdot)$ is an operator that creates a column vector from a matrix by stacking the column vectors of the matrix. Since vec$(\boldsymbol{J}_f)$ and $\boldsymbol{g}_{ij}^{\omega}$ are represented as a linear combination of the vertex positions after retargeting $\boldsymbol{x} = [v'_{1,x}, ..., v'_{\alpha,x}, v'_{1,y}, ..., v'_{\alpha,y}]^T$, E_a can be written as the following quadratic form in $\boldsymbol{x}$:

$$||\boldsymbol{Bt} - \boldsymbol{Ax}||_2^2. \tag{14}$$

The optimal $\boldsymbol{t}$ can be found by solving $\boldsymbol{Ax} = \boldsymbol{Bt}$ in a least square sense: $\boldsymbol{t} = \boldsymbol{B}^+ \boldsymbol{Ax}$. Thus, E_a can be written as follows:

$$\left|\left|(\boldsymbol{I} - \boldsymbol{BB}^+)\boldsymbol{Ax}\right|\right|_2^2. \tag{15}$$

Since E_r and E_s are also quadratic with respect to the vertex positions after retargeting, the solution of Eq. (11) can be found by solving a sparse linear system.

3. EXPERIMENTAL RESULTS AND DISCUSSION

To evaluate our method, we have implemented the proposed method and tested it on a variety of images. We used a uniform grid mesh divided by 20×20 faces. Hence, each face is of $W/20 \times H/20$ pixels. Weight parameters are set to $\lambda_a = 20$, $\lambda_r = 2$, and $\lambda_s = 1$. We compared our method with Rubinstein's seam carving method [4] and Wang's warping method [5]. Our method uses a user input for selection of regions subject to protection of the relative arrangement. To achieve a fair comparison, the pixel importance on the selected regions was set to a maximum value to reduce distortion in the comparison methods [4] and [5].

Fig. 3 represents retargeting results of image reduction. Both our proposed method and the conventional methods protect the shape of the objects. In addition, our method maintains the relative arrangement among objects while the conventional methods change it. As shown in Fig. 4, our method also produces satisfactory results when enlarging images. While the conventional methods expand the relative distance between objects, our method appropriately maintains it.

Fig. 5 represents retargeting results using quadrilateral and triangle mesh. As can be seen from Fig. 5, our energy definition can be used for both types of mesh. In addition, all the energies including the arrangement energy are expressed as a quadratic form, which is commonly used in warping methods. Therefore, the arrangement energy can be incorporated into conventional warping methods.

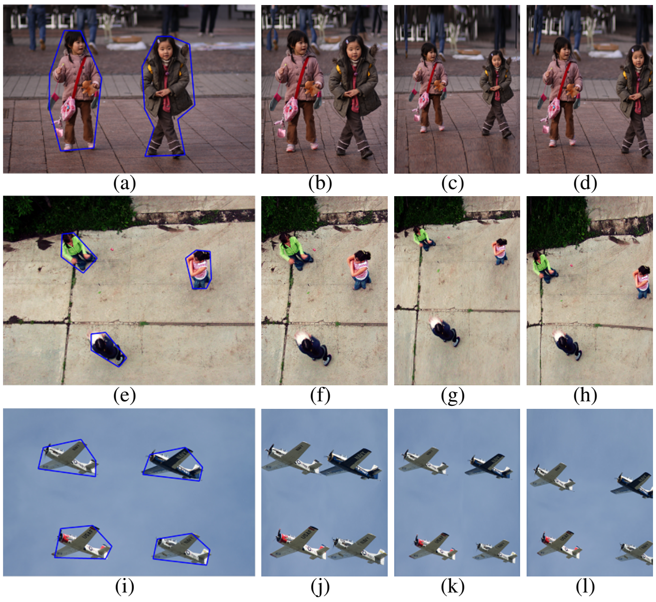

Fig. 3. Retargeting results of image reduction. (a), (e) and (i) Source images with selected regions (blue polygons). (b), (f) and (j) Retargeting image using seam carving [4]. (c), (g) and (k) Retargeting image using warping method [5]. (d), (h) and (l) Retargeting image using our method.

