

Perception-Driven Resizing for Dynamic Image Sequences

Lingling Zi^{1,2}, Junping Du^{1*}, Lisha Hou¹, Xiangda Sun¹, and Jangmyung Lee³

¹ Beijing Key Lab of Intelligent Telecommunication Software and Multimedia
School of Computer Science, Beijing University of Posts and Telecommunications
100876 Beijing, China

{lingling19812004,junpingdu,mssim1205,sunxiangda88}@126.com

² School of Electronic and Information Engineering, Liaoning Technical University
125105 Huludao, China

³ Department of Electronics Engineering, Pusan National University, Busan, Korea

Abstract. With the development of multimedia display devices, dynamic image sequence resizing, which can adapt image sequences to be displayed on devices with different resolutions, is becoming more important. However, existing approaches do not resize results from the viewpoint of the user. In this paper, we present a new resizing framework, which uses the feature descriptor technique and the image interpolation technique, that aims to improve the resizing quality of the important content perceived by user. To accomplish this, we use a coarse-to-fine detection approach to determine the important content of image sequences, and construct a partition interpolation model to improve the definitions of important content. By adopting a region energy protection approach we can obtain high quality image displays. Compared to representative algorithms in image resizing, our method can achieve satisfactory performance not only in terms of image visualization but also in terms of quantitative measures.

Keywords: image resizing, human visual perception, image Interpolation, feature descriptor, dynamic image sequence.

1. Introduction

Image sequences resizing has gained significant importance because of rapid increases in the diversity and versatility of display devices. Because of the limited screen sizes of different devices, the same image sequence is frequently required to be displayed in different sizes [16]. Therefore, it is common to resize the dynamic image sequences to fit each target display device. Standard resizing approaches include scaling and cropping. The former relies on a uniform ratio to resize the images, which results in obvious distortion [18]. The latter can only remove pixels from the image boundary and therefore inevitably discards information [17]. To overcome the above shortcomings, the seam carving

* Corresponding Author

method [2] has been proposed to resize image sequences through removing or duplicating seams [8]. Through analyzing those approaches, we find that they have only focused on inherent characteristics of image sequences and lack enough attention to resizing results from the viewpoint of the user. In this paper, we propose a novel perception-driven resizing algorithm (PDR). Two main issues need to be addressed to enhance the quality of the important content perceived by user. The first is the need to determine the important content from the users' viewpoint. The second is the need to improve the definitions of the important content during resizing. Therefore, we introduce an image processing technique based on human visual perception to address the first issue and an image interpolation technique to address the second issue. This allows us to obtain high quality resizing results displayed on devices with different resolutions.

Fig.1 shows an example of our proposed algorithm. Given the input image sequence (a), attention regions (b) are acquired before the corresponding interpolated images (c) are obtained. Next, we compute the energy of pixels in (c) as shown in (d), and the seams (including low energy pixels) are removed, as shown by the red lines in (e). Finally, resizing results are shown in (f).

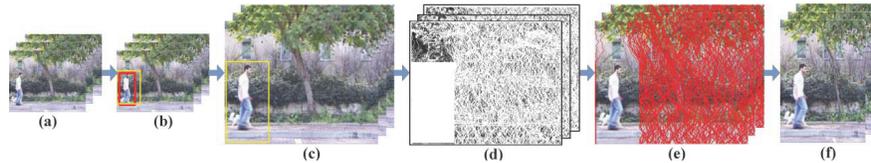


Fig. 1. An example of perception-driven resizing displaying the (a) input, (b) attention region, (c) interpolated image, (d) pixel energy display, (e) removed seams, and (f) output

Our main contributions can be summarized as follