# High quality image resizing

Qi Wang [a], Yuan Yuan [b],*

[a] Northwestern Polytechnical University, Xi′an 710072, Shaanxi, PR China.
[b] Center for OPTical IMagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi′an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi′an 710119, Shaanxi, PR China

A B S T R A C T

An increasing amount of display devices with fixed sizes call for an adaptive strategy for optimal display. For this purpose, content aware image resizing techniques are developed. Previous works mainly lay their attention on the shrinkage operation of the examined image. Less efforts are paid on the expansion manipulation. Though some literatures claim an extension of their shrinkage operation to expanding the image in a similar way, the obtained results are not satisfying. In this paper, a high quality image resizing method is proposed to retain the details when stretching an image. Instead of using interpolation based techniques which are taken for granted by existing methods, an expansion model is first learned from a set of training images. Then the future enlargement is based on this principle. Experiments on two publicly available datasets demonstrate the effectiveness of the presented method. A further extension on video enlargement is also presented as an example. Though the proposed method is formulated in the context of seam carving, it can be readily extended to other techniques such as cropping, segmentation and warping based resizing methods.

© 2014 Published by Elsevier B.V.

## 1. Introduction

With the development of imaging technologies, an increasing amount of display devices are emerging up. For instance, PDA, cell phone, tablet PC, and laptop are the most familiar facilities in our daily life and each of them has a function of presenting visual media for users. Normally, images have a fixed resolution or are authorized once. But they need to be consumed in different circumstances, especially on various display devices. If the image aspect ratio differs from that of the available display area, the image should be adjusted to adapt to the layout. This process of changing the image size for optimal display is generally named as *image resizing* or *image retargetting* [1,2].

A common method for resizing is to scale the image ununiformly according to the desired aspect ratio. This is a simple and straightforward operation but the resulting visual effect might not be satisfying. The reason behind the consequence is mainly that most of the time, the inequivalent aspect ratios before and after resizing will lead to content distortion, which is unacceptable by most people. To tackle the problem, *content-aware* techniques are developed to maintain the important or salient objects undistorted. These objects are mostly inferred from saliency or attention model [3–6]. Generally, these techniques can be categorized as *content-aware cropping, seam carving, segmentation-based* and *warping-based* methods [1,2,7–9]. Though many efforts have been spent from the four main aspects during the past years, they only concentrate on how to keep the important content intact. As for the quality of the remaining areas in the resized results, few discussion is incurred by existing literatures. Besides, most methods focus on the shrinkage manipulation of the input image. As for the image′s expansion, less work has touched on this subject. Though some methods may extend the shrinking operation to expanding images in a similar manner [10,11], this will result in a significant degradation of image quality. Take Fig. 1 for example. The first column is the original images. The second one is the resized images by the seam carving method [11]. In this example, input images are expanded by 1.5 times in width. It is manifest that the resized results are stretched with a fuzzy resolution, especially in the green rectangled areas.

Based on this consideration, a high quality image resizing method is proposed in this work. The primary focus is on the image expansion procedure aiming at maintaining the high resolution property when stretching the input image. A typical result of the presented method is shown in Fig. 1(c).

### 1.1. Related work

The following discussion in this section is mainly restricted to the image expansion process. To the best of the authors′ knowledge, existing image resizing methods have not directly addressed the problem of high quality expansion. Almost all of them assume

* Corresponding author.
  *E-mail addresses:* crabwq@gmail.com (Q. Wang), yuany@opt.ac.cn (Y. Yuan).
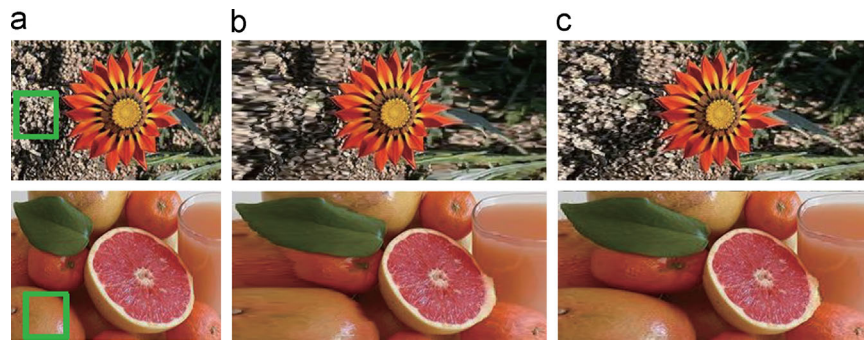
**Fig. 1.** Illustration of image resizing. In this example, images are enlarged 1.5 times in width. (a) Input images. Resized results by (b) the seam carving method and (c) the proposed method. It is clear that seam carving method generates an unsatisfying results because the stretched areas are fuzzy, especially for the green rectangled area. But the proposed method in this work ensures a high quality enlarging results with clear details. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

an interpolation for the stretched area [1,2] and take it for granted. But the unavoidable defect with this operation is a blurred result. This consequence cannot be accepted by most users, or at least it is unexpected.

For the *content-aware cropping* method, a smaller area compatible with the desired output is first cropped from the original image. Then the cropped sub-image is rescaled to the target size. For this purpose, firstly, either a searching window is employed or a user interaction is involved. Then the interpolation is used to enlarge the obtained sub-image. The whole procedure can be considered as a special interpolation that keeps and interpolates the target area and discards the unnecessary parts. The work by [12] is among the earliest ones to approach resizing with an automatic cropping method. Their focus is mainly on shrinking the images after finding the minimum rectangle containing adequate salient content. Ciocca et al. [13] fulfill the resizing task by first classifying the input images into three categories and second cropping them in different approaches. Nishiyama et al. [14] introduce a quality classifier to evaluate the agreement of the cropped window with human's perception. These methods mainly concentrate on the image shrinkage. When the images need to be expanded, interpolation based techniques are usually employed.

*Seam carving* method tries to insert the seams with minimum energy paths into the input image. With the inserted seams, the image is enlarged to the target size. When each seam is added, its color values are interpolated by its neighbors. For example, Avidan and Shamir [15] are the first to introduce seam carving concept. By defining a gradient related energy function, the seams with lowest energy can be identified. Then pixels are inserted at these positions by equally averaging their neighbors. Later, Rubinstein et al. [11] improve this method by using graph cut to find seams. Kumar et al. [16] devote to preserve edges and reduce aliasing artifacts by following the seam carving paradigm. Zhang et al. [17] combine the horizontal and the vertical seam carving with the uniform scaling to produce a bi-directional method. Basha et al. [18] put seam carving in the context of stereo vision and generate resized images compatible with stereo constraints. Besides, [19,20] address seam carving as an application of saliency detection. Though these works are improvements of the seminar work [15], the interpolation operation is used by all of them.