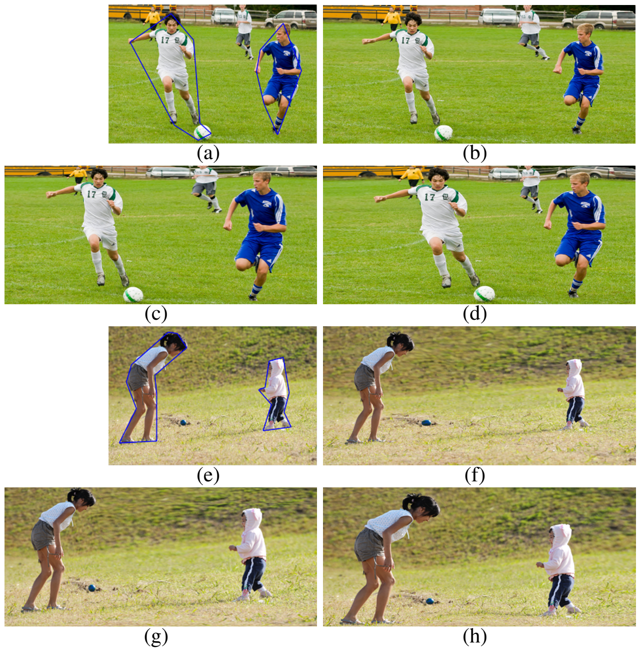

Fig. 4. Retargeting results of image enlargement. (a) and (e) Source images with selected regions (blue polygons). (b) and (f) Retargeting image using seam carving [4]. (c) and (g) Retargeting image using warping method [5]. (d) and (h) Retargeting image using our method.

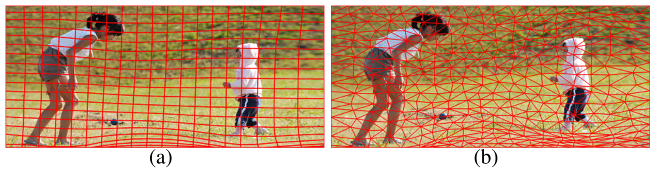

Fig. 5. Retargeting images with quadrilateral mesh (a) and triangle mesh (b).

4. CONCLUSIONS

We proposed a novel method for retargeting images while maintaining the relative arrangement among objects. To achieve protection of object arrangement, we introduced the arrangement energy that enforces all the objects and the relative positions among these objects to be transformed by the same transformation in the retargeting process. Experimental results demonstrate that our proposed method maintains the relative arrangement while protecting important regions.

Acknowledgment

We thank the following Flickr (http://www.flickr.com/) users for Creative Commons imagery: Stephan Pabst (horse race), gwenael.piaser (two little tourists), shioshvili (old witch), Mike Miley (Formation Flight), B Mlry (Soccer). This work was supported by JSPS KAKENHI Grant Number 25820159.

5. REFERENCES

[1] D. Vaquero, M. Turk, K. Pulli, M. Tico, and N. Gelfand, "A survey of image retargeting techniques," in *Proc. SPIE*

Applications of Digital Image Processing XXXIII, 2010, vol. 7798.

[2] A. Shamir and O. Sorkine, “Visual media retargeting,” in *SIGGRAPH ASIA Courses*, New York, NY, USA, 2009, pp. 1–13, ACM.

[3] S. Avidan and A. Shamir, “Seam carving for content-aware image resizing,” *ACM Trans. Graph.*, vol. 26, no. 3, pp. 10, 2007.

[4] M. Rubinstein, A. Shamir, and S. Avidan, “Improved seam carving for video retargeting,” *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–9, 2008.

[5] Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee, “Optimized scale-and-stretch for image resizing,” *ACM Trans. Graph.*, vol. 27, no. 5, pp. 118:1–118:8, 2008.

[6] Y. Guo, F. Liu, J. Shi, Z.-H. Zhou, and M. Gleicher, “Image retargeting using mesh parametrization,” *IEEE Trans. Multi.*, vol. 11, no. 5, pp. 856–867, 2009.

[7] Xiaodi Hou and Liqing Zhang, “Saliency detection: A spectral residual approach,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2007, pp. 1–8.