

## Technical Section

Aesthetic photo composition by optimal crop-and-warp<sup>☆</sup>Yong Jin, Qingbiao Wu<sup>\*</sup>, Ligang Liu

Department of Mathematics, Zhejiang University, Hangzhou 310027, China

## ARTICLE INFO

## Article history:

Received 7 November 2011

Received in revised form

24 July 2012

Accepted 24 July 2012

Available online 21 August 2012

## Keywords:

Computational aesthetics

Image composition

Image retargeting

Quadratic programming

Triangular mesh

## ABSTRACT

We present a high-efficiency approach to optimally enhance image composition using a crop-and-warp-based algorithm. We have designed a piecewise quadratic aesthetic energy function with linear inequality constraints that measures the distance between original positions and the aesthetic target positions of visual elements, to assess the target image composition. The target image is cropped by a sub-window that can zoom and move within the original image window. The optimal cropped window position can be obtained by solving a piecewise quadratic program that attempts to minimize the energy function. The salient object and feature line locations in the cropped image are further optimally adjusted by adopting a triangular mesh-based warping technique, which also allows fitting the proposed approach for image retargeting. A quadratic optimization controls the mesh warping and can be achieved by solving a sparse linear system. We illustrate the effectiveness of our approach in several experimental results and compare them to previous approaches.

Crown Copyright © 2012 Published by Elsevier Ltd. All rights reserved.

## 1. Introduction

Enhancing the aesthetic appearance of input photos utilizing computer technology is a research focus in a growing field called Computational Aesthetics, which emphasizes computational techniques for making similar aesthetic decisions as humans [1]. Nevertheless, judging photographic aesthetics is always subjective and includes many factors, such as color, tone, illumination and composition [2–4].

This paper focuses on enhancing image composition, as composition represents the main harmony that an image/photo seeks to express to viewers. Expert photographers prefer to use professional composition rules to which amateurs are rarely sensitive when taking photos. Such rules have been systematically described in photography courses or textbooks [5,6] as guidelines to increase aesthetic image appreciation.

Photos taken without considering compositional guidelines might lack aesthetics or even be powerless to demonstrate what the photographers intend to tell the viewers. After a camera produces a photo, users find it tedious to enhance the composition without complicated digital editing. Assisted by commercial tools such as Adobe Photoshop, a user might crop the image, extract foreground objects and paste them back into the image at targeted positions to enhance the composition. However, the

procedure might be filled with inaccuracy and complexity, while being un-reproducible for another image.

Automatic techniques for aesthetic image composition enhancement have also been researched and developed. Liu et al. [7] have translated several basic composition guidelines into quantitative aesthetic scores, based on which an automatic crop-and-retarget approach to enhancing the image composition has been developed. The method presented by Liu et al. [7] searches for the optimal composition result in a 4D space; the 4D space contains all cropped windows with various widths and heights (2D) as well as the cropping window positions (2D), which include all possible cropping and warping results. This is actually a passive methodology. Such a passive methodology guarantees obtaining the optimal solution, but it is slow and inefficient.

To speed up the algorithm, a real-time image composition enhancement technique based on triangular mesh warping has also been presented [8]. This method [8] computes the optimal composition result by warping the salient objects to target positions and can be considered an active methodology. Such a method is efficient and fast. However, it may cause large distortions in the results, as it does not perform any cropping operator to remove unnecessary parts from the original photo.

To consider effectiveness [7] and efficiency [8] simultaneously, we present an improved approach to optimally enhance image composition using a crop-and-warp-based algorithm. We have designed an active crop operator to replace the passive crop operator [7] to speed up the algorithm without sacrificing effectiveness. Moreover, the warp operator, which was derived from a previous version [8], has been improved to significantly reduce

<sup>☆</sup>This article was recommended for publication by Tobias Isenberg.

<sup>\*</sup> Corresponding author. Tel.: +86 571 87953665.

E-mail addresses: [jsnjkl@zju.edu.cn](mailto:jsnjkl@zju.edu.cn) (Y. Jin), [qbwu@zju.edu.cn](mailto:qbwu@zju.edu.cn) (Q. Wu), [ligangliu@zju.edu.cn](mailto:ligangliu@zju.edu.cn) (L. Liu).

the opportunity to cause severe distortions and self-intersections. The crop operator can be transformed into a piecewise quadratic program that attempts to minimize the energy function, while a quadratic optimization can control the warp operator, which can be achieved by solving a sparse linear system. Our method combines the advantages of both methods (e.g., [7,8]) and can thus efficiently obtain good results.

## 2. Related work

This section briefly reviews state-of-the-art techniques related to aesthetic image composition and image retargeting.

*Aesthetic Image Composition:* Image composition is determined by the combinations and locations of visual elements inside an image frame, including salient objects and feature lines. Professional photographers prefer to take photos using many defined composition rules [2] to enhance the aesthetic appearance of their products. Although no absolute rules can ensure aesthetic composition in arbitrary circumstances, some heuristic principles exist that provide strategies to achieve eye-pleasing composition when utilized properly, including the rule of thirds, shapes and lines, amputation avoidance, visual balance and diagonal dominance [5–7].

Attempts to allow automatic image cropping or capturing for visual quality enhancement have been made throughout the last decade. Suh et al. [9] have proposed a set of fully automatic image cropping techniques based on the visual saliency model proposed by Itti et al. [10]. Gooch et al. [11] utilize the rules of thirds and fifths to find the aesthetic viewpoint and layout for an image of a 3D object. Byers et al. [12] have developed an autonomous robot system that takes well-composed photographs of people using the rule of thirds. Another compositional guideline that describes how features should be balanced from left to right has been used to arrange images and texts properly in one window [13]. Santella et al. [14] present an interactive method for cropping photographs given minimal information about the location of important content based on eye tracking.

Recently, Liu et al. [7] have proposed an approach to optimize image composition. Several well-known composition guidelines, including the rule of thirds, diagonals, visual balance, and region size, are quantified as aesthetic composition scores. The algorithm searches for the optimized composition with the highest aesthetic score in the 4D space. Their subsequent work [8] presents a triangular mesh warping-based technique to achieve the proposed aesthetic composition by moving visual elements towards their aesthetic target positions.

*Image Retargeting:* Image Retargeting seeks to display images on resized screens or frames with different aspect ratios than the original while preserving salient objects and features, which has been discussed in a work called Seam Carving [15].

The Seam Carving methods [15,16] iteratively remove less noticeable seams to fit the target image size aspect ratio. Rubinstein et al. [17] have proposed a multi-operator technique combined with Seam Carving and Scaling. Setlur et al. [18] segment an image into regions and identify salient objects. Salient objects are then cut and pasted into the resized image frame, in which missing background regions are filled using inpainting techniques. Quad mesh warping-based techniques have been widely applied to achieve content-aware image retargeting [19–21]. A quad mesh is placed over the original image, and an energy function is employed to preserve the aspect ratios of salient objects during the warping while fitting the mesh boundaries to the target image size. It is also feasible to develop algorithms based on triangular meshes [22,23]. Jin et al. [23] propose a method that can preserve feature lines and curves during image retargeting. More intensive descriptions of image retargeting techniques can be found in a related survey [24].

Our algorithm is an improved combination of the crop [7] and warp operators [8]. It overcomes the efficiency bottleneck in the passive methodology [7] and the effectiveness bottleneck in the active methodology [8] simultaneously. We suggest replacing the aesthetic score [7] with a piecewise quadratic energy function that can accelerate the crop operator, which has been abandoned due to its lack of efficiency [8]. We also modify the warp operator to preserve the relative locations of visual elements inside the image, which could improve the warp operator effectiveness and reliability. We try to denote the more relevant benefits and shortcomings of our approach and related techniques throughout the paper.

The rest of the paper is organized as follows. The next section overviews the algorithm. Before presenting the detailed algorithm, Section 4 briefly reviews the basic composition guidelines and discusses related image pre-processing techniques. Sections 5 and 6 propose and detail the crop and warp operators, respectively. Section 7 presents the experimental results and comparisons with previous techniques, and the last section summarizes our conclusions.

## 3. Overview

The proposed approach to optimize image composition combines two steps: crop and warp. Fig. 1 shows a concise overview of the approach. The crop operator tries to choose a sub-window in which visual elements optimally coincide with the basic composition guidelines. A sub-image (Fig. 1(b)) chosen by the optimal cropping window (the pink window in Fig. 1(a)) can be considered a composition improved version of the original image (Fig. 1(a)). The warp operator further optimizes the composition of the cropped image to obtain the final version (Fig. 1(c)).