*Segmentation-based* method is another kind of resizing technique. It first segments the *region of interest* (ROI) combining with saliency detection results [21–23] or other higher level priors (e.g., face detection result). After that, the remaining image without ROI is scaled to the output size and the ROI is finally put back to the enlarged background image. In the rescaling step, interpolation is the common choice for previous literatures. The work by Setlur et al. [24] is the pioneer of this type. Its last step of scaling the ROI and background also involves an interpolation manipulation.

*Warping-based* method allows the important content to distort less and the unimportant content to twist more. By this means, the unpleasant transformation is primarily absorbed by the unimportant regions. The consequent expansion results are relatively better than those of other methods because the visual context is well preserved. However, the interpolation is also utilized in the process. Zhang et al. [25] define handles on the image grid mesh and adjust the aspect ratio by minimizing the associate distortion energy. Guo et al. [26] construct a mesh representation according to the image structure and address the resizing problem as mesh parametrization. Wolf et al. [27] formulate image resizing as solving a restricted linear system. Wang et al. [10] approach the resizing problem by distributing distortion in both horizontal and vertical directions. Kim et al. [28] partition the image into strips according to a gradient map and adjust them individually by an optimization process. In addition to these, there are also other techniques performing this nonlinear warping. But all of them encounter a problem of texture mapping. To solve this problem, interpolation is naturally adopted by these methods. Nevertheless, the interpolated results are not satisfying.

Moreover, there is another type of methods called *multi-operator* resizing. These methods assume that every method has both sides of advantages and disadvantages. An alternative option is to combine them optimally in the *resizing space*. Rubinstein et al. [29] seek an optimal transformation sequence of cropping, scaling, and seam carving. Dong et al. [30] jointly use seam carving and scaling to resize an image, whose key principle is a bidirectional similarity function that simultaneously incorporates image Euclidean distance and seam energy variation. However, when applied to image expansion, interpolation or duplication is also adopted.

### 1.2. Proposed method

In this paper, a high quality resizing method is proposed to retain the details when expanding the image. Instead of utilizing interpolation techniques to stretch the image, we first examine the relationship between the local smaller region and the expanded larger region. By learning the image expansion model, a more precise principle is obeyed in the future resizing operation. The consequent resizing results can enrich more details, making a pleasant visual effect.

As for the expansion model, it is learned from a set of training image pairs. There are about 30 images selected from the internet, containing a variety of textures. With this principle, we can ensure that the basic textual elements are included in the training procedure. Each pair contains a low resolution image and a high resolution image, with a fixed scale ratio. Then a dictionary pair is learned for mapping the low resolution image patch to a high resolution patch. This idea is inspired by the super-resolution

paradigm [31,32] but is different from super-resolution. In the future process, local expansion ratio is first estimated. A corresponding dictionary pair is then selected according to the estimated ratio. Finally, for the pixels (patches) that need to enlarged, they are expanded by the learned mapping principle. For the pixels (patches) that should stay unchanged, they are kept intact. Fig. 2 shows the flowchart of the proposed method.

The rest of this paper is organized as follows. Section 2 introduces the proposed high quality resizing method. Section 3 conducts experiments on two publicly available datasets to prove the effectiveness of the presented method. Section 4 extents the image resizing framework to video resizing. Discussions are given in Section 5 and conclusions are made in Section 6.

## 2. High quality image resizing

In this section, the details of the proposed high quality image resizing method are introduced. There are mainly three steps for the processing: learning expansion model, estimating local expansion ratio, and stretching the input image.

### 2.1. Learning expansion model

Traditional interpolation based method (e.g., bicubic, bilinear or nearest neighbor interpolation) is a simple and straightforward choice for enlarging an image. However, the obtained results will be overly smoothed or have ringing and jagged artifacts. This will result in a poor visual experience. Actually, the input image and its expanded counterpart are highly correlated, especially for those textured areas. The correlation is more complex than interpolation. Inspired by the success in super-resolution of learning a dictionary pair to map a low resolution image to a high resolution image [32], this paper approaches the image expansion problem on the basis of such a dictionary learning technique. But the super-resolution technique cannot be readily fit for our purpose because the image expansion ratio for image resizing is different in distinct areas. Even for one particular area, its expansion ratios in horizontal and vertical directions may be different. Therefore, several modifications have to be made to adapt to this application. In the following discussion, we will follow the notations of [32].

(1) *Problem formulation*: Given a low resolution image, the objective is to infer the high resolution image from it. The reconstruction is based on patches and sparse representation. To be more specific, for a low resolution patch $y$ from the input image, it is first represented by an over-complete dictionary $D_l$. This

representation is sparse and the constraint should be satisfied:

$$\min\|\alpha\|_0 \quad \text{s.t.} \quad \|FD_l\alpha - Fy\|_2^2 \leq \varepsilon, \tag{1}$$

where $\alpha$ is a coefficient vector with most entries being zeros, and $F$ is a feature extraction operator imposing perceptual constraint. Generally, $F$ is selected as some kind of high-pass filters, which is consistent with human perception. In this work, we choose the first and second order derivatives as the features. Optimizing Eq. (1) is NP-hard and it can be approximately approached by Lasso [33] as

$$\min_\alpha \|FD_l\alpha - Fy\|_2^2 + \lambda\|\alpha\|_1. \tag{2}$$

Solving the above optimization problem, we can get the desired $\alpha^*$. Accordingly, the corresponding high resolution patch $x$ is recovered by

$$x = D_h\alpha^*, \tag{3}$$

where $D_h$ is the high resolution dictionary jointly learned with $D_l$.

(2) *Joint dictionary training*: Training the low and high resolution dictionaries independently will cause large dimensionality and inefficient speed. A joint learning strategy is then adopted.

Given two sets of patches $X^h = \{x_1, x_2, ..., x_n\}$ and $Y^l = \{y_1, y_2, ..., y_n\}$ sampled from high and low resolution image pairs, the training process is formulated as

$$D_h = \arg\min_{\{D_h, Z\}}\|X_h - D_hZ\|_2^2 + \lambda\|Z\|_1, \tag{4}$$

and

$$D_l = \arg\min_{\{D_l, Z\}}\|Y_l - D_lZ\|_2^2 + \lambda\|Z\|_1, \tag{5}$$

where $Z$ is the representative coefficient that enforces sparsity. Jointly learning the two dictionary leads to

$$\min_{\{D_h, D_l, Z\}} \frac{1}{N}\|X_h - D_hZ\|_2^2 + \frac{1}{M}\|Y_l - D_lZ\|_2^2 + \lambda\left(\frac{1}{N} + \frac{1}{M}\right)\|Z\|_1, \tag{6}$$

where $N$ is the dimension of high resolution patches and $M$ is that of low resolution patches. By solving Eq. (6) with an alternative manner, the corresponding dictionary pair is obtained which represents the low resolution basic elements and its counterpart high resolution basic elements.

### 2.2. Estimating expansion ratio

Since the target aspect ratio differs with that of the input image, there must be distortions in the resized result. Content aware methods seek to minimize the distortions of important objects and make the unimportant regions absorb the most



Training data

*Expansion model*

Input image · Calculate local expansion ratio · Local mapping · Resized image
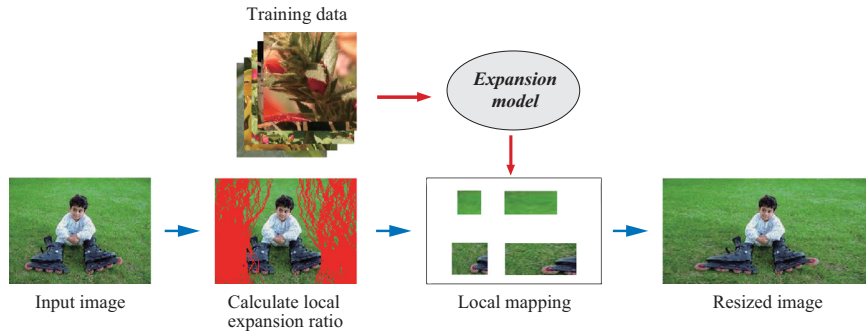
**Fig. 2.** Flowchart of the proposed method. In the training stage, an expansion model that projects a low resolution patch to a high resolution patch is learned from a set of training data. In the testing stage, pixels/patches that need to be expanded are first detected. At the same time, the local expansion ratio is determined according to the employed specific resizing technique. In this example, the seam carving is utilized to find the seams that should be expanded along the horizontal axis of the input image. Since the red seams imply the positions that should insert the same number of pixels, the expansion ratio is 2 in the horizontal direction. Thus they are locally stretched according to the learned model. After all these pixels/patches are enlarged, the final results are the desired images. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

distortions. This will cause the transformations of distinct operative elements vary with each other. (Element here may be one pixel in seam carving, or one grid in warping.) And for the same element, the transformation may be inhomogeneous in both horizontal and vertical directions.

To estimate the local expansion ratio, we should refer to the exact resizing algorithm. Assume the element set to be enlarged is denoted as $E = \{e_1, e_2, \ldots, e_n\}$, where $n$ is the total element number. For each $e_i$, suppose the expansion ratios in the horizontal and vertical directions are denoted as $r_{hi}$ and $r_{vi}$, respectively. Then its local expansion ratio is set as $r_i = \max(r_{hi}, r_{vi})$. That means the examined element is uniformly enlarged according to the largest expansion ratio, which would surely lead to a larger resized target patch than desired. This mismatch will be solved in the next step.

As for the estimation of $r_{hi}$ and $r_{vi}$, it depends on the employed specific technique. Here we will illustrate its determination in the context of seam carving as shown in Fig. 3, in which case the basic element is pixel. Suppose to enlarge the image in the horizontal direction. Seams are first identified according to [11]. Then each pixel of the seam will be expanded in double size along the horizontal direction. Therefore, $r_{hi} = 2$ and $r_{vi} = 1$ is established. This means $r_i = \max(r_{hi}, r_{vi}) = 2$. If the employed technique is warping based method, the treatment will be different. Extensions about this topic will be discussed in Section 5.

### 2.3. Stretching input image

For the input image, if every local element is enlarged independently, the computational cost will be very high and the results will not ensure the comparability between two adjacent elements. To handle this problem, we propose to define *element patch* as the basic processing unit instead of individual element. An element patch is a square containing the maximum number of elements with the same expansion ratio. Patches are obtained by a searching process, which starts from top-left to bottom-right by row order. Suppose the input image can be divided into a set of patches $P = \{p_1, p_2, \ldots, p_m\}$, where $m$ is the total patch number. Since $p_i$ is a squared shape composed of basic elements, the patch expansion
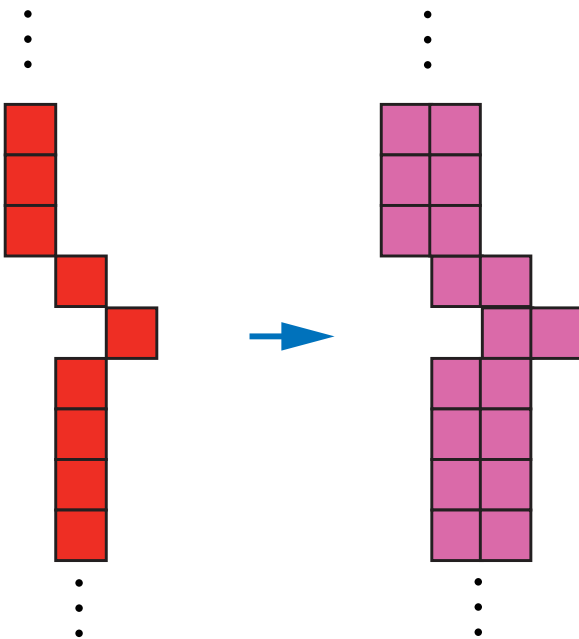


**Fig. 3.** Illustration of image resizing in the horizontal direction. Left: a path of detected seams in the input image. Right: the expanded result. In this seam carving example, $r_{hi} = 2$ and $r_{vi} = 1$.
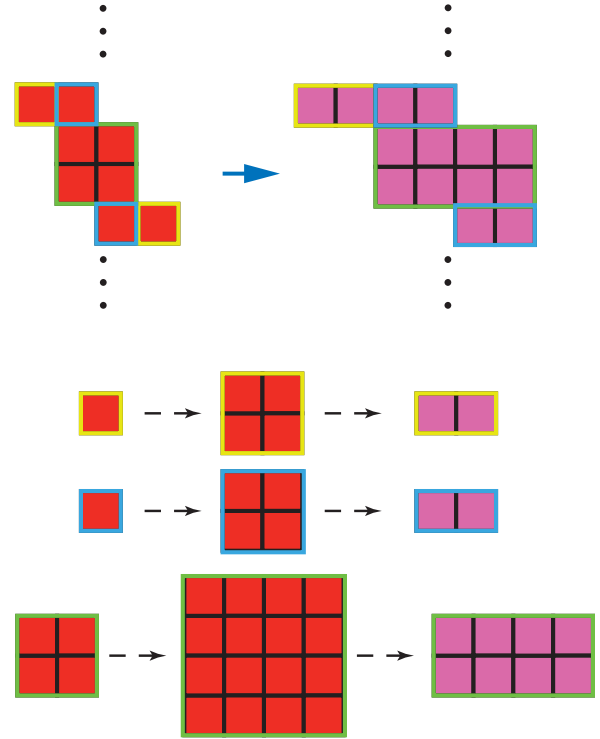


**Fig. 4.** Illustration of resizing each patch. The patches are identified as the maximum square containing elements with the same expansion ratio. In this figure, the squares with different boundary colors are the patches to be processed. For each patch $p_i$, it is first enlarged according to its $rp_i$ by the super-resolution technique. Then it is squeezed to the desired size. This illustration is also drawn in the context of seam carving. In this example, there are only $1 \times 1$ and $2 \times 2$ patches. For areas with high density seams, patches may be larger, such as $3 \times 3$ and $5 \times 5$. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

ratio is defined as its component expansion ratio, which is denoted as $rp_i$.

The element patch is first enlarged with super-resolution technique according to the expansion ratio $rp_i$ as $q_i = E(p_i)$, where $E$ represents the expansion operation. But due to the unevenly expected enlargement (In the seam carving example, $r_{hi} = 2$ and $r_{vi} = 1$. But $r_i$ is set to 2.), there must be one direction that stretches much more than desired. Therefore, after the patch is enlarged by the super-resolution method, it is then squeezed to the desired aspect ratio. In other words, the patch is reshaped in one dimension to the desired size. The final obtained patch is $q_i = S(E(p_i))$, where $S$ is the squeeze operation.

On one hand, this process can guarantee a finer detail of the resized image. On the other hand, it conforms to the target aspect ratio. The procedure is illustrated in Fig. 4 in the context of seam caring.

## 3. Experiments

In order to validate the proposed method, experiments are done on two publicly available datasets. The first one contains a total number of 1000 images, which is widely used in saliency detection and image resizing [34,35]. The second one is a collection of 80 images for a comparative study of image resizing algorithms [36]. On both datasets, seam carving [11] is employed to calculate local expansion ratio. Then the proposed method is utilized to expand the input image to 1.5 times in width. The results are finally compared with that of the original seam carving method. In addition, to guarantee a compatible result, neighboring

patches are allocated one pixel overlap. This makes a smooth transition between patches.

## 3.1. Results

On the first dataset, we employ a saliency detection algorithm to detect the saliency map of the input image. Then the obtained saliency map is used to guide the seam carving operation. Fig. 5 shows some typical examples. It is manifest that the detected seams are mainly in the background region, avoiding the salient objects. When expanding the input images, the original seam carving method employs interpolation to enlarge the image. This makes the results fuzzy and lack of details, especially for the regions with high density seams and obvious textures. But the proposed method can ensure a much better results with less distortions and rich of textures.

On the second dataset, the input image is directly used for seam carving. Since no prior information is incorporated in the resizing procedure, the detected seams scatter around the image. As shown in Fig. 6, there are less regions with high density seams compared with the above experiments. Consequently the improved performance may not be significantly manifest. But a close investigation

will find that the details are more abundant than the original seam carving method. For example, areas with a better performance are the Colosseum in the first row, the roof of the Japanese house and its frontal trees in the second row, the middle building in the background of third row, and the trees on the right in the fourth row. Therefore, the conclusion is consistent with the above experiments.

## 3.2. Subjective evaluation

The above analysis is a kind of intuitive. To get a more fair evaluation, we design an experiment for subjective assessment. Twenty participants are involved in this procedure and they are asked to rate the stretched images on the two datasets. Firstly, the original image is shown to the participant. Then the stretched images by seam carving method [11] and the proposed method are simultaneously displayed. The participant then records his scores on each result. The score is a real number ranging from 1 to 7, where 1 represents the stretched result is extremely poor and 7 highly pleasing. The score is given according to the participant's own visual experience and no other constraints are put on them. After every participant finishes his own evaluation independently,
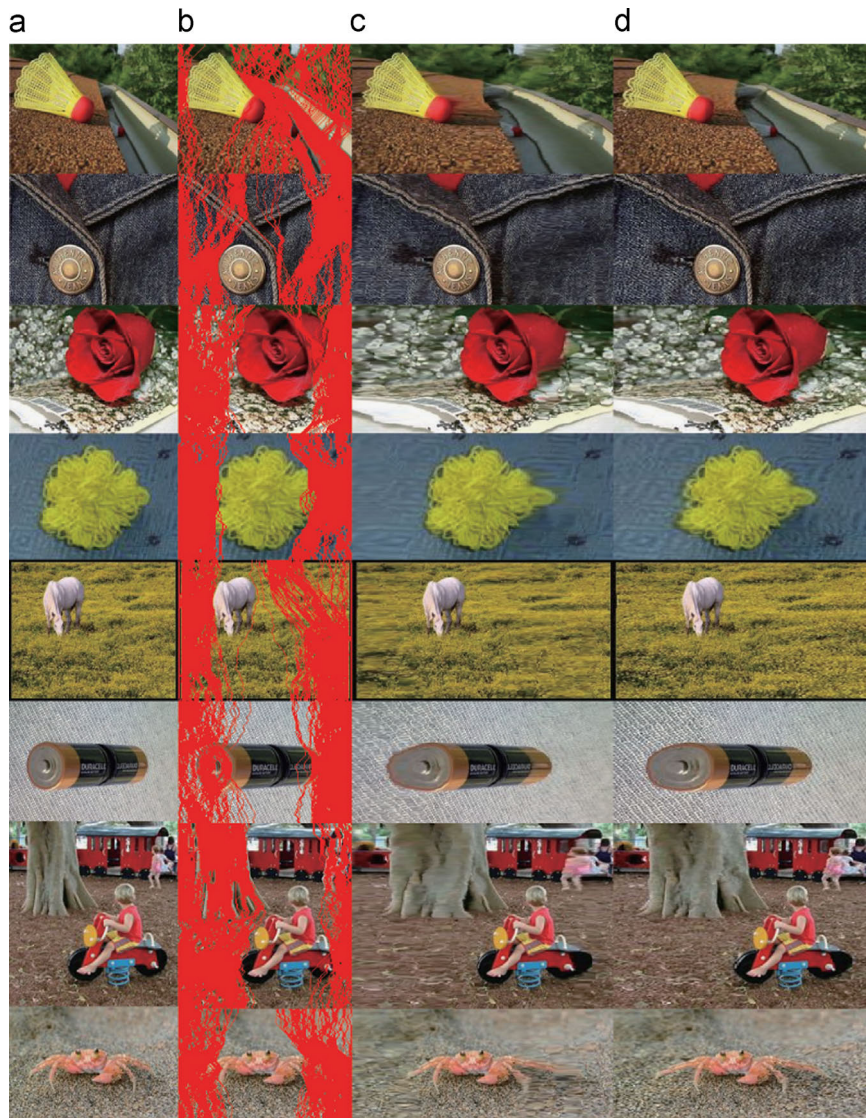


**Fig. 5.** Experiments on the first dataset. (a) Input images. (b) Detected seams. Resized results by (c) the seam carving method and (d) the proposed method.
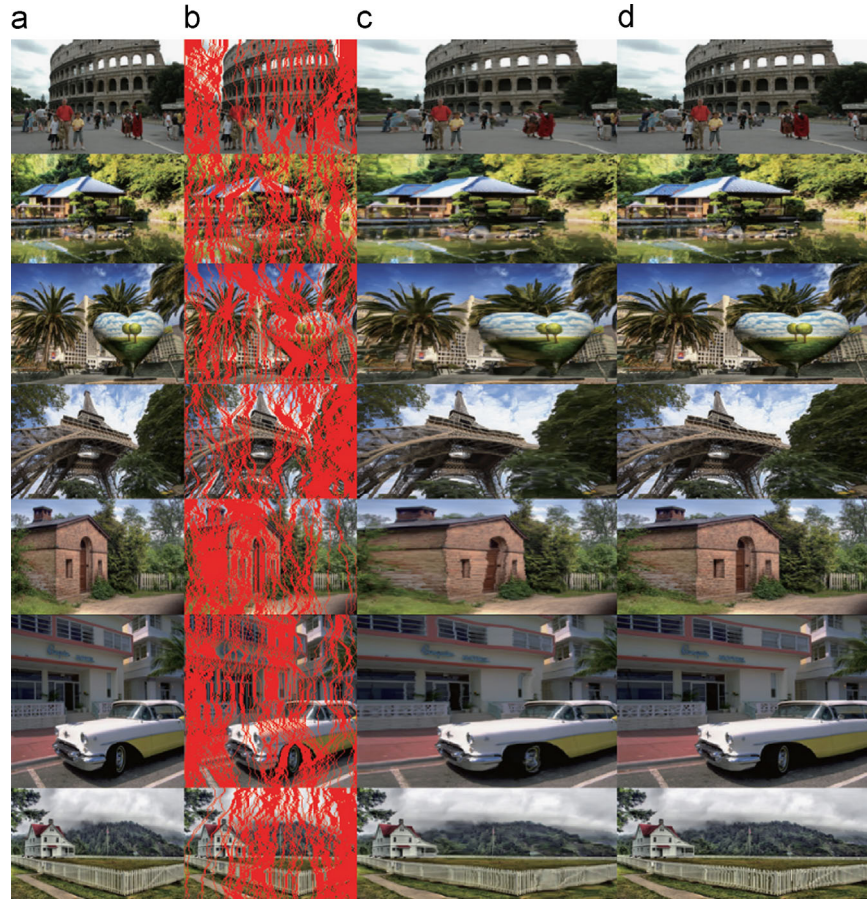
**Fig. 6.** Experiments on the second dataset. (a) Input images. (b) Detected seams. Resized results by (c) the seam carving method and (d) the proposed method.

the averaged scores for the two methods are then calculated. The final results are original seam carving (3.5) and the proposed method (5.5). It is clear that the proposed method generates a more pleasant resizing result than the original seam carving method.

To validate the effectiveness of the subjective evaluation results, we need to measure the consistency across the participants. This is estimated using one form of the *Intraclass Correlation Coefficient* psychological model [37,38]. To be specific, the $ICC(3, k)$ is employed for the task. It measures the consistency of the $k$ participants' mean ratings and is defined as

$$ICC(3, k) = \frac{bms - ems}{ems}, \qquad (7)$$

where $bms$ represents mean square of the ratings between targets, $ems$ means total error mean square, and $k$ is the number of participants. The values of $ICC$ range from 0 to 1, where 0 implies no consistency and 1 complete consistency. This model is used for every method to examine the correlation of participants' ratings. Since 20 participants are involved in the procedure, the $ICC(3, 20)$ for the original seam carving is *0.93* and the proposed method *0.94*. This indicates a high consensus among the participants, which proves that the obtained statistical results are meaningful and convincing.

### 3.3. Computational time

As for the computational time, it depends on the sizes of the input image and the target image. After the dictionary is trained beforehand and obtaining the high resolution frame and saliency map, we resize a frame of $200 \times 133$ to an output size of $r$ times in

**Table 1**
Comparison of computational time on different expansion ratios.

| Methods | Computational time (s) | | | |
|---|---|---|---|---|
| | $r=1.2$ | $r=1.4$ | $r=1.6$ | $r=1.8$ |
| Seam carving | 0.9568 | 1.9006 | 2.6811 | 3.5797 |
| Proposed method | 4.1756 | 7.0063 | 9.1285 | 10.377 |

width. All these computations are conducted in MatLab environment of a 2G memory and 2.93 GHz CPU computer. The obtained average time for each frame is listed in Table 1. It is clear that the proposed method is much slower than the original seam carving method. But the resized quality is improved a lot. Therefore, though it cannot be used in real time resizing applications,[1] it can be employed in off-line conditions which abound much in daily life.

### 4. Extension on video resizing

For video sequences, there are also works aiming at resizing for optimal display [39,40]. Nevertheless, the blurring effect of enlargement operation also exists. The proposed method can be straightforwardly applied to video resizing. However, we restrict our focus on simple videos that have static or slower background

---

[1] In fact, the processing can be speeded up by other means such as GPU.

motion compared with the foreground. More complicated videos require more advanced models of video contents, which is beyond the scope of this paper.

The first thing to stretch a video is to determine where to conduct this operation. Different areas of the input video have different importance. Our hope is to maintain the interesting objects undisturbed and to expand the unimportant contents. At the same time, to ensure the smoothness and coherence, the expanded areas in adjacent frames should be continuously near each other. For these purposes, we employ two clues from the video: saliency and optical flow. The first one considers the information in the individual frame and the second one incorporates the information between frames. Our assumptions are that (1) the salient areas are prone to be preserved, and (2) the areas with relatively larger optical flow fields are more likely to be the interesting objects.



**Fig. 7.** Illustration of identifying seams in a video clip. First row: frames from a video clip. (The three ones are taken every twenty frames.) Second row: seams to be expanded. Third row: stretched frames. The seams for each frame is the same and they are established as the most frequently appeared seams in the whole sequence. The interesting object can be maintained in the stretching processing and the obtained results are coherent.
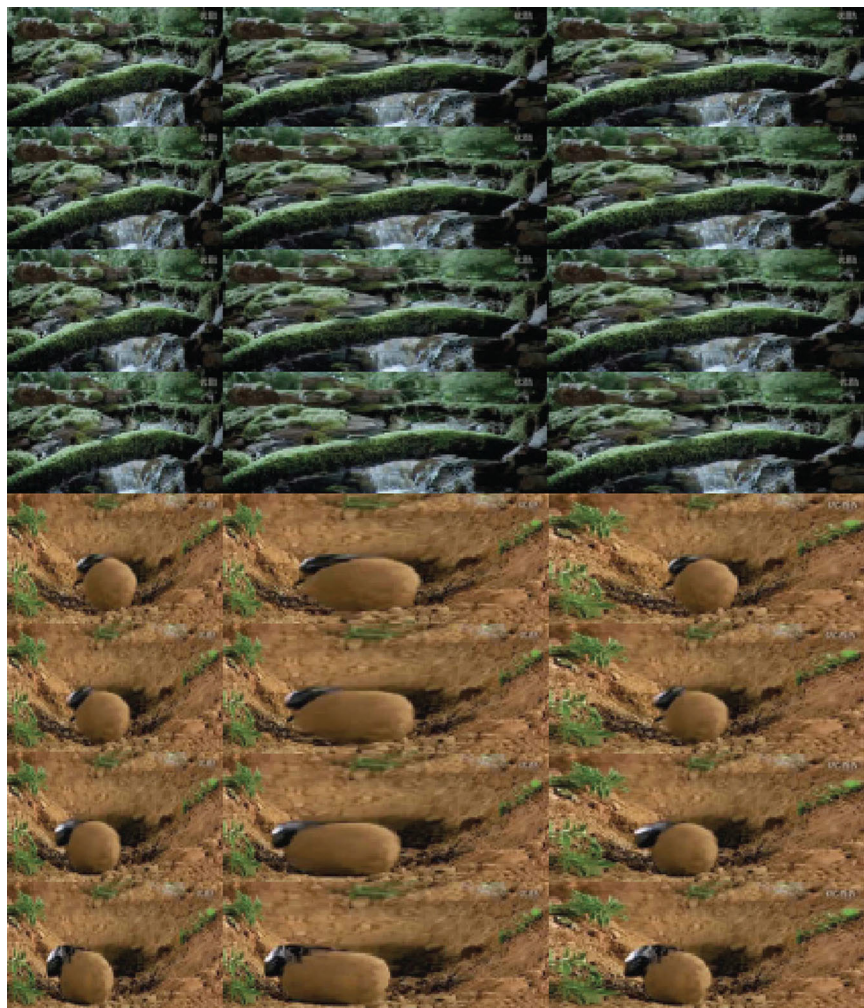


**Fig. 8.** Experimental results on two video clips are displayed here. For each video clip, four adjacent frames are illustrated. The first column is the input video frames. The second is the resized results by the seam carving method and the third is the resized results by the proposed method.
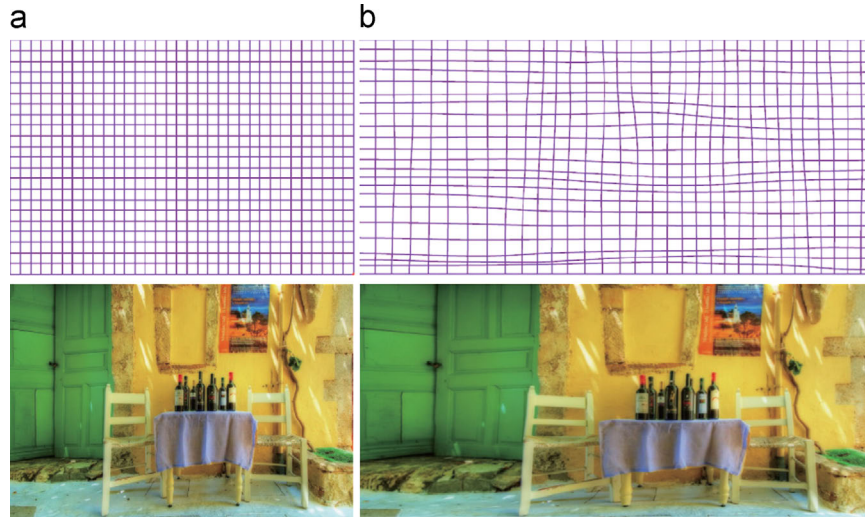
**Fig. 9.** Illustration of using warping based method to resize one image. (a) Input image and its original mesh grids. (b) Target result and its resized mesh grids. This image and its corresponding result are cited from Wang et al. [10].

Since the video content is changing continuously between adjacent frames, the detected saliency and calculated optical flow are also smoothly varied. But directly stretching the frames on the unsalient or small motion areas will still cause flickering effects. To solve this problem, we first determine the stretched areas in each frame. Then the areas most frequently adjusted in the whole sequence are chosen as the area to be stretched in the whole video. In this case, the video sequence will appear smoothly because the resized areas are kept identical among frames.

In the following, details about the energy definition by employing saliency and optical flow are introduced. The two clues act as guidance for defining the energy function in seam carving, where smaller value indicates less important regions in the video.

(1) Saliency: Humans will selectively pay more attention to some objects in the scene while unconsciously ignore other things. This phenomenon is studied in computer vision field as saliency detection. The detected salient objects are the ones that are prone to be maintained in visual processing. For the seaming carving, the motivation is similar that the seams should be located at the unsalient pixels. As for the specific saliency detection algorithm, we employ a *multiple instance learning* based work [41]. The obtained saliency value for each pixel $(i,j)$ is $S(i,j) \in [0,1]$. A larger value indicates a greater possibility to be salient.

(2) Optical flow: The motion field of object is a clue extracted between neighboring frames. Violent motion comes with a large optical flow value. Generally, the interesting objects have a larger movement than the background. Therefore, the seams should avoid these regions as much as possible. For calculating optical flow, two constraints should be satisfied, intensity constancy and smoothness constraint. An energy function combining the data term and smoothness term is finally minimized. Details about this topic can be found in [42]. After obtaining the optical flow field, their magnitudes are also normalized to $OF(i,j) \in [0,1]$. A larger value implies more violent motion.

(3) Combining saliency and optical flow: In seam carving, the energy function is defined based on the above two factors. Experimental results show that a linear combination of them suffices for our task

$$E(i,j) = \gamma S(i,j) + OF(i,j), \qquad (8)$$

where $\gamma$ is the balance between the two terms and is experimentally set to 0.3. A typical example by utilizing this energy function

to identify seams is illustrated in Fig. 7. It is clear that the final stretched results are coherent among consecutive frames.

To demonstrate the effectiveness of the proposed method, experiments are done on video sequences. There are eight video clips collected from the Internet and two typical examples are shown in Fig. 8. For each video sequence, it is processed by (1) applying the original seam carving method to each frame and (2) the proposed energy definition to get seams and dictionary based stretching method to expand each frame. After that, frames are combined into a video. Since we use the original seam carving method[2] to tackle individual images and no temporal constraint is considered, the obtained results are not coherent among frames and the interesting objects are unsatisfactorily distorted. On the contrary, the proposed method can ensure a smooth resizing visual experience. Most importantly, more details are maintained in the resizing procedure by employing the proposed method. The conclusion is consistent with the above experiments.

## 5. Discussion

In the above formulation, the local expansion ratio is calculated in the context of seam carving framework. This does not mean that the proposed method is only suitable for seam carving. We claim that the presented method can be embedded in various types of techniques but should be implemented differently. Take warping based methods for instance. Fig. 9 shows the resized result by *scale and stretch* [10], which belongs to the warping based paradigm. It is obvious that the individual mesh grid is resized differently, which indicates that the local expansion ratio for each grid is unequal. According to the definition that the *element patch* should be collection of grid elements with the same expansion ratio, the individual grid might be difficult to be clustered together as a patch because almost every grid has a distinct distortion. In this case, the patches are mostly equal to the single grids. As a result, the stretching process is conducted grid by grid. This is different from previously discussed seam carving.

As for cropping and segmentation based methods, the treatment is a little simpler. If the cropped sub-image or the

---

[2] The code is provided on the author's homepage. http://people.csail.mit.edu/mrub/.

background without segmented objects needs to be enlarged, each pixel has the same expansion ratio. Therefore, the following processing is similar to what seam carving undergoes.

## 6. Conclusion

In this paper, a high quality resizing method is proposed to retain the texture details when stretching an image. Different from traditional operations that directly utilize interpolation based methods, we first learn a mapping principle that projects the low resolution image patch to the corresponding high resolution one. Then local expansion ratio for the input image is calculated and patches are identified for further processing. In the end, each patch is mapped to the target size according to the learned dictionary. Experiments on various images demonstrate that the presented method is effective. Besides, the proposed method is also extended to simple video sequences with static or small background motion.

There are several remaining issues for future work. The first one is how to evaluate the experiments objectively. To the best of the authors' knowledge, there are no influential work for objective evaluation of resizing results. Though Rubinstein et al. [36] try to explore this subject with great efforts, the result is limited. We think that the key point is how to construct a ground truth benchmark. The second one is to make the proposed method much faster in an real application. This work is under way.

## Acknowledgments

## References

[1] A. Shamir, O. Sorkine, Visual media retargeting, in: SIGGRAPH ASIA Courses, ACM, 2009.
[2] V. Daniel, T. Matthew, P. Kari, T. Marius, G. Natasha, A survey of image retargeting techniques, in: Applications of Digital Image Processing XXXIII, vol. 7798, SPIE, 2010.
[3] W. Yang, Y.Y. Tang, B. Fang, Z. Shang, Y. Lin, Visual saliency detection with center shift, Neurocomputing 103 (2013) 63–74.
[4] S.-W. Ban, Y.-M. Jang, M. Lee, Affective saliency map considering psychological distance, Neurocomputing 74 (2011) 1916–1925.
[5] J. Han, E.J. Pauwels, P. de Zeeuw, Fast saliency-aware multi-modality image fusion, Neurocomputing (2013).
[6] Y. Niu, F. Liu, X. Li, M. Gleicher, Warp propagation for video resizing, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2010, pp. 537–544.
[7] Y. Liang, Z. Su, X. Luo, Patchwise scaling method for content-aware image resizing, Signal Process. 92 (2012) 1243–1257.
[8] H. Lim, H.W. Park, A ringing-artifact reduction method for block-DCT-based image resizing, IEEE Trans. Circuits Syst. Video Technol. 21 (2011) 879–889.
[9] D.-Y. Chen, Y.-S. Luo, Content-aware video resizing based on salient visual cubes, J. Vis. Commun. Image Represent. 22 (2011) 226–236.
[10] Y.-S. Wang, C.-L. Tai, O. Sorkine, T.-Y. Lee, Optimized scale-and-stretch for image resizing, ACM Trans. Graph. 27 (2008) 118.
[11] M. Rubinstein, A. Shamir, S. Avidan, Improved seam carving for video retargeting, ACM Trans. Graph. 27 (2008).
[12] B. Suh, H. Ling, B.B. Bederson, D.W. Jacobs, Automatic thumbnail cropping and its effectiveness, in: 16th Annual ACM Symposium on User Interface Software and Technology, ACM, 2003, pp. 95–104.
[13] G. Ciocca, C. Cusano, F. Gasparini, R. Schettini, Self-adaptive image cropping for small displays, IEEE Trans. Consumer Electron. 53 (2007) 1622–1627.
[14] M. Nishiyama, T. Okabe, Y. Sato, I. Sato, Sensation-based photo cropping, in: ACM Multimedia, ACM, 2009, pp. 669–672.
[15] S. Avidan, A. Shamir, Seam carving for content-aware image resizing, ACM Trans. Graph. 26 (2007) 10.
[16] M. Kumar, D.D. Conger, R.L. Miller, J. Luo, H. Radha, A distortion-sensitive seam carving algorithm for content-aware image resizing, Signal Process. Syst. 65 (2011) 159–169.
[17] X. Zhang, S. Zhong, Z. Pan, K. Wong, R. Yun (Eds.), Optimal Bi-directional Seam Carving for Content-Aware Image Resizing, vol. 6249, Berlin Heidelberg, Springer, 2010.
[18] T. Basha, Y. Moses, S. Avidan, Geometrically consistent stereo seam carving, in: IEEE International Conference on Computer Vision, IEEE, 2011, pp. 1816–1823.
[19] R. Achanta, S. Süsstrunk, Saliency detection using maximum symmetric surround, in: International Conference on Image Processing, IEEE, 2010, pp. 2653–2656.
[20] S. Goferman, L. Zelnik-Manor, A. Tal, Context-aware saliency detection, in: IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2010, pp. 2376–2383.
[21] Q. Wang, P. Yan, Y. Yuan, X. Li, Multi-spectral saliency detection, Pattern Recognit. Lett. 34 (2013) 34–41.
[22] Q. Wang, Y. Yuan, P. Yan, Visual saliency by selective contrast, IEEE Trans. Circuits Syst. Video Technol. 23 (2013) 1150–1155.
[23] Q. Wang, G. Zhu, Y. Yuan, Multi-spectral dataset and its application in saliency detection, Comput. Vis. Image Underst. 117 (2013) 1748–1754.
[24] V. Setlur, S. Takagi, R. Raskar, M. Gleicher, B. Gooch, Automatic image retargeting, in: International Conference on Mobile and Ubiquitous Multi-media, ACM, 2005, pp. 59–68.
[25] G.-X. Zhang, M.-M. Cheng, S.-M. Hu, R.R. Martin, A shape-preserving approach to image resizing, Comput. Graph. Forum 28 (2009) 1897–1906.
[26] Y. Guo, F. Liu, J. Shi, Z.-H. Zhou, M. Gleicher, Image retargeting using mesh parametrization, IEEE Trans. Multimed. 11 (2009) 856–867.
[27] L. Wolf, M. Guttmann, D. Cohen-Or, Non-homogeneous content-driven video-retargeting, in: IEEE International Conference on Computer Vision, IEEE, 2007, pp. 1–6.
[28] J.-S. Kim, J.-H. Kim, C.-S. Kim, Adaptive image and video retargeting technique based on fourier analysis, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2009, pp. 1730–1737.
[29] M. Rubinstein, A. Shamir, S. Avidan, Multi-operator media retargeting, ACM Trans. Graph. 28 (2009).
[30] W. Dong, N. Zhou, J.-C. Paul, X. Zhang, Optimized image resizing using seam carving and scaling, ACM Trans. Graph. 28 (2009).
[31] H. Chang, D.-Y. Yeung, Y. Xiong, Super-resolution through neighbor embed-ding, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2004, pp. 275–282.
[32] J. Yang, J. Wright, T.S. Huang, Y. Ma, Image super-resolution via sparse representation, IEEE Trans. Image Process. 19 (2010) 2861–2873.
[33] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. Ser. B (Methodological) 58 (1996) 267–288.
[34] R. Achanta, S. Hemami, F. Estrada, S. Susstrunk, Frequency-tuned salient region detection, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2009, pp. 1597–1604.
[35] M.M. Cheng, G.X. Zhang, N.J. Mitra, X. Huang, S.M. Hu, Global contrast based salient region detection, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2011, pp. 409–416.
[36] M. Rubinstein, D. Gutierrez, O. Sorkine, A. Shamir, A comparative study of image retargeting, ACM Trans. Graph. 29 (2010) 160.
[37] P.E. Shrout, J.L. Fleiss, Intraclass correlation: uses in assessing rater reliability, Psychol. Bull. 86 (1979) 420–428.
[38] Q. Wang, Z. Wang, A subjective method for image segmentation evaluation, in: ACCV (3), Springer, 2009, pp. 53–64.
[39] H. El-Alfy, D. W. Jacobs, L. S. Davis, Multi-scale video cropping, in: ACM Multimedia, ACM, 2007, pp. 97–106.
[40] F. Liu, M. Gleicher, Video retargeting: automating pan and scan, in: ACM Multimedia, ACM, 2006, pp. 241–250.
[41] Q. Wang, Y. Yuan, P. Yan, X. Li, Saliency detection by multiple-instance learning, IEEE Trans. Cybern. 43 (2013) 660–672.
[42] T. Brox, A. Bruhn, N. Papenberg, J. Weickert, High accuracy optical flow estimation based on a theory for warping, in: ECCV, Springer, 2004, pp. 25–36.

**Qi Wang** received the B.E. degree in automation and Ph.D. degree in pattern recognition and intelligent system from the University of Science and Technology of China, Hefei, China, in 2005 and 2010 respectively. He is currently an associate professor with the North-western Polytechnical University, Xi'an, China. His research interests include computer vision and pattern recognition.

**Yuan Yuan** is a full professor with the Chinese Academy of Sciences (CAS), China. Her major research interests include visual information processing and image/video content analysis. She has published over a hundred papers, including about 70 in reputable journals, like IEEE transactions and Pattern Recognition, as well as conferences papers in CVPR, BMVC, ICIP, ICASSP, etc.