Shrink Image by Feature Matrix Decomposition

Qi Wang[a], Xuelong Li[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.*

Abstract

With the development of multimedia technology, image resizing has been raised as a question when the aspect ratio of an examined image should be displayed on a device with a different aspect ratio. Traditional nonuniform scaling for tackling this problem will lead to distortion. Therefore, content-aware consideration is mostly incorporated in the designing procedure. Such methods generally defines an energy function indicating the importance level of image content. The more important regions would be retained in the resizing procedure and distortion could be avoided consequently. The definition of the related energy function is thus the critical factor that directly influences the resizing results. In this work, we explore the definition of energy function from another aspect, matrix decomposition of *Low-rank Representation*. In our processing, a feature matrix that reflects the texture prior of object contour is firstly constructed. Then the matrix is decomposed into a low-rank one and sparse one. The energy function for resizing is then inferred from the sparse one. We illustrate the proposed method through seam carving framework and image shrinkage operation. Experiments on a dataset containing 1,000 images demonstrate the effectiveness and robustness of the proposed method.

Keywords: computer vision, machine learning, image resizing, seam carving, low-rank representation, sparse coding

Email addresses: crabwq@gmail.com (Qi Wang), xuelong_li@opt.ac.cn (Xuelong Li)

1. Introduction

With the widespread growth of digital images, a variety of display devices are developed. For example, laptop, cell phone, iPad, and TV screen are the most popular ones. When images are displayed on these devices, it is desirable that they could be automatically adapted to the different sizes of display devices. This technique of adjusting images for optimal display is named as *image resizing*.

Because of its actual usefulness, image resizing has been extensively studied in computer vision and graphics. Earlier and straightforward work towards this problem is nonuniform scaling and cropping. However, transforming the image isotropically or merely cropping a fixed portion will lead to distortion and unbalance of image content. Trying to keep the interesting objects unaffected and compatible with the background is thus the main focus of recent research [1], [2]. According to the techniques employed, these content-aware methods can be categorized into *segmentation* based, *warping* based and *seam carving* based.

For segmentation based method [3], [4], the region of interest (ROI) [5], [6], [7] is firstly segmented from the background. Then the remaining background is scaled to the desired output size. In the end, the cropped foreground object is placed back to the resized image. Warping based method [8] treats the image as a group of grids and stretches each grid individually according to its distinctive property. The interesting objects are less affected and the background absorbs most distortions. As for seam carving based method [9], seams representing the minimum energy paths are removed or added in the image. In this way, the desired output dimension is obtained.

In this work, however, we are not intent on covering all the three aspects. The focus is mainly related to the seam carving paradigm [9]. As a comprehensive method, it contains several challenging issues, such as feature selection, energy definition and minimum cost path search. But the critical point is to define a suitable energy function to guide the seam searching process. Unfortunately, existing definitions are mostly restricted to local intensity gradients. The global statistics are not considered. This makes the extracted seams suboptimal in describing the interesting objects. Take Fig.1 for example. Original seam carving method emphasizes on the edges in an image, due to its gradient based nature. Since the seams crossing the ROI might have smaller energy (because the inside of ROI has no variations), the important objects are not protected well in the produced results [10].

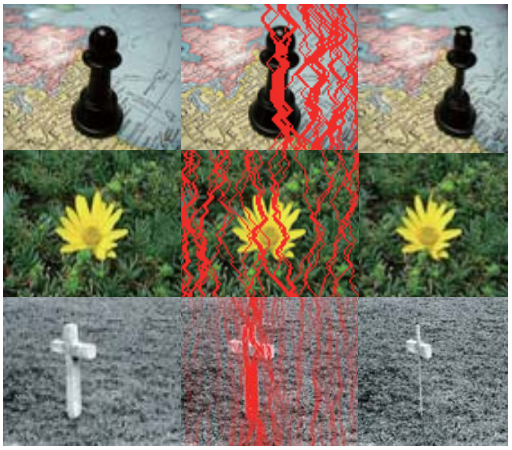

Figure 1: Examples for image resizing by Avidan and Shamir [9]. The first column is the original images, the second column is the removed seams, and the third column is the final resized results.

Therefore, a more feasible energy needs to distinguish the foreground with background [11], [12]. Motivated by this deficiency and inspired by [13], a novel energy definition is introduced in this paper. We propose to decompose the feature matrix into a low-rank one and sparse one. Then the sparse matrix is considered to be reflecting the ROI information. This decomposition treats the image as a whole in the first place and then splits it into two parts of different properties. It integrates the local and global statistics at the same time. Therefore, it is assumed to be more reasonable and effective.

1.1. Related work

Seam carving is one type of resizing methods that has achieved great popularity in recent years. It was proposed by Avidan and Shamir [9] and extended by many other researchers. Since the proposed method mainly focuses on the redefinition of the energy function for seam carving, the review is only restricted on the energy related literatures. A thorough survey can be found in [11].

There are many proposed methods for the modified energy function in the framework of seam carving. For ease of presentation, we classify them into three categories, which are respectively *straightforward* seam carving, *human detection* based seam carving and *saliency* based seam carving.

Straightforward seam carving is the direct improvements of the seminal

work by Avidan and Shamir [9]. Rubinstein *et al.* [14] defined a new energy function, which is called forward energy in the process of graph-cut. The new definition protects the interesting content better than the original backward energy in [9]. Han *et al.* [15] modified the energy function by weighing wavelet subband components. Conger *et al.* [16] generalized seam carving under the perspective of filter banks and redefined the backward energy and forward energy in the framework of the filter banks. Besides the redefined energy function, there are other ways to directly extend seam carving. For example, utilizing multi-operator [17] instead of single operator is a typical improvement. Rubinstein *et al.* [18] combined scaling, cropping, and seam carving together to resize images. Dong *et al.* [19] also combined seam carving with scaling by modifying the measure of image distance. However, these improved methods still result in aliasing [10] and the multi-operator ordering is difficult to determine either.

Human detection based seam carving is a transitional stage between the straightforward seam carving and the saliency based seam carving. The main operation is to extract human related features first such as face and gesture, and then to treat it as the ROI content that should be retained. Kopf *et al.* [20] proposed a method to pay more attention to objects, faces, and texts in images. In order to protect those regions, most of them are detected at first. Then the seams are ensured to round these regions. Subramanian *et al.* [21] employed neural network to detect the skin as the sign of human occurrence. Then the subsequent carving can avoid the possible human related region. However, not all images have a human related region. As for the images without human faces or gestures, this kind of method seems to be ill-suited. Thus more universal features expressing the salient region are needed.

Saliency based seam carving directly utilizes the output of saliency detection [22], [23]and treat it as the guidance to conduct seam carving. Achanta and Svsstrunk [24] presented an efficient seam carving method by utilizing saliency map. The employed saliency map, which is computed by intensity and color features, can be used to prevent seams passing through the ROI. Later after this work, various saliency detection methods are utilized in the application of image resizing. Typical methods include RC[25], AC[26], CA[12], FT[27], SC[28], and so on. These saliency based methods perform better than the ones merely based on gradient. The main reason is that the saliency measure tries to get the continuous and consistent expression of salient regions. It can reflect the ROI more accurately than the low level gradient.

1.2. The proposed method

To get an appealing resizing result, the fundamental importance is to identify the object of attention to human vision system. In order to represent these regions, lots of researchers have been working for a better description. There are mainly two ways for the efforts. One of them is to find more effective features to distinguish the salient regions from their surroundings. The other is to define more reasonable energy measure, which can make those regions more prominent. In this paper, the two aspects are simultaneously explored.

For the energy model, we adopt a mathematically motivated principle. Since the ROI usually holds part of an image or involves an inconsistent change with the background, which implies sparsity, we suppose the retargeting process may benefit from sparse signal analysis. To be specific, inspired by the recently proposed robust principle component analysis (RPCA) [29], we decompose the feature matrix F into a low-rank one L and a sparse one S:

$$\begin{aligned} \min_{L,S} \quad & \|L\|_* + \lambda \|S\|_1 \\ s.t. \quad & F = L + S \end{aligned} \tag{1}$$

where $\|\cdot\|_*$ represents the matrix nuclear norm, $\|\cdot\|_1$ denotes the ℓ_1 norm, and λ is a balance between them. The background pixels are prone to share similar properties. Therefore, they are highly correlated and correspond to the low-rank matrix. The foreground ROI mostly takes a small part of the image or has an inconsistent property with the background. Thus it can be explained by the sparse matrix.

For the features employed in the energy calculation, we adopt the single *texture*. Though the integrated multiple features might be more informative and descriptive, the computational cost is also a problem that should not be neglected. In this work, textures are expressed on an over-complete dictionary learned from a set of training images. In fact, the dictionary can be thought of as a series of prototype elements. Each patch can be expressed by a sparse linear combination of such basic elements according to research on image statistics [30].

To sum up, the main contribution of this work is an improved feature expression and energy formulation in seam carving procedure. Firstly, an over-complete dictionary is learned from a set of training images. These images are of various textures and contents. Secondly, a feature matrix is

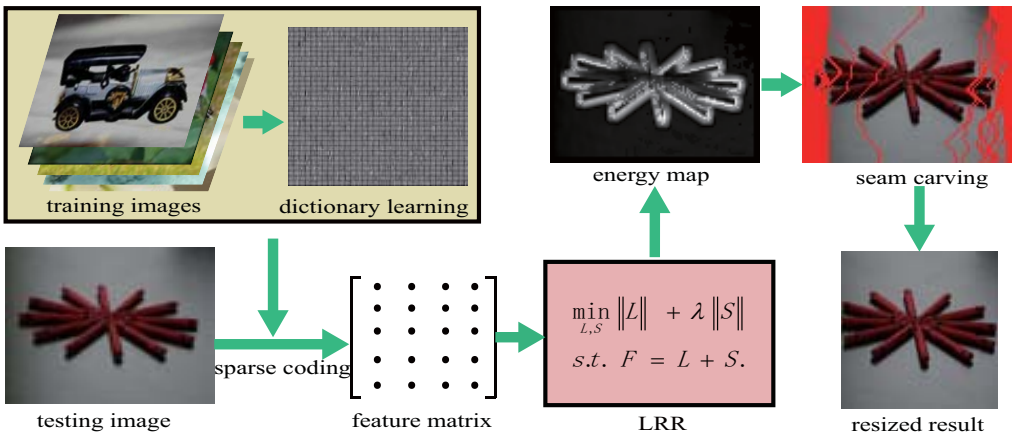

Figure 2: Flowchart of the proposed method.

constructed by concatenating each pixel's coefficient vector of texture expression on the dictionary. Thirdly, the obtained feature matrix is decomposed according to the *Low-rank Representation* (LLR) technique [13]. In the end, each pixel's energy is calculated by the decomposed sparse matrix and seam carving is conducted accordingly. Though the framework is mainly formulated in the context of image shrinkage, it can be readily extended to image enlargement. The flowchart of the proposed method is shown in Fig. 2.