**Fig. 1.** Overview of optimizing image composition. (a) Input image  $I_I$ ; the pink window is the optimal cropping window; (b) the optimally cropped and uniformly scaled image  $I_S$  from (a); (c) The optimally warped image  $I_U$  from (b). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

## 4. Preliminaries

In this section, we introduce basic composition guidelines for composing aesthetic images and image pre-analyses, including how to extract visual elements.

### 4.1. Basic composition guidelines

Various guidelines for taking well-composed photos can be found in photography textbooks [5,6]. We consider a primary subset of basic composition guidelines that are well-defined and widely disseminated, as in previous research [7,8].

**Rule of thirds:** The rule of thirds [5,6] maintains that the frame should be equally partitioned into thirds both vertically and horizontally. The four lines that separate the thirds are called power lines. The four intersection points formed by the thirds lines are called power points. Photographers are encouraged to place the main subjects around the power points. Moreover, photographers are encouraged to align the salient vertical and horizontal components or lines with the power lines in the image.

**Diagonal Dominance:** Salient diagonal objects or lines should be located along the corresponding diagonal line, thus creating a dynamic emphasizing effect [5]. The two diagonal lines are referred to as an additional two power lines.

**Visual Balance:** Visual balance is a vital component to image composition [6]. Salient objects should be placed evenly around the center of the image frame to achieve a visually balanced effect.

**Region Size:** Professional photographers maintain their principle for determining salient region sizes. Liu et al. [7] have surveyed over 200 professional photos by studying the area fractions of the salient objects using the original frames. The fractions have three dominant peaks in statistics:  $\mathcal{F} = \{\mathbf{F}_1 = 0.1, \mathbf{F}_2 = 0.56, \mathbf{F}_3 = 0.82\}$ . This suggests that the area fractions of salient objects should be near the three peaks in a well-composed image.

### 4.2. Image pre-processing

The image composition should be assessed by analyzing the spatial structure and visual element distribution, including salient objects and feature lines inside the image. We introduce a brief algorithm to detect such elements.

**Salient Object Detection:** We employ an image segmentation and saliency patch by expanding another technique [7,18] to extract salient objects. We adopt a saliency map called GBVS [25] to assign a saliency value between 0 and 1 to each image pixel. The saliency value suggests which parts of the image are more important, which helps us recognize salient objects and determine their importance. Patches with high saliency scores (over a given threshold) are recognized as part of salient objects. The salient objects can be obtained by greedily expanding the corresponding patches.

**Feature Lines Detection:** Feature lines can be automatically detected using a Hough transformation [26], while small line segment fragments should be discarded. We consider feature lines that form an angle with an  $x$ -axis less than  $15^\circ$  as horizontal lines, and if they form an angle with a  $y$ -axis less than  $15^\circ$ , we consider them as vertical lines; otherwise, they are considered as diagonal lines. We also fit line segments to elongated salient objects, as in a previous study [7]. The algorithm should treat such elongated salient objects as feature lines.

Fig. 2 shows the saliency map and detected visual elements of the space shuttle image in Fig. 1(a). Because recognizing visual elements in images is difficult, our guideline for extracting a saliency map and visual elements may not work for arbitrary images. If our algorithms do not make sense in some cases, users

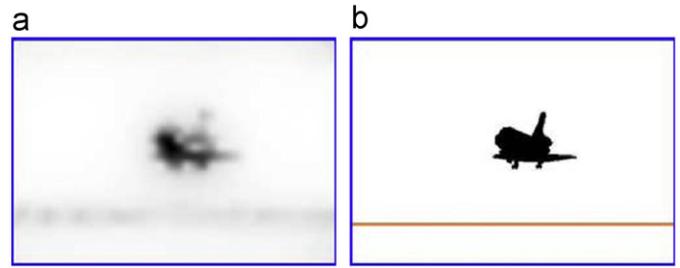


Fig. 2. Image pre-processing for the space shuttle image. (a) Saliency map (Darker regions have larger saliency values); (b) detected salient object and feature line.

can try other algorithms for this procedure, including a saliency map constructed from eye-tracking data [27,28] and visual element masks computed using the global contrast [29]. Both the saliency map and visual elements are allowed to be semi-automatically provided by the user, while the rest of the algorithm does not need to be modified. The following algorithm for enhancing image composition is designed under the assumption that visual element information has been successfully provided.

### 4.3. Notations

Denote  $\mathcal{I}$  as the image. The width and height of image  $\mathcal{I}$  are represented as  $W$  and  $H$ . The aspect ratio  $R$  could be computed as  $R = H/W$ .

The sets of detected salient objects and feature lines (e.g., the shuttle plane and horizon in Fig. 2(b)) are expressed as  $\mathcal{O}$  and  $\mathcal{L}$ , respectively. The sets of the four power points and six power lines are indicated as  $\mathcal{P}$  and  $\mathcal{Q}$ .

For each visual element  $\mathbf{e} \in \mathcal{O} \cup \mathcal{L}$ ,  $\sigma(\mathbf{e}) \in [0,1]$  denotes its saliency value (as in Fig. 2(a)).  $A(\mathbf{e})$  represents the area of the object if  $\mathbf{e} \in \mathcal{O}$ , and  $\bar{\sigma}(\mathbf{e}) = \sigma(\mathbf{e})A(\mathbf{e})$  represents its area-weighted saliency.  $G(\mathbf{e})$  represents the feature line length if  $\mathbf{e} \in \mathcal{L}$ , and  $\bar{\sigma}(\mathbf{e}) = \sigma(\mathbf{e})G(\mathbf{e})$  represents its length-weighted saliency.

$C(\mathbf{e}) = (C^x(\mathbf{e}), C^y(\mathbf{e}))$  denote the representative coordinates of elements in  $\mathcal{I}$ . For each salient object  $\mathbf{e} \in \mathcal{O}$ ,  $C(\mathbf{e})$  is its mass center in the image. For each line element  $\mathbf{e} \in \mathcal{L}$ ,  $C(\mathbf{e})$  are the coordinates of its midpoint in the image.

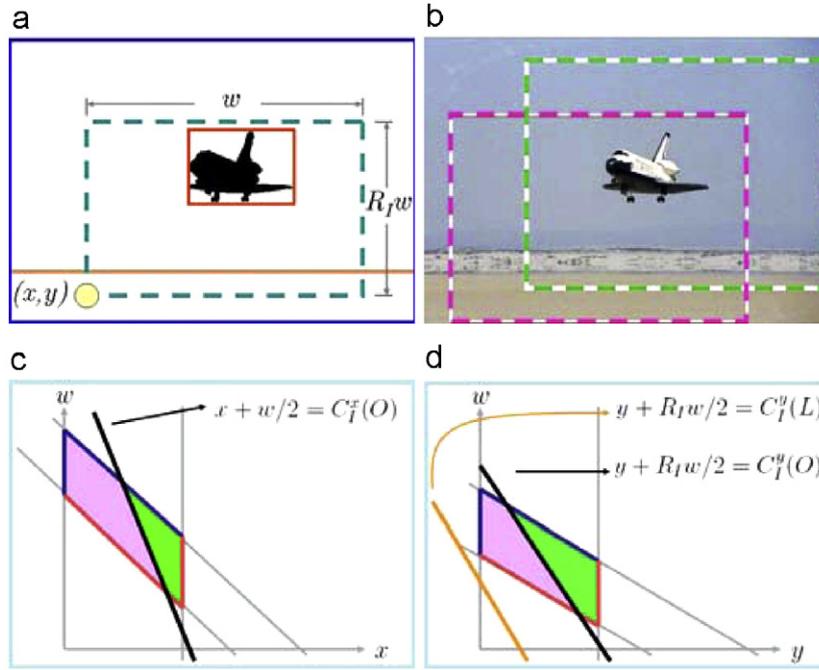
Denote  $\mathcal{I}_I$ ,  $\mathcal{I}_S$  and  $\mathcal{I}_U$  as the input image (Fig. 1(a)), optimally cropped sub-image (Fig. 1(b)) and output image (Fig. 1(c)), respectively. For identification purposes, we can add subscripts  $I, S, U$  to the defined symbols.  $W_I$  and  $H_I$  can stand for the width and height of the input image,  $\mathcal{P}_S$  and  $\mathcal{Q}_S$  for power point and power line sets in the cropped image  $\mathcal{I}_S$  and  $C_U(\mathbf{e})$  for the coordinates of element  $\mathbf{e}$  in the output image  $\mathcal{I}_U$ .

Sets of salient objects,  $\mathcal{O}$ , and sets of feature lines,  $\mathcal{L}$ , are generally identical in all three image forms. Without loss of generality, we do not add subscripts to them, including their saliency values, areas and lengths, all of which denote their counterparts in  $\mathcal{I}_I$ .

## 5. Optimal aesthetic crop

An image that contains obvious salient objects that are not well composed usually has abundant background spaces, as in Fig. 1(a). It is thus possible to find a sub-image with a better composition by discarding abundant background spaces.

Given input image  $\mathcal{I}_I$  with detected salient objects  $\mathcal{O}$ , feature lines  $\mathcal{L}$  and a saliency map, we attempt to find a sub-image  $\mathcal{I}_S$  whose composition optimally obeys the defined basic composition rules (Section 4.1). We establish the sub-window  $\mathcal{D}$  for cropping using three parameters: the coordinates of the left-bottom corner  $(x,y)$  and the width  $w$  of the sub-window, as in



**Fig. 3.** Optimal Aesthetic Crop. (a) The sub-window (cyan) uses three parameters: the coordinates of the left-bottom corner  $(x, y)$  (yellow) and the window width  $w$ . The sub-window must contain the bounding box  $B(\mathcal{O})$  (the red box) of the object set  $\mathcal{O}$ . (b) In the pink sub-window, i.e., the optimized solution, the space shuttle belongs to the right-top power point. In the green one, the space shuttle belongs to the left-bottom power point. (c) The active region, which is bounded by an image boundary (blue) and an object bounding box (red), is partitioned into two parts by  $x + w/2 = C_I^x(\mathcal{O})$  (black) and  $y + R_I w/2 = C_I^y(L)$  (orange). (d) The active region is also partitioned into 2 parts by  $y + R_I w/2 = C_I^y(\mathcal{O})$  (black) and  $x + w/2 = C_I^x(L)$  (orange). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

**Fig. 3(a).** We keep the aspect ratio of the sub-window consistent with the original image frame. The height of the sub-window could thus be computed as  $R_I w$ .

We adopt a strategy wherein salient objects cannot be even partly carved. Consequently, the sub-window for cropping the image can move freely and zoom homogeneously inside the original image frame only if the sub-window contains the bounding box  $B(\mathcal{O})$  of object set  $\mathcal{O}$  (the red box in Fig. 3(a)).

### 5.1. Aesthetic composition energy

Basic composition guidelines cannot play a direct role in numerical computation. We thus utilize quadratic functions to quantify the composition guidelines, as in previous work [8]. Though this section emphasizes analyzing the coordinate system of  $\mathcal{I}_S$ , the energy function is written without loss of generality for reuse in Section 6.

**Object term:** For each object  $\mathbf{o} \in \mathcal{O}$ , we denote  $\mathbf{P}_S(\mathbf{o}) \in \mathcal{P}_S$  as the power point closest to its mass coordinate  $C_S(\mathbf{o})$  in  $\mathcal{I}_S$ . We define the object term as the sum of the squared distances between  $C_S(\mathbf{o})$  and  $\mathbf{P}_S(\mathbf{o})$ :

$$E_{ob} = \frac{\sum_{\mathbf{o} \in \mathcal{O}} \bar{\sigma}(\mathbf{o}) \|D(C(\mathbf{o}), \mathbf{P}(\mathbf{o}))\|^2}{\sum_{\mathbf{o} \in \mathcal{O}} \bar{\sigma}(\mathbf{o})}, \quad (1)$$

where  $C(\mathbf{o}) = C_S(\mathbf{o}) = C_I(\mathbf{o}) - (x, y)$ ,  $\mathbf{P}(\mathbf{o}) = \mathbf{P}_S(\mathbf{o})$  and  $D(\cdot)$  indicates the Euclid distance between two points.

**Line term:** Consider a feature line  $\mathbf{l} \in \mathcal{L}$ . We denote its corresponding power line in  $\mathcal{I}_S$  as  $\mathbf{Q}_S(\mathbf{l}) \in \mathcal{Q}_S$ , which might be either a thirds line or a diagonal line. The line term is defined thus:

$$E_{li} = \frac{\sum_{\mathbf{l} \in \mathcal{L}} \bar{\sigma}(\mathbf{l}) \|\bar{D}(C(\mathbf{l}), \mathbf{Q}(\mathbf{l}))\|^2}{\sum_{\mathbf{l} \in \mathcal{L}} \bar{\sigma}(\mathbf{l})}, \quad (2)$$

where  $C(\mathbf{l}) = C_S(\mathbf{l}) = C_I(\mathbf{l}) - (x, y)$ ,  $\mathbf{Q}(\mathbf{l}) = \mathbf{Q}_S(\mathbf{l})$  and  $\bar{D}(\cdot)$  indicates the Euclid point-to-line distance.

**Visual balance term:** The visual balance guideline implies that the mass of all objects should be located around the image center

$C_S(\mathcal{I}) = (x + w/2, y + R_I w/2)$ . The visual balance term is defined thus:

$$E_{vb} = \left\| D\left(\frac{\sum_{\mathbf{o} \in \mathcal{O}} \sigma A(\mathbf{o}) C(\mathbf{o})}{\sum_{\mathbf{o} \in \mathcal{O}} \sigma A(\mathbf{o})}, C(\mathcal{I})\right) \right\|^2. \quad (3)$$

Here,  $C(\mathbf{o}) = C_S(\mathbf{o})$  and  $C(\mathcal{I}) = C_S(\mathcal{I})$ .

**Region size term:** For each object  $\mathbf{o} \in \mathcal{O}$ , we assume that its corresponding optimal region size fraction is  $\mathbf{F}_S(\mathbf{o}) \in \mathcal{F}$ , which is nearest to the fraction of area it actually captures in  $\mathcal{I}_S$  ( $f_S(\mathbf{o}) = A(\mathbf{o}) / (W_S H_S)$ ). Consequently, the optimal sub-image width corresponding to  $\mathbf{F}_S(\mathbf{o})$  could be computed as  $w_S(\mathbf{o}) = \sqrt{A(\mathbf{o}) / (R_I \mathbf{F}_S(\mathbf{o}))}$ . The region size term is defined as thus

$$E_{sz} = \frac{\sum_{\mathbf{o} \in \mathcal{O}} \bar{\sigma}(\mathbf{o}) \|w - w_S(\mathbf{o})\|^2}{\sum_{\mathbf{o} \in \mathcal{O}} \bar{\sigma}(\mathbf{o})}. \quad (4)$$

The total energy function for assessing the composition of the cropped image  $\mathcal{I}_S$  is defined as the weighted sum of the aesthetic composition energy terms:

$$E_{cr} = \omega_{ob} E_{ob} + \omega_{li} E_{li} + \omega_{vb} E_{vb} + \omega_{sz} E_{sz}, \quad (5)$$

where  $\omega_{ob}$ ,  $\omega_{li}$ ,  $\omega_{vb}$  and  $\omega_{sz}$  are weights.

The solution to minimize (5) defines the optimal position of sub-window  $\mathcal{D}$  in which visual elements best obey basic composition lines.

### 5.2. Space partition

The energy terms are not well determined when the sub-window  $\mathcal{D}$  moves and zooms within the input image frame, as the corresponding power points  $\mathbf{P}_S(\mathbf{o})$  in (1) and the power lines  $\mathbf{Q}_S(\mathbf{l})$  in (2), accompanied by the optimal sub-image width  $w(\mathbf{o})$  in (4), might indicate different targets in  $\mathcal{I}_S$ .

In Fig. 3(b), the space shuttle belongs to the right-top power point in the pink sub-window, which is the optimized solution. In the green one, the space shuttle belongs to the left-bottom power point.

To determine the target power point, consider each object  $\mathbf{o} \in \mathcal{O}$  with its coordinate  $C(\mathbf{o}) = (C_1^x(\mathbf{o}), C_1^y(\mathbf{o}))$  in  $\mathcal{I}_I$ . By defining the plane  $C_1^x(\mathbf{o}) - x = w/2$  that is parallel to the  $y$ -axis, we can partition the  $x-y-w$  space into two sub-spaces:

$$\Omega_R = \{\mathbb{R}^3, C_1^x(\mathbf{o}) - x \geq w/2\},$$

$$\Omega_L = \{\mathbb{R}^3, C_1^x(\mathbf{o}) - x < w/2\}.$$

$\mathbf{P}_S(\mathbf{o})$  in  $\Omega_R$  can be determined as the right power point in  $I_S$ , and  $\mathbf{P}_S(\mathbf{o})$  in  $\Omega_L$  can be determined as the left power point in  $I_S$ . Similarly, the  $x-y-w$  space can also be partitioned by the plane  $C_1^y(\mathbf{o}) - y = R_I w/2$ .

For horizontal or vertical lines, we can also define space partitions as for objects. We only need to judge the left or right side for a vertical line and the up or down side for a horizontal line. Moreover, we do not need to define space partitions for diagonal lines, as their corresponding power lines can be pre-determined by their slopes.

Analogously, to determine the corresponding optimal sub-image width  $w_i(\mathbf{o})$  for each object  $\mathbf{o} \in \mathcal{O}$ , we compute the two median value of the optimal widths:  $w_{12}(\mathbf{o}) = \sqrt{A(\mathbf{o})/2(R_I(F_1 + F_2))}$  and  $w_{23}(\mathbf{o}) = \sqrt{A(\mathbf{o})/2(R_I(F_2 + F_3))}$ .  $w(\mathbf{o})$  can thus be determined in the sub-spaces partitioned by the plane  $w = w_{12}(\mathbf{o})$  and  $w = w_{23}(\mathbf{o})$ .

We sort the  $x$ -coordinates of objects and vertical feature lines as  $\mathcal{X} = \{x_0 = -\infty, x_1, \dots, x_l, x_{l+1} = +\infty, x_i \leq x_{i+1}\}$ , where  $l$  is the total number of objects and vertical lines. Likewise, the  $y$ -coordinates of objects and horizontal feature lines are sorted as  $\mathcal{Y} = \{y_0 = -\infty, y_1, \dots, y_m, y_{m+1} = +\infty, y_i \leq y_{i+1}\}$ , and the median values  $w_{12}, w_{23}$  of optimal widths are sorted as  $\mathcal{W} = \{w_0 = -\infty, w_1, \dots, w_{2n}, w_{2n+1} = +\infty, w_i \leq w_{i+1}\}$ , where  $m$  indicates the total number of objects and horizontal lines and  $n$  denotes the number of objects. The  $x-y-w$  space can be partitioned thus:

$$\Omega_{i,j,k} = \{\mathbb{R}^3 \mid x_i \leq x + w/2 < x_{i+1}, y_j \leq y + R_I w/2 < y_{j+1}, w_k \leq w < w_{k+1}\}$$

$$i = 0, \dots, l, j = 0, \dots, m, k = 0, \dots, 2n. \quad (6)$$

In any region  $\Omega_{i,j,k}$ , the corresponding power points  $\mathbf{P}_S(\mathbf{o})$ , power lines  $\mathbf{Q}_S(\mathbf{l})$  and optimal sub-image width  $w(\mathbf{o})$  are all well determined.

The sub-window  $\mathcal{D}(x,y,w)$  is only allowed to move and zoom inside the original image  $I_I$ . It must contain the object bounding box  $B$ . The active region for computing the sub-window position  $\mathcal{D}$  can be expressed as follows:

$$\Omega_{ac} = \{\mathbb{R}^3, B(\mathcal{O}) \subseteq \mathcal{D}(x,y,w) \subseteq \mathcal{I}_I\}. \quad (7)$$

Fig. 3(c,d) displays the space partition considering the object coordinates  $(C_1^x(\mathcal{O}), C_1^y(\mathcal{O}))$  and the  $y$ -coordinates  $C_1^y(L)$  of the feature line. For visualization purposes, we show the  $x-w$  and  $x-y$  plane side views. In Fig. 3(c), the active region, which is bounded by the image boundary (blue) and object bounding box (red), is partitioned into two parts by  $x + w/2 = C_1^x(\mathcal{O})$  (black). The pink cropping window in Fig. 3(b) belongs to the pink region in Fig. 3(c), while the green one belongs to its green counterpart. The space is partitioned into three parts, and the active region is also partitioned into two parts by  $y + R_I w/2 = C_1^y(\mathcal{O})$  (black) and

$y + R_I w/2 = C_1^y(L)$  (orange) in Fig. 3(d). The partitioned regions can be further partitioned by the median value of optimal widths  $w_{12}(\mathcal{O}), w_{23}(\mathcal{O})$ , which are not shown in the figures to maintain the clarity of the figures.

### 5.3. Implementation

Because all Euclid distances included are linear combinations of the parameters  $x,y,w$  equipped with the space partition (6), the total energy (5) can be considered a piecewise quadratic function. Both region constraints (6) and (7) are linear inequality constraints. The solution to minimize (5) in the active region  $\Omega_{ac}$ , which defines the optimal position of sub-window  $\mathcal{D}$ , can be easily obtained by solving a piecewise quadratic program.

The size of the optimally cropped image  $\mathcal{I}_S$  may be smaller than that of the input image  $\mathcal{I}_I$ ; however, their aspect ratios are the same. We can homogeneously magnify  $\mathcal{I}_S$  to fit the size of  $\mathcal{I}_I$ . Consequently,  $\mathcal{I}_S$  can be considered a composition-enhanced version of  $\mathcal{I}_I$ , as in Fig. 1(a,b).

## 6. Optimal aesthetic warp

The composition of the cropped image  $\mathcal{I}_S$  is usually an obviously enhanced contrast to the input image  $\mathcal{I}_I$ . However, the cropping window  $\mathcal{D}$  cannot move out of the original frame during the crop operator and may miss the optimal solution to the composition energy (5) without active region constraints (7). Moreover, tradeoffs may exist in minimizing (5) if multiple objects and lines crowd into one frame.

We employ a mesh warping technique to actively push visual elements towards their target power lines or power points, as in a previous study [8]. The mesh vertices are sampled uniformly from  $\mathcal{I}_S$ , cooperating with the priorities of boundary pixels, Canny edges [30] and pixels contained by feature lines. The mesh is constructed using a constrained Delaunay triangulation algorithm [31], which asserts that feature lines should be part of the mesh edges, as in Fig. 4(a,b).

Fig. 4 illustrates the workflow of the aesthetic warp operator. Blue mesh triangles correspond to triangles in salient objects  $\mathcal{O}$ . Orange mesh edges correspond to feature lines  $\mathcal{L}$ .

We denote the constructed mesh from  $\mathcal{I}_S$  as  $\mathcal{M}_S$ . Our goal is to compute the vertex coordinates of the output mesh  $\mathcal{M}_U$ , which has the same topology as  $\mathcal{M}_S$ , and compute the final enhanced image composition  $\mathcal{I}_U$ . This can be formulated as a problem of warping the input mesh  $\mathcal{M}_S$  to the output mesh  $\mathcal{M}_U$ . The target image  $\mathcal{I}_U$  can thus be obtained by texture mapping between the corresponding triangles in  $\mathcal{M}_S$  and  $\mathcal{M}_U$ , as in Fig. 4(c,d). To achieve warping, we employ a quadratic optimization framework that considers several components of energies and is described below. The width and height of  $\mathcal{I}_U$  do not need to be the same as their counterparts in  $\mathcal{I}_S$ , as in Fig. 4(e,f). We denote the scaling factor of the image frame as  $s^x = W_U/W_S, s^y = H_U/H_S$ .

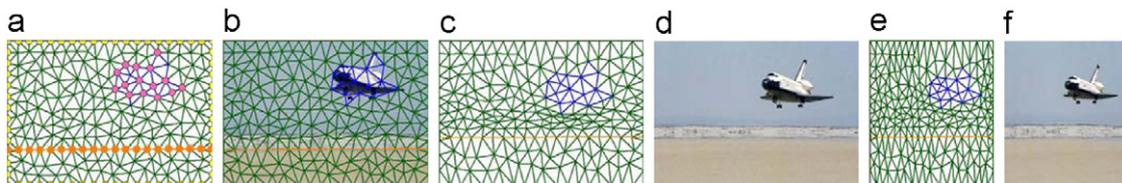


Fig. 4. Workflow of an optimal aesthetic warp. (a) A triangular mesh is constructed by constrained Delaunay triangulation from  $\mathcal{I}_S$ . Some mesh vertices are sampled from Canny edges (pink), feature lines (orange) and image boundaries (yellow). The blue triangles denote the salient object in the image. (b) Mesh over image. (c, d) Optimally warped mesh and the final enhanced image composition. (e, f) Optimally warped mesh by 60% width and the corresponding image result. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

### 6.1. Warp energy

**Composition term:** We attempt to actively push visual elements in  $\mathcal{I}_S$  towards their power points or lines. We set  $\mathbf{C}(\mathbf{e}) = C_U(\mathbf{e})$ ,  $\mathbf{P}(\mathbf{o}) = \mathbf{P}_U(\mathbf{o})$ ,  $\mathbf{Q}(\mathbf{l}) = \mathbf{Q}_U(\mathbf{l})$ , and  $C(\mathcal{I}) = C_U(\mathcal{I})$  in the Object (1), Line (2) and Visual balance terms (3). The composition term is defined as follows:

$$E_{co} = \omega_{ob}E_{ob} + \omega_{li}E_{li} + \omega_{vb}E_{vb}. \quad (8)$$

Differing from Section 5, the visual element coordinates  $C_U(\mathbf{e})$  can be represented as linear combinations of vertex coordinates in the target mesh  $\mathcal{M}_U$ . Moreover, target power points and power lines  $\mathbf{P}_U(\mathbf{o})$  and  $\mathbf{Q}_U(\mathbf{l})$  are well determined using  $\mathcal{I}_S$ .

We add two mesh deformation terms, i.e., scaling and smoothness, to control the mesh warp operator behavior, as in a previous study [8]:

**Scaling term:** Denote the set of all mesh triangles as  $\mathcal{T}$ . We assign each triangle  $t$  an auxiliary linear transformation matrix:

$$G_t = \begin{pmatrix} s_t^x & 0 \\ 0 & s_t^y \end{pmatrix}$$

Here,  $s_t^x$  is the x-scale factor, and  $s_t^y$  is the y-scale factor.

Conversely, the rotation and scaling portions  $J_t$  of the actual affine mapping of triangle  $t$  from mesh  $\mathcal{M}_S$  to its counterpart in mesh  $\mathcal{M}_U$ , which is a  $2 \times 2$  Jacobian matrix, linearly depend on the vertex coordinates of (triangles in)  $\mathcal{M}_U$ . We define the scaling term thus

$$E_{sc} = \sum_{t \in \mathcal{T}} \bar{\sigma}(t) \|J_t - G_t\|_F^2. \quad (9)$$

Here  $\bar{\sigma}(t)$  represents the area-weighted saliency of triangle  $t$ , and  $\|\cdot\|_F$  is the Frobenius norm.

Minimizing the scaling term  $E_{sc}$  makes the rotation and scaling portion  $J_t$  of the affine mapping be as close as possible to the allowed scaling transformation  $G_t$ .

**Smoothness term:** To prevent obvious discontinuities or distortions in the result image, we suggest that the scaling transformations applied within a region of the mesh be as similar to each other as possible. Define the following smoothness term:

$$E_{sm} = \sum_{s,t \in \mathcal{T}; s,t \text{ are adjacent}} \bar{\sigma}(s,t) \|G_t - G_s\|_F^2, \quad (10)$$

where  $\bar{\sigma}(s,t) = (\bar{\sigma}(s) + \bar{\sigma}(t))/2$ .

The total energy defined for the aesthetic warp operator is defined thus:

$$\mathbf{E}_{wp} = E_{co} + \lambda E_{sc} + \mu E_{sm}, \quad (11)$$

where  $\lambda$  and  $\mu$  are weights.

### 6.2. Constraints

We must add essential constraints to perform the optimization.

**Boundary constraints:** Boundary mesh vertex coordinates should fit the rectangular boundary of output image  $\mathcal{I}_U$ . For each vertex  $v$  on the left side of the image frame, we impose the positional constraint  $v^x = 0$ . For each triangle  $t$  that contains an edge on the left side of the image frame, we impose the scaling constraint  $s_t^y = s^y$ . The other boundary constraint sets are similar.

**Feature line constraints:** The feature lines should be preserved in a straight line in the output mesh. We predict the normal  $\mathbf{n}(\mathbf{l})$  for feature line  $\mathbf{l} \in \mathcal{L}$  based on the aspect ratio of the output image frames. For each mesh edge  $g \in \mathbf{l}, \mathbf{l} \in \mathcal{L}$ , we impose the following constraints to preserve the straightness of  $\mathbf{l}$ :

$$g \cdot \mathbf{n}(\mathbf{l}) = 0 \quad \forall g \in \mathbf{l} \quad \forall \mathbf{l} \in \mathcal{L}. \quad (12)$$

**Salient object constraints:** There are two requirements for imposing salient object constraints. One is that triangles contained by salient objects should better preserve their aspect ratio, i.e., they ought to be scaled homogeneously.

The other is that the framework should also be able to zoom into object  $\mathbf{o} \in \mathcal{O}$  to fit the optimal region size fraction  $F_U(\mathbf{o}) \in \mathcal{F}$  that is closest to the fraction of the area that the object actually captured in  $\mathcal{I}_U$ :  $f_U(\mathbf{o}) = A(\mathbf{o})/(H_U W_U)$ . Although the optimal zoom factor for each object  $\mathbf{o}$  should be  $s(\mathbf{o}) = \sqrt{F(\mathbf{o})/f(\mathbf{o})}$ , objects should be considered integral when zooming to preserve their relative proportions. We define the optimal zoom factor as the saliency-weighted average

$$\alpha = \frac{\sum_{\mathbf{o} \in \mathcal{O}} \bar{\sigma}(\mathbf{o}) s(\mathbf{o})}{\sum_{\mathbf{o} \in \mathcal{O}} \bar{\sigma}(\mathbf{o})}.$$

Moreover, we do not allow objects to significantly zoom out, as doing so might cause mesh edge intersections. We thus define the actual object zoom factor using the image frame scaling factor as follows:

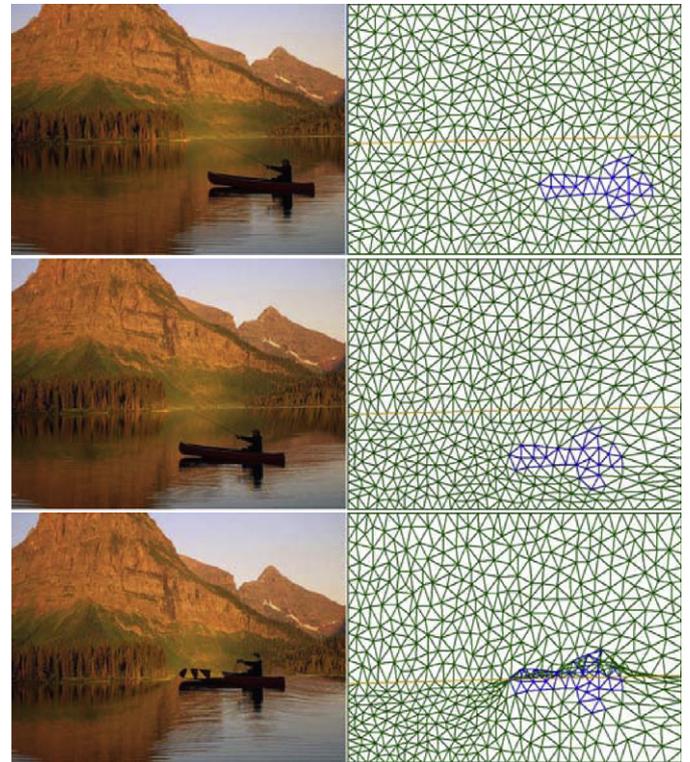
$$\beta = \min(\alpha, \sqrt{s^x s^y}).$$

The salient object constraints are defined as follows:

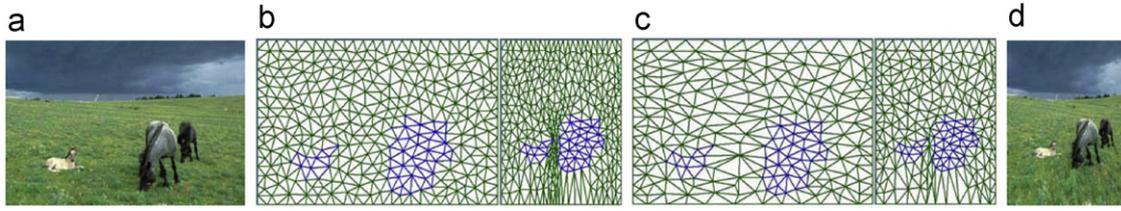
$$G_t = \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} \quad \forall t \in \mathbf{o} \quad \forall \mathbf{o} \in \mathcal{O} \quad (13)$$

**Relative position constraints:** The composition term (8) attempts to push visual elements towards their target power points or power lines. However, if two objects belong to the same target power point or two lines belong to the same target power line, a rough push causes mesh self-intersection. Moreover, objects may also conflict with lines, as in the bottom row of Fig. 5.

We therefore try to preserve their relative positions to avoid such defects. The relative position constraints on both objects and



**Fig. 5.** Relative positions of visual elements should be preserved to prevent mesh self-intersection. Upper row: Image  $\mathcal{I}_S$  and its mesh  $\mathcal{M}_S$ ; middle row: composition enhanced image  $\mathcal{I}_U$  and warped mesh  $\mathcal{M}_U$ ; bottom row: rough warp [8] causes severe mesh self-intersection.



**Fig. 6.** Adaptive mesh simplification. (a) Image  $\mathcal{I}_S$ ; (b) Meshes  $\mathcal{M}_S$  and  $\mathcal{M}_U$ . Meshes between the two objects are heavily crowded. (c) Simplified meshes  $\overline{\mathcal{M}}_S$  and  $\overline{\mathcal{M}}_U$ ; (d) Output image  $\mathcal{I}_U$ .

lines are defined thus

$$V_{\perp}(C_U(e_1) - C_U(e_2)) = V_{\perp}G(C_S(e_1) - C_S(e_2)), \quad \forall e_1, e_2 \in \mathcal{O} \cup \mathcal{L}; e_1, e_2 \text{ are conflicted}, \quad (14)$$

where  $V_{\perp}(\cdot)$  denotes the vector if both  $e_1$  and  $e_2$  are objects or a component of the vector that is perpendicular to the target power line if one or both of  $e_1$  and  $e_2$  are feature lines.

$$G = \begin{pmatrix} s^x & 0 \\ 0 & s^y \end{pmatrix}$$

is the scaling matrix of the image frame.

### 6.3. Implementation

The energy function (11) is quadratic with respect to the vertex coordinates of  $\mathcal{M}_U$  and scale factors  $\{s_t^x, s_t^y \mid t \in \mathcal{T}\}$  for all triangles. The boundary and salient object constraints (13) can be considered hard constraints, while the feature line (12) and relative position constraints (14) are heavily weighted soft constraints for the minimization problem. Minimizing (11) can be achieved by solving a sparse linear system.

Although we solve for both the mesh vertex coordinates of  $\mathcal{M}_U$  and the scale factors for all triangles, we are only interested in the mesh vertex coordinates; the scale factors play an auxiliary role.

### 6.4. Adaptive mesh simplification

Because we allow users to modify the target image frame size, the frame size might be reduced in one dimension significantly, which may concentrate mesh edges in one region, as in Fig. 6(b). This causes mesh self-intersections and numerical instability when computing  $\mathcal{I}_S$  from  $\mathcal{M}_U$ .

Consequently, we provide an optional procedure to enhance the mesh quality of mesh  $\mathcal{M}_U$ . After  $\mathcal{M}_S$  warps to  $\mathcal{M}_U$ , we can simplify the topology of the two meshes simultaneously for a defined edge threshold in  $\mathcal{M}_U$  by employing the mesh edge collapse algorithm [32] to obtain the simplified mesh  $\overline{\mathcal{M}}_S$ . The vertex produced by each collapsed edge is set as the edge midpoint. We then utilize  $\overline{\mathcal{M}}_S$  to perform the warp operator and produce the final result image  $\mathcal{I}_U$ , as in Fig. 6(c, d).

The opportunity to cause self-intersection is significantly reduced using the adaptively simplified mesh  $\overline{\mathcal{M}}_S$ . Moreover, it is more numerically stable when computing the warped image  $\mathcal{I}_U$  from  $\overline{\mathcal{M}}_S$  than computing it from  $\mathcal{M}_S$ .

## 7. Experimental results and discussions

All experimental results presented in this paper were made on a PC with a Duo-Core 1.8 GHz CPU and 2 GB memory. The piecewise quadratic program contains three parameters; it can thus be instantly solved using arbitrary convex optimization algorithms, such as interior-point methods [33]. We also employ Intel MKL [34] to solve the involved sparse linear system, which is sensitive to mesh size.

The average edge length in a triangular mesh is within 15–35 pixels. The main computational cost is solving the sparse linear system. It takes about 50–150 ms to optimize the composition when choosing a common mesh size for an image with a resolution around  $1024 \times 768$ .

Our algorithm contains a few energy weights. In our experience, the whole algorithm is robust over a large range of all parameters. We set  $\omega_{ob} = 0.1$ ,  $\omega_{li} = 0.1$ ,  $\omega_{vb} = 0.03$ ,  $\omega_{sz} = 0.01$  in (5), (8) and  $\lambda = 1.0$ ,  $\mu = 0.5$  in (11) for their magnitudes and importance.

Fig. 7 shows retargeting images to different frame sizes while the image composition has been enhanced simultaneously during the retargeting. Several images selected from the RetargetMe database [35] have been used to assess image retargeting algorithms. Within the selected examples, our algorithm is especially effective for images comprising obvious visual elements with adequate background space and can intrinsically find an enhanced composition version of the image. We have also prepared supplemental material to show more representative images that fit our algorithm, which can be found on our website.

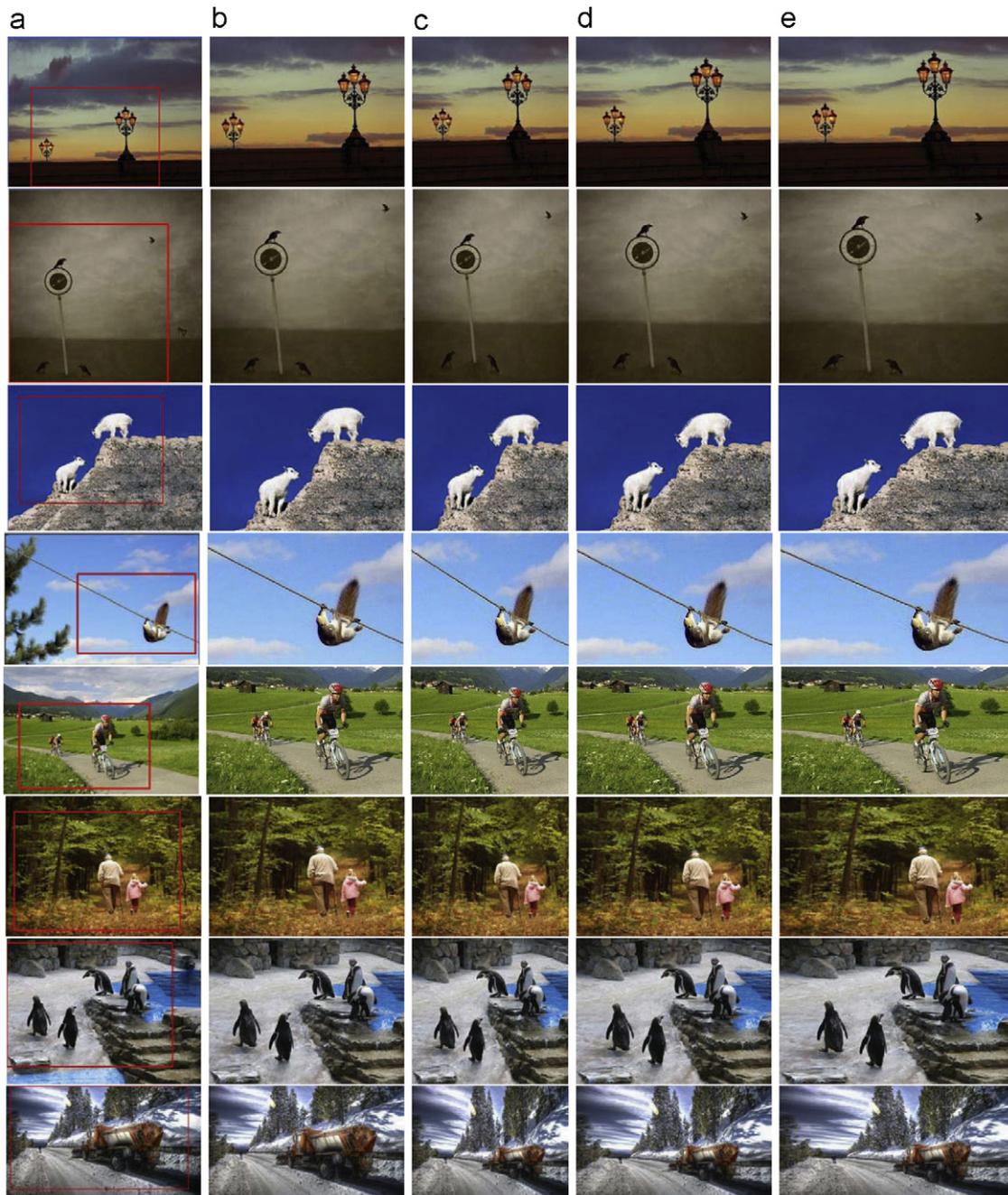
We compare our method and two state-of-the-art image composition enhancement techniques with each other: the crop-and-retarget [7] and warp methods [8] in Fig. 8. The three algorithms perform similarly, including in the first row.

In the second row, the salient objects in the image, including the sun, boat and girl, are well composed both in our result and those of Liu et al. [7]. Moreover, the horizon in our result is closer to its corresponding thirds line, as the warp operator has further optimally adjusted the composition of our result. Because it lacks a crop operator, the composition of the result of Liu et al. [8] seems odd, and an obvious trace exists, caused by warping the sun to the left-bottom power point.

It is similar to the second row in the example of the third row: the scene showing more of the lake in both our results and those of Liu et al. [7] seems more satisfactory than the scene with more sky. The lake is more essential for emphasizing the duck, which is the unique subject of the image.

The compositions of the three results are fairly good in the last three rows. However, the salient object sizes (e.g., boat, building and wood) in the results of Liu et al. [8], without considering the region size, seem small. The other two results do not contain such defects.

We also conducted a user study to compare the results produced in previous studies [7,8] as well as those produced using our algorithm. We invited students in our research group to attend the user study, as all students in our group had been trained in computer graphics. They possessed basic knowledge about digital image processing algorithms. In our user study questionnaire, we showed training images to clarify the composition rules, including the rule of thirds, visual balance and initial region size. The six rows of images in Fig. 8 were then shown to the participants, and the algorithms were kept anonymous. The participants were requested to select their favorite and least favorite images. The most favorite image gained 2 points for its corresponding algorithm, while the least favorite one gained



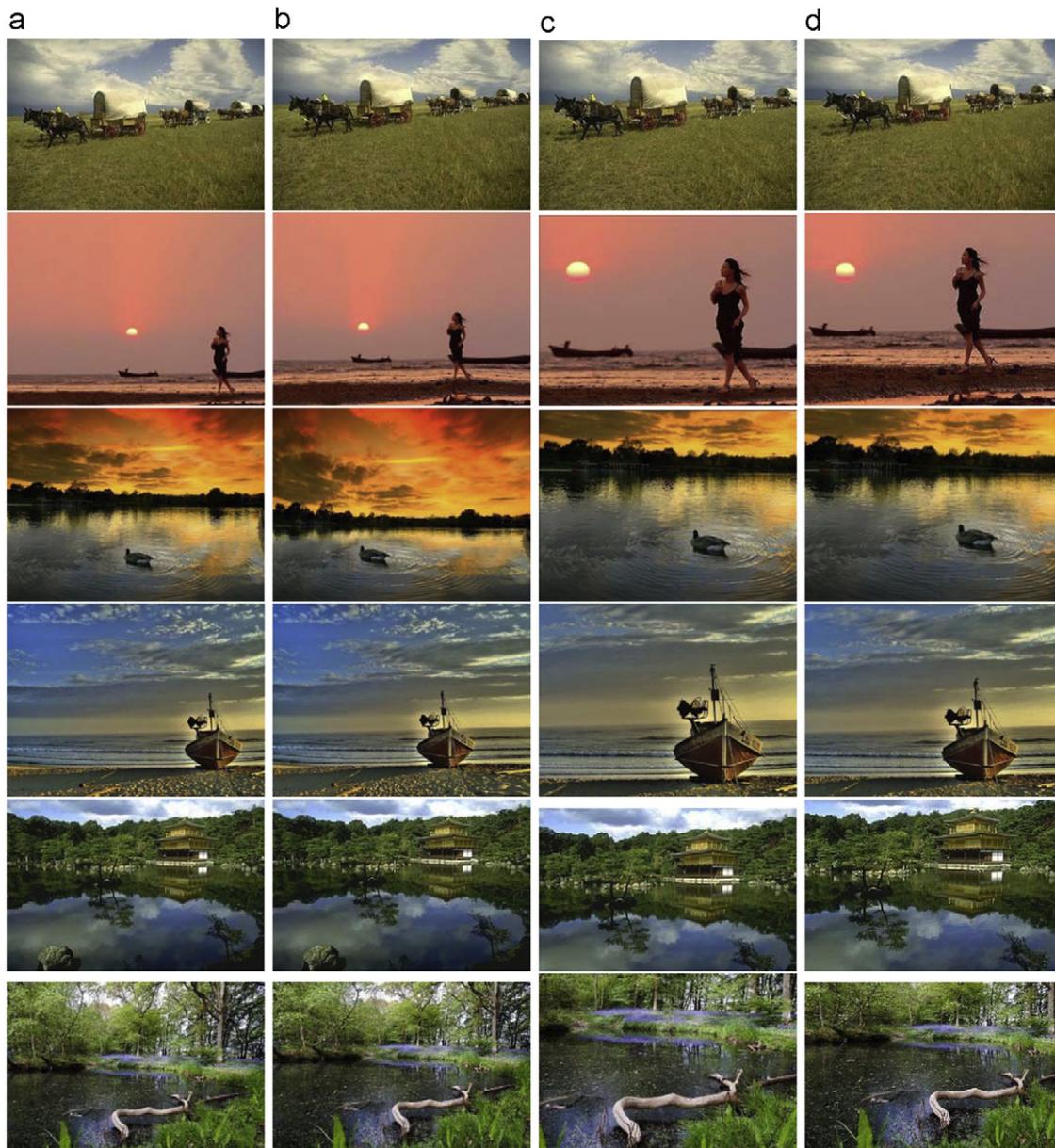
**Fig. 7.** Retargeting the images into different frame sizes while enhancing the composition simultaneously. (a) The input image and the optimal sub-window for cropping; (b) the cropped image; the image is retargeted to 80% (c), 100% (d), and 120% (e), respectively, of the original frame width.

0 points and the remaining one gained 1 point. We received 15 responses. The algorithm [8] received 19% of the total points, the algorithm [7] received 46% points, and our algorithm received 35% points. The algorithm presented by Liu et al. [7] and our algorithm obtained similar vote percentages. To check their significant difference, the votes of the six test images were accumulated for each participant. We then performed a one-way ANOVA analysis [36] on the accumulated points using Matlab [37]. Fig. 9 presents the results. Although the average value of our algorithm is lower than that of Liu et al. [7] by approximately 1 point, the  $F$ - and  $p$ -values of the ANOVA analysis are 4.15 and 0.0513, respectively, illustrating that the statistical differences between the algorithms are not significant. The user study shows that our method obtains results comparable to one previous algorithm [7] and slightly better than another [8].

Though our approach and the previous method of [7] performed remarkably, our approach is much more efficient (Table 1). Our method takes approximately 100 ms to enhance the composition of one image, while the previous method [7] claims that it takes 2–14 s to optimize the composition of a photo with a resolution of  $1024 \times 768$ .

Table 1 shows the image size, mesh size and timing statistics for most examples in this paper to illustrate the performance of our algorithm. The running time depends on the mesh resolution adopted by the warp operator and the number of visual elements, while it is independent of the image resolution. Our approach can also handle large-size images.

**Limitation:** Our algorithm especially fits images with obvious focuses or visual elements, as it is easier to compose visual elements using the aesthetic guidelines in such images.



**Fig. 8.** Comparison with two state-of-the-art image composition enhancement techniques. (a) Input Image; (b) result of the warping method [8]; (c) crop-and-retarget method results [7]; (d) results of our approach.

Consequently, our algorithm cannot work on images with multiple visual elements and limited background space, as in Fig. 10(a). The cropped image is usually the original image itself; hence, the composition is hardly enhanced.

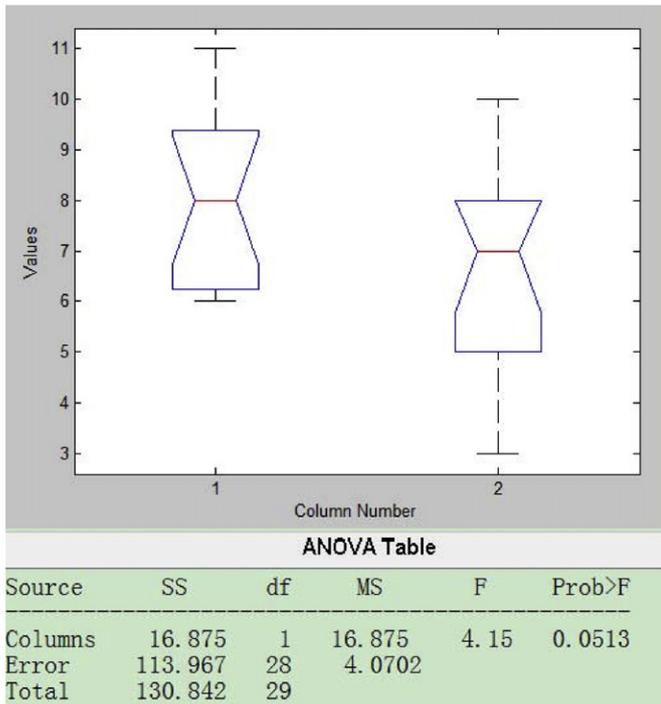
Moreover, although the improved warp operator significantly reduces the opportunity to cause self-intersections and severe distortions, obvious distortion may still exist in images with strongly identifiable background spaces that contain a long-distance warp operator in the result image, as in Fig. 10(b). For such cases, we suggest that the crop operator be omitted.

## 8. Conclusions and future work

We have presented an improved approach to enhance image composition using a crop-and-warp-based algorithm. A sub-image

in which visual elements optimally coincide with basic composition guidelines is cropped. The locations of visual elements in the cropped image are then further optimally adjusted by adopting a triangular mesh-based warping technique. The algorithm is performed by solving a piecewise quadratic program and a sparse linear system in succession. Several experimental results have shown the effectiveness and efficiency of our approach, which can be integrated into image processing tools.

It would be interesting to consider more guidelines for producing aesthetic images, such as colors or tones, in future work. Moreover, it might be possible to better judge the optimal positions of visual elements and the image aspect ratios using eye-tracking-based techniques [27,28]. For images with multiple visual elements, i.e., Fig. 10(a), it might be possible to inpaint the background space and perform the crop operator to obtain a well composed image using the proposed composition rules.



**Fig. 9.** One-way ANOVA analysis on accumulated points for a previous algorithm [7] and for ours. Column 1 represents Liu et al. [7], and column 2 represents our algorithm.

**Table 1**  
Timing statistics.

Fig.	Image size	Mesh size	Time (ms)
4	544 × 382	155 V/258 T	47
5	640 × 480	594 V/1078 T	125
7 (1st. Row)	472 × 472	313 V/548 T	109
7 (2nd. Row)	800 × 600	205 V/347 T	94
7 (3rd. Row)	800 × 600	251 V/431 T	94
7 (4th. Row)	472 × 318	124 V/200 T	47
7 (5th. Row)	460 × 300	220 V/372 T	52
7 (6th. Row)	600 × 429	187 V/315 T	31
7 (7th. Row)	615 × 461	202 V/240 T	62
7 (8th. Row)	1024 × 594	452 V/802 T	110



**Fig. 10.** Two kinds of failure cases. (a) An image (left) with multiple visual elements with limited background space. The optimal cropped image (right) is the original image; (b) an image (left) with a strongly identifiable background. The warp operator (right) may cause noticeable distortion.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (10871178, 61070071), the Important National Science & Technology Specific Projects (2009ZX07424-001), and the 973 Program (2009CB320801, 2011JB105001).

## Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version, at <http://dx.doi.org/10.1016/j.cag.2012.07.007>.

## References

- [1] Hoenig F. Defining computational aesthetics. In: Gooch B, Neumann L, Sbert M, Purgathofer W, editors. International symposium on computational aesthetics in graphics, visualization and imaging. Girona: Eurographics Association; 2005. p. 13–8.
- [2] Martinez B, Block J. Visual forces, an introduction to design. New York: Prentice-Hall; 1998.
- [3] Peters G. Aesthetic primitives of images for visualization. In: International conference on information visualization. Zurich: IEEE Computer Society; 2007. p. 316–25.
- [4] Rivotti V, Proenaa JJ, Jorge J, Sousa M. Composition principles for quality depiction and aesthetics. In: Cunningham DW, Meyer G, Neumann L, Dunning A, Paricio R, editors. International symposium on computational aesthetics in graphics, visualization and imaging. Banff: Eurographics Association; 2007. p. 37–44.
- [5] Grill T, Scanlon M. Photographic composition. Watson-Guptill; 1990.
- [6] Krages B. Photography: the art of composition. Allworth Press; 2005.
- [7] Liu L, Chen R, Wolf L, Cohen-Or D. Optimizing photo composition. Comput Graph Forum 2010;29:469–78.
- [8] Liu L, Jin Y, Wu Q. Realtime aesthetic image retargeting. In: Deussen O, Jepp P, editors. International symposium on computational aesthetics in graphics, visualization and imaging. London: Eurographics Association; 2010. p. 1–8.
- [9] Suh B, Ling H, Bederson B, Jaobs D. Automatic thumbnail cropping and its effectiveness. In: ACM conference on user interface and software technology. ACM Press; 2003. p. 95–104.
- [10] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Pattern Anal Mach Intell 1998;20:1254–9.
- [11] Gooch B, Reinhard E, Moulding C, Shirley P. Artistic composition for image creation. In: Eurographics workshop on rendering technique. Eurographics Association; 2001. p. 83–8.
- [12] Byers Z, Dixon M, Smart WD, Grimm C. Say cheese! experiences with a robot photographer. AI Mag 2004;25:37–46.
- [13] Lok S, Feiner S, Ngai G. Evaluation of visual balance for automated layout. In: International conference on intelligent user interfaces. New York: ACM Press; 2004. p. 101–8.
- [14] Santella A, Agrawala M, Decarlo D, Salesin DH, Cohen MF. Gaze-based interaction for semiautomatic photo cropping. In: ACM human factors in computing systems (CHI). ACM Press; 2006. p. 771–80.
- [15] Avidan S, Shamir A. Seam carving for content-aware image resizing. ACM Trans Graph 2007;26:267–76.
- [16] Rubinstein M, Shamir A, Avidan S. Improved seam carving for video retargeting. ACM Trans Graph 2008;27:1–9.
- [17] Rubinstein M, Shamir A, Avidan S. Multi-operator media retargeting. ACM Trans Graph 2009;28:1–11.
- [18] Setlur V, Takagi S, Raskar R, Gleicher M, Gooch B. Automatic image retargeting. In: Proc. of mobile and ubiquitous multimedia. ACM Press; 2005. p. 59–68.
- [19] Karni Z, Freedman D, Gotsman C. Energy-based content-aware image deformation. Comput Graph Forum 2009;28:1257–68.
- [20] Wang YS, Tai CL, Sorkine O, Lee TY. Optimized scale-and-stretch for image resizing. ACM Trans Graph 2008;27:1–8.
- [21] Wolf L, Guttman M, Cohen-Or D. Nonhomogeneous content-driven video-retargeting. In: Proceedings of the ICCV. IEEE; 2007. p. 1–6.
- [22] Guo Y, Liu F, Shi J, Zhou Z, Gleicher M. Retargeting using mesh parametrization. IEEE Trans Multimed 2009;11:856–67.
- [23] Jin Y, Liu L, Wu Q. Nonhomogeneous scaling optimization for realtime image resizing. Visual Comput 2010;26:769–78.
- [24] Shamir A, Sorkine O. Visual media retargeting. In: SIGGRAPH Asia course. ACM Press; 2009.
- [25] Harel J, Koch C, Perona P. Graph-based visual saliency. In: Proceedings of the neural information processing systems (NIPS), vol. 19. MIT Press; 2006. p. 545–52.
- [26] Fernandes L, Oliveira M. Real-time line detection through improved hough transform voting scheme. Pattern Recogn 2008;41:299–314.
- [27] Judd T, Ehinger K, Durand F, Torralba A. Learning to predict where humans look. In: IEEE International conference on computer vision (ICCV). IEEE; 2009.

- [28] Castillo S, Judd T, Gutierrez D. Using eye-tracking to assess different image retargeting methods. In: Symposium on applied perception in graphics and visualization (APGV). ACM Press; 2011.
- [29] Cheng M-M, Zhang G-X, Mitra NJ, Huang X, Hu S-M. Global contrast based salient region detection. In: IEEE CVPR. IEEE; 2011. p. 409–16.
- [30] Canny J. A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 1986;8:678–98.
- [31] Seidel R. Constrained Delaunay triangulations and Voronoi diagrams with obstacles, Technical report, Institute for Information Processing, Graz, Austria; 1988.
- [32] Garland M, Heckbert PS. Surface simplification using quadric error metrics In: Proceedings of ACM SIGGRAPH; p. 209–16.
- [33] Boyd S, Vandenberghe L. Convex optimization. Cambridge University Press; 2004.
- [34] Intel math kernel library, <<http://software.intel.com/en-us/articles/intel-mkl/>>; 2011.
- [35] Rubinstein M, Gutierrez D, Sorkine O, Shamir A. A comparative study of image retargeting. ACM Trans Graphics (Proc SIGGRAPH Asia) 2010; 29:160:1–10.
- [36] Hogg RV, Ledolter J. Engineering Statistics. New York: MacMillan; 1987.
- [37] Matlab, <<http://www.mathworks.com>>; 2011.