The rest of this paper is organized as follows. Section 2 introduces the new energy definition. Section 3 presents the texture feature for constructing the feature matrix. Section 4 revisits the seam carving framework by putting the new energy and feature matrix together. To verify the effectiveness and robustness of the proposed method, experiments are conducted in Section 5. In the end, conclusions are made in Section 6.

2. Energy Definition

In this Section, we will introduce the proposed definition of energy function for seam carving purpose. Firstly, we formulate the problem in a formal manner. Secondly, the algorithm to find the solution is discussed.

2.1. Formulation

The problem we discussed is that given an input image $I_{m\times n}$, an energy function $E(i,j)$ should be defined. The assignment of $E(i,j)$ reflects the

importance of pixel (i, j). A bigger value indicates greater importance and larger possibility to be retained in the resizing processing. To obtain the desired energy function, we first construct a feature matrix $F_{d\times N}$ by concatenating each pixel's feature vector in row-major order, where $N = m \times n$ is the number of pixels in the image and each column of F is a $d \times 1$ feature vector representing a pixel's texture expression. Then the feature matrix is decomposed as two parts:

$$F = L + S, \tag{2}$$

where L is a low-rank matrix and S is a sparse matrix. To get the decomposition, RPCA [29] claims to minimize

$$\begin{aligned} \min_{L,S} \quad & \|L\|_* + \lambda\|S\|_1 \\ s.t. \quad & F = L + S. \end{aligned} \tag{3}$$

However, this formulation implicitly assumes that the underlying data lie on a single low-rank subspace. The actual situation, on the contrary, might not be the case. To better model the mixed data, [13] suggests a more general rank minimization problem:

$$\begin{aligned} \min_{Z,S} \quad & \|Z\|_* + \lambda\|S\|_{2,1} \\ s.t. \quad & F = AZ + S, \end{aligned} \tag{4}$$

where A is a dictionary spanning the data space, Z is the low-rank representation with respect to the dictionary A, and $\|\cdot\|_{2,1}$ is the $\ell_{2,1}$ norm defined as the sum of ℓ_2 norm of each column for a matrix:

$$\|S\|_{2,1} = \sum_j \sqrt{\sum_i (S(i,j))^2}. \tag{5}$$

In [31], they also prove that setting $A = F$ is more appropriate, leading to

$$\begin{aligned} \min_{Z,S} \quad & \|Z\|_* + \lambda\|S\|_{2,1} \\ s.t. \quad & F = FZ + S. \end{aligned} \tag{6}$$

After obtaining the sparse matrix S^* by decomposing the feature matrix, the energy for each pixel is calculated as

$$E(i,j) = \|S^*(:,k)\|_2, \tag{7}$$

where the kth column in S^* corresponds to the pixel (i,j) and for row-major order, $k = (i-1) \times n + j$.

2.2. Solving Eq.6

In the above Section, the energy is derived from decomposing the feature matrix. In this part, the solution to Eq.6 is discussed. According to [13], [32], this problem is convex and can be solved by Augmented Lagrange Multiplier (ALM) method [33]. We first convert Eq.6 to the following equivalent problem:

$$\begin{aligned} \min_{Z,S} \quad & \|J\|_* + \lambda\|S\|_{2,1} \\ s.t. \quad & F = FZ + S, Z = J. \end{aligned} \tag{8}$$

Then the augmented Lagrange function is defined as

$$\begin{aligned} \mathcal{L} = \|J\|_* + \lambda\|S\|_{2,1} + tr(Y_1^T(F - FZ - S)) + \\ tr(Y_2^T(Z - J)) + \frac{\mu}{2}(\|F - FZ - S\|_F^2 + \|Z - J\|_F^2), \end{aligned} \tag{9}$$

where $\mu > 0$ is a penalty parameter. This problem is unconstrained and thus it can be minimized in terms of J, Z and S, respectively. At the same time, the Lagrange multipliers Y_1 and Y_2 are updated, too. The optimization is conducted iteratively until the convergence condition is satisfied (The convergence of this algorithm is proved in [32]). This stepwise procedure is outlined in Algorithm 1. In order to clearly distinguish the notations in this Section, we list them together in Table 1.

3. Image feature

The feature matrix F is composed of columns of feature vectors, each column of which corresponds to a pixel's feature vector. In this work, only *texture* feature is utilized.

Algorithm 1 Solving Eq.6 by inexact ALM

Input: Feature matrix F, parameter λ.

Initialize: Set Z, J, S, Y_1, Y_2 equal to 0 and $\mu = 10^{-6}$, $\mu_{max} = 10^6$, $\rho = 1.1$, $\varepsilon = 10^{-8}$.

while not converged **do**

1: Fix the others and update J by

$$J = \arg\min \frac{1}{\mu}\|J\|_* + \frac{1}{2}\|J - (Z + Y_2/\mu)\|_F^2.$$

2: Fix the others and update Z by

$$Z = (I + F^T F)^{-1}(F^T(F - S) + J + (F^T Y_1 - Y_2)/\mu).$$

3: Fix the others and update S by

$$S = \arg\min \frac{\lambda}{\mu}\|S\|_{2,1} + \frac{1}{2}\|S - (F - FZ + Y_1/\mu)\|_F^2.$$

4: Update the multipliers

$$Y_1 = Y_1 + \mu(F - FZ - S),$$
$$Y_2 = Y_2 + \mu(Z - J).$$

5: Update the parameter μ by $\mu = min(\rho\mu, \mu_{max})$.

6: Check the convergence conditions:

$$\|F - FZ - S\|_\infty < \varepsilon \; and \; \|Z - J\|_\infty < \varepsilon.$$

end while

Table 1: List of notations

Notations	Descriptions
F	Feature matrix
L	Decomposed low-rank matrix from F
S	Decomposed sparse matrix from F
A	Dictionary spanning the data space
Z, J	Descriptive coefficients on A
E	Energy matrix associated with the pixel's saliency
$\mathcal{L}$	Lagrange function
Y_1, Y_2	Lagrange multipliers
μ	Penalty parameter

Texture is a type of effective feature to describe image characteristics. Typically, textures are represented by decomposing its spectrum through a group of filters [34], which reflect different orientations and spatial frequencies [35], [36]. But in this work, texture is sparsely expressed by a discriminative over-complete dictionary, which is learned by randomly selected patches in the training images. This is motivated by the success of [30], [37],[38],[39],[40]. In fact, this dictionary is a collection of basic texture elements. Each pixel can be projected onto the dictionary to get a feature vector of expressive coefficients. At the same time, this expression is sparse, i.e., most coefficients are zero and only a few are nonzero.

To learn such a dictionary, we adopt the method proposed by Yang *et al.* [30], which was firstly utilized in super-resolution application. To be specific, given the training data $X = \{x_1, x_2, \cdots, x_l\}$ composed of image patches, the dictionary D can be obtained by solving the following problem

$$\begin{aligned} D = arg \min_{D,Z} \|X - DZ\|_2^2 + \lambda \|Z\|_1 \\ s.t. \quad \|D_i\|_2^2 \leq 1, i = 1, 2, \cdots, l. \end{aligned} \tag{10}$$

where l is the number of feature atoms and the ℓ_1 norm is enforced in order to get the sparse expression. By solving the optimization problem in Eq.10, a sparse dictionary is learned. The optimization procedure is as follows:

1) Initialize D with a Gaussian random matrix, with each column unit normalized.

2) Fix D, update Z by

$$Z = \arg\min_{Z} \|X - DZ\|_2^2 + \lambda\|Z\|_1. \tag{11}$$

This can be solved by linear programming.

3) Fix Z, update D by

$$\begin{aligned} D &= \arg\min_{D} \|X - DZ\|_2^2 \\ s.t.\ \ &\|D_i\|_2^2 \leq 1, i = 1, 2, \cdots, l. \end{aligned} \tag{12}$$

This is Quadratically Constrained Quadratic Programming and can be solved by existing packages.

4) Iterate 2) and 3) until convergence.

In our image resizing application, the aim of energy definition is to distinguish the ROI from background. For this purpose, the training patches are only selected from the contours of ROIs in the training images. We hope to produce texture prototypes depicting the ROI-background differences. Fig. 3 shows the generated dictionary, which is learned by the above mentioned sparse coding. The patches are 5×5 in size and there are in total $l = 1024$ texture prototypes in the learned dictionary. With this dictionary, every pixel's texture is firstly represented by the patch around it. Then this patch is further projected to the dictionary, corresponding to a vector of coefficients.

4. Seam carving

For seam carving technique, a seam is defined as a connected path of pixels from top to bottom or left to right, having the minimum energy:

$$\begin{aligned} & S = \{s_i\}_{i=1:n}, \\ s.t.\ \ & \forall i,\ |s_i - s_{i-1}| \leq 1, \\ & S = arg\min_{s} \sum_{i=1}^{n} E(s_i), \end{aligned} \tag{13}$$

where s_i is a pixel in seam S, n is the number of pixels contained in seam S, and $E(s_i)$ the energy of pixel s_i defined by Eq.7.

An input image is resized by removing one seam at a time until appropriate number of seams are removed. Traditional seam carving defines the energy of a pixel as follows:

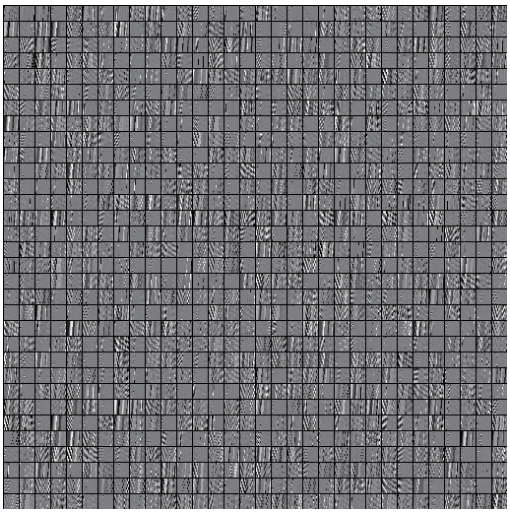

Figure 3: Illustration of the learned dictionary, with sparsity of $\lambda = 0.1$ and 1024 prototypes.

$$E(s_i) = |\frac{\partial}{\partial x} I(s_i)| + |\frac{\partial}{\partial y} I(s_i)|, \tag{14}$$

which is the sum of intensity gradients in horizontal and vertical directions. It is reasonable to a certain extent, because the larger gradient value might indicate change between the ROI and background. But most of the time, the local gradient cannot reflect the global information of the ROI. For this reason, we search for an alternative definition of the energy function.

In our proposed energy formulation, the feature matrix is decomposed into the low-rank one and sparse one. The low-rank one accounts for the stable background with high correlation between pixels, and the sparse one stands for the ROI which is different from most of the image pixels. This decomposition not only considers the local information of the pixels, but also treats the feature matrix as a whole to decompose. Therefore, the global statistics is also utilized.

With the obtained sparse matrix, the energy is calculated according to Eq.7. To get a visualized understanding, we scale the energy value to [0 255] and display it as a gray image. Fig. 4 demonstrates several typical examples.

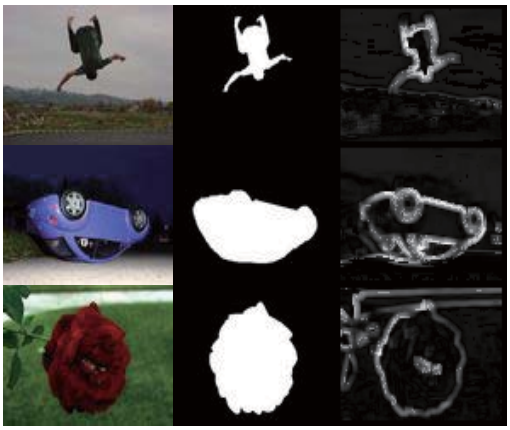

Figure 4: Visualization of the calculated energy map. First column: original images. Second column: the ground truth labelings of ROIs. Training patches are selected around the black-white contours. Third column: visualized energy maps as gray scale images.

5. Experiments

In this Section, several experiments are conducted to illustrate the effectiveness of the proposed method. The experiments utilize a dataset including 1,000 images. By shrinking the images to 75% width of the original size, the proposed method is evaluated. Experimental results are assessed not only subjectively, but also objectively.

5.1. Dataset

The experiments for the proposed method utilizes a dataset containing 1,000 images. It is constructed by Achanta *et al.* [27] and widely used in image processing applications. Every image in the dataset has a ground truth labeling of the ROI. This makes the objective evaluation possible. 500 images are randomly chosen to train the dictionary, and the remaining 500 images are used for testing. Fig.5 gives some examples for the dataset.

5.2. Evaluation measure

Most evaluations for image resizing are subjective. This is greatly influenced by different individuals. Few researchers have proposed quantitative evaluations for image resizing. The typical work is by Tong *et al* [41]. However, though it has some statistical analysis, it is still based on participants'

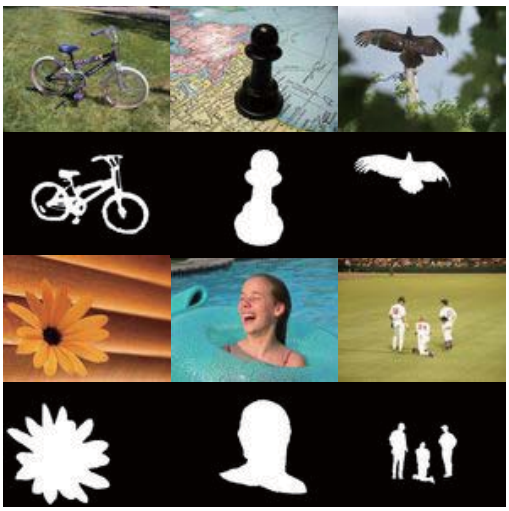

Figure 5: Several example images in the employed dataset. Each image in the dataset has a ground truth labeling of the ROI.

subjective judgements. For this reason, we propose an objective evaluation measure to test the failure error rate:

$$error\ rate = \frac{crossed\ seams}{total\ seams} \times 100\%, \tag{15}$$

where *crossed seams* are the seams cutting through the ground truth ROI, and the *total seams* represent all the removed seams. It reflects the percentage of *bad* seams in the resizing procedure. A higher value of this *error rate* indicates the worse resizing results.

5.3. Comparison with traditional seam carving

In this part, the proposed method is compared with the traditional seam carving method by Avidan and Shamir [9]. Fig. 6 shows some typical examples for the image resizing results. It can be seen that in the proposed method, the seams try to round the important objects in an image. There are much fewer seams passing through the ROIs compared with the traditional method. This is mainly because the new energy map can extract the salient contours of ROIs as the guiding clue for resizing procedure, which is much more informative than the gradient based energy definition. The strengthened energy around the contours prevents the seams passing through the ROIs.

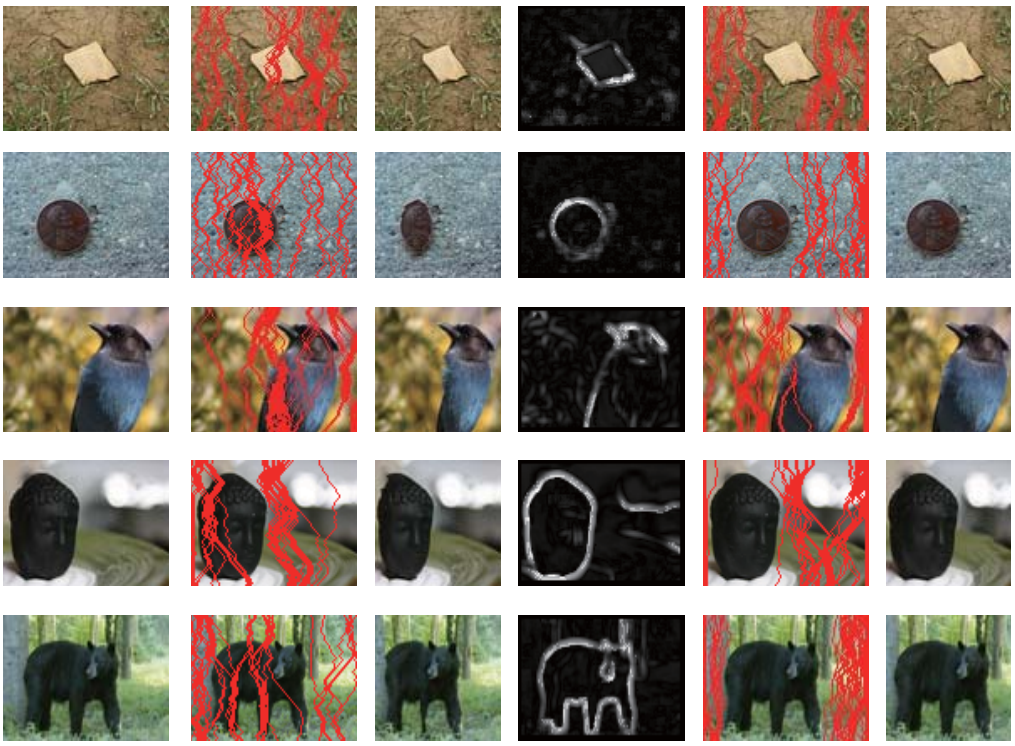

Figure 6: Comparison between the original seam carving method [9] and the proposed method. (a) Original images; (b) Results of removed seams by [9]; (c) Results of retargeted images by [9]; (d) Energy map by the proposed method; (e) Results of removed seams by the proposed method; (f) Results of retargeted images by the proposed method.

At the same time, in order to evaluate the results objectively, we plot the statistical results in Fig. 7. In this figure, we count the number of images with an error rate below a particular threshold. Though the image number for the seam carving method is larger than that of the proposed method at a lower error rate, the trend changes with the increase of error rate. This is because most of the results generated by seam carving have a larger error. It is thus evident that the proposed method outperforms the original seam carving.

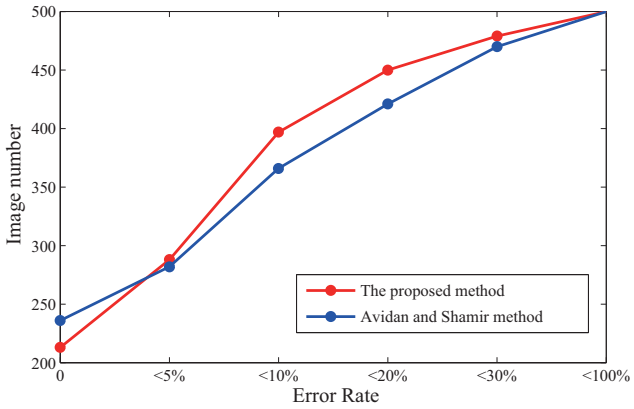

Figure 7: A quantitative comparison of the proposed seam carving method with [9] on 500 testing images. Horizontal axis: error rate below a particular threshold ; Vertical axis: the number of images with an error rate below the specified threshold.

5.4. Robustness to noise

Experiments are also conducted in order to evaluate the robustness of the proposed method. We add different levels of noises to the original images and then apply the proposed method and [9] to see their performance. The added noise is Gaussian with mean $\mu = 0$ and variance $\sigma^2 = 20, 40, 60, 80, 100$. Fig. 8 shows typical examples with added noises for the proposed method. It is obvious that with the increased noise level, the results of the image resizing are becoming worse, which is reasonable to understand. But all the time, the proposed method outperforms the original seam carving method. This quantitative comparison is shown in Fig. 9. We can see clearly that the heavy noise will induce more increased error. But the performance of the original seam carving decrease more rapidly than the proposed method.

5.5. Failure cases

The above experiments demonstrate the effectiveness and robustness of the proposed method. However, it is not always successful. Sometimes, the obtained results are not satisfying. There are mainly two reasons for this failure. The first one is that the ROI fills most of the image. No matter from which the seams are removed, the ROI will be impaired. The first image in Fig.10 illustrates this case. The second one is that occlusion often exists in

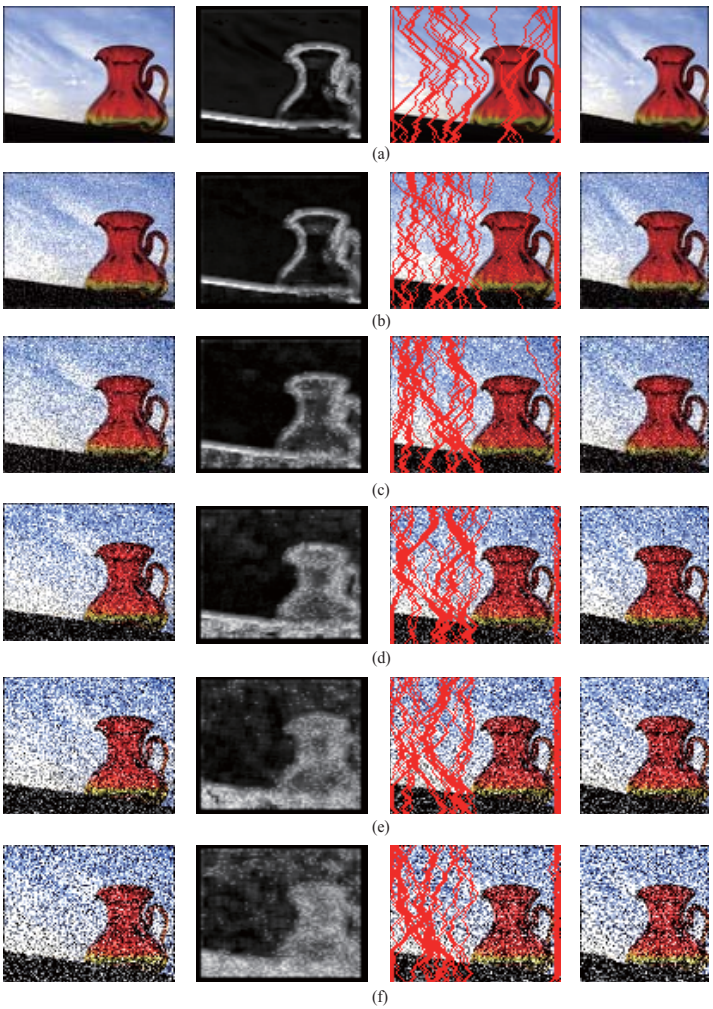

Figure 8: Resizing examples on the noisy data by the proposed method. The added Gaussian noise is (a) $\sigma^2 = 0$, (b) $\sigma^2 = 20$, (c) $\sigma^2 = 40$, (d) $\sigma^2 = 60$, (e) $\sigma^2 = 80$, and (f) $\sigma^2 = 100$. For each row, the first image is the original image, the second one is the energy map, the third one is the removed seams, and the last one is the retargeted result.

the image. If an object is occluded by the other, then the calculated energy might not reflect the actual situation. Take the second image in Fig.10 for example. The orange is partially covered by the leaves. But our method cannot know this spacial relationship. The consequent results are not well

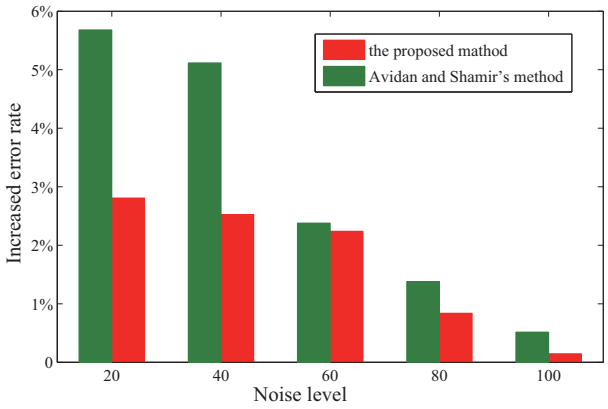

Figure 9: Comparison of increased error rates under different Gaussian noises. Horizontal axis: noise level; Vertical axis: increased error rate compared to the previous error (in percentage). It is obvious that with the increase of noisy level, the average error rate is increasing. But the original seam carving performs worse than the proposed one.

enough. The above two problems are difficult to resolve for current state-of-the-art method, which deserve further research in the future.

6. Conclusion

In this paper, the problem of energy redefinition is explored in the seam carving framework. Instead of relying on the local contrast based efforts, we explore the global statistics of an image through matrix decomposition. To be specific, the popular LRR technique and dictionary based texture expression are explored simultaneously. It decomposes the feature matrix into a low rank one and a sparse one. The sparse one is then assumed to reflect the ROI information, which guides the seam carving procedure. Experiments on a dataset of 1,000 images demonstrate the assumption is effective and robust.

Our work primarily focus on the energy redefinition step. We think it is a critical step for the resizing manipulation. Though the method is formulated in the framework of seam carving, it can be extended to other resizing techniques. For example, the warping based resizing also involves an energy

Figure 10: Two typical failures cases for the proposed image resizing method. First row: original images; Second row: resized results.

definition for each grid. The matrix based decomposition is readily fit for this case. Other improvements based on this idea can be derived similarly.

In the future, we plan to apply this framework to specific applications. More informative priors are supposed to be incorporated in these conditions. We think it might be a plausible way to tackle the problem raised in Section 5.5.

References

1. Tanaka, Y., Hasegawa, M., Kato, S.. Seam carving with rate-dependent seam path information. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*. 2011, p. 1449 – 1452.

2. Wang, Q., Yan, P., Yuan, Y., Li, X.. Multi-spectral saliency detection. *Pattern Recognition Letters* 2013;**34**(1):34 – 41.

3. Ma, M., Guo, K.. Automatic image cropping for mobile device with built-in camera. In: *IEEE Consumer Communications and Networking Conference*. IEEE; 2004, p. 710–711.

4. Liang, Y., Su, Z., Luo, X.. Patchwise scaling method for content-aware image resizing. *Signal Processing* 2012;**92**:1243–1257.

5. Cui, X., Liu, Q., Zhang, S., Yang, F., Metaxas, D.N.. Temporal spectral residual for fast salient motion detection. *Neurocomputing* 2012; **86**(0):24–32.

6. Ban, S.W., Jang, Y.M., Lee, M.. Affective saliency map considering psychological distance. *Neurocomputing* 2011;**74**(11):1916–1925.

7. Maybank, S.. A probabilistic definition of salient regions for image matching. *Neurocomputing* 2013;(0).

8. Ren, T., Liu, Y., Wu, G.. Image retargeting using multi-map constrained region warping. In: *ACM International Conference on Multimedia*. 2009, p. 853–856.

9. Avidan, S., Shamir, A.. Seam carving for content-aware image resizing. *ACM Transactions on Graphics* 2007;**26**(3).

10. Conger, D.D., Kumar, M., Miller, R.L., Luo, J., Radha, H.. Improved seam carving for image resizing. In: *IEEE International Workshop on Signal Processing Systems*. 2010, p. 345–349.

11. Vaquero, D., Turk, M., Pulli, K., Tico, M., Gelfand, N.. A survey of image retargeting techniques. In: *Society of Photographic Instrumentation Engineers*. 2010, p. 779814–779814–15.

12. Goferman, S., Zelnik-Manor, L., Tal, A.. Context-aware saliency detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2010, p. 2376–2383.

13. Liu, G., Lin, Z., Yu, Y.. Robust subspace segmentation by low-rank representation. In: *International Conference on Maching Learning*. 2010, p. 663–670.

14. Rubinstein, M., Avidan, S., Shamir, A.. Improved seam carving for video retargeting. *ACM Transactions on Graphics* 2008;**27**(3).

15. Han, J.W., Choi, K.S., Wang, T.S., Cheon, S.H., Ko, S.J.. Wavelet based seam carving for content-aware image resizing. In: *IEEE International Conference on Image Processing*. 2009, p. 345–348.

16. Conger, D.D., Kumar, M., Radha, H.. Generalized multiscale seam carving. In: *IEEE International Workshop on Multimedia Signal Processing*. 2010, .

17. Luo, S., Zhang, J., Zhang, Q., Yuan, X.. Multi-operator image retargeting with automatic integration of direct and indirect seam carving. *Image and Vision Computing* 2012;.

18. Rubinstein, M., Avidan, S., Shamir, A.. Multi-operator media retargeting. *ACM Transactions on Graphics* 2009;**28**(3).

19. Dong, W., Zhou, N., Paul, J.C., Zhang, X.. Optimized image resizing using seam carving and scaling. *ACM Transactions on Graphics* 2009; **29**(5).

20. Kopf, S., Kiess, J., Lemelson, H., Effelsberg, W.. Fscav-fast seam carving for size adaptation of videos. In: *ACM international conference on Multimedia*. 2009, p. 321–330.

21. Subramanian, S., Kumar, K., Mishra, B.P., Banerjee, A., Bhattacharya, D.. Fuzzy logic based content protection for image resizing by seam carving. In: *IEEE Conference on Soft Computing in Industrial Applications*. 2008, .

22. Wang, Q., Yuan, Y., Yan, P., Li, X.. Saliency detection by multiple-instance learning. *IEEE Transactions on Cybernetics* 2013;**43**(2):660–672.

23. Wang, Q., Zhu, G., Yuan, Y.. Multi-spectral dataset and its application in saliency detection. *Computer Vision and Image Understanding* 2013;**117**(12):1748 – 1754.

24. Achanta, R., sstrunk, S.S.. Saliency detection for content-aware image resizing. In: *IEEE International Conference on Image Processing*. 2009, p. 1005–1008.

25. Cheng, M., Zhang, G., Mitra, N.J., Huang, X., Hu, S.. Global contrast based salient region detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2011, p. 409–416.

26. Achanta, R., Estrada, F.J., Wils, P., Süsstrunk, S.. Salient region detection and segmentation. *Computer Vision Systems* 2008;:66–75.

27. Achanta, R., Hemami, S., Estrada, F., Susstrunk, S.. Frequency-tuned salient region detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2009, p. 1597–1604.

28. Wang, Q., Yuan, Y., Yan, P., Li, X.. Visual saliency by selective contrast. *IEEE Transactions on Circuits and Systems for Video Technology* 2013;**23**(7):1150–1155.

29. Wright, J., Ganesh, A., Rao, S., Peng, Y., Ma, Y.. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In: *Adv. Neural Inf. Process. Syst.* 2009, .

30. Yang, J., Wright, J., Huang, T., Ma, Y.. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing* 2010; **19**(11):2861–2873.

31. Lang, C., Liu, G., Yu, J., Yan, S.. Saliency detection by multitask sparsity pursuit. *IEEE Transactions on Image Processing* 2012; **21**(3):1327–1338.

32. Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y., Ma, Y.. Robust recovery of subspace structures by low-rank representation. *IEEE Transactions Pattern Analysis and Machine Intelligence* 2013;**35**(1):171–184.

33. Lin, Z., Chen, M., Ma, Y.. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. UIUC Technical Report UILU-ENG-09-2215; 2009.

34. Malik, J., Belongie, S., Leung, T., Shi, J.. Contour and texture analysis for image segmentation. *International Journal of Computer Vision* 2001;**43**(1):7–27.

35. Srinivasu, P., Avadhani, P.S., Satapathy, S.C., Pradeep, T.. A modified kolmogorov-smirnov correlation based filter algorithm for feature selection. *Advances in Intelligent and Soft Computing* 2012;**132**:819–826.

36. Ahonen, T., Pietikainen, M.. Image description using joint distribution of filterbank responses. *Pattern Recognition Letters* 2009;**30**(4):368–376.

37. Yu, J., Tao, D.. *Modern Machine Learning Techniques and Their Applications in Cartoon Animation Research*; vol. 4. John Wiley & Sons; 2013.

38. Yu, J., Tao, D., Wang, M.. Adaptive hypergraph learning and its application in image classification. *IEEE Transactions on Image Processing* 2012;**21**(7):3262–3272.

39. Yu, J., Wang, M., Tao, D.. Semisupervised multiview distance metric learning for cartoon synthesis. *IEEE Transactions on Image Processing* 2012;**21**(11):4636–4648.

40. Yu, J., Liu, D., Tao, D., Seah, H.S.. Complex object correspondence construction in two-dimensional animation. *IEEE Transactions on Image Processing* 2011;**20**(11):3257–3269.

41. Ren, T., Wu, G.. Automatic image retargeting evaluation based on user perception. In: *IEEE International Conference on Image Processing*. 2010, p. 1569–1572.

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 Northwestern Polytechnical University, Xi'an, China. His research interests include computer vision and pattern recognition.

Xuelong Li is a full professor with the 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, P. R. China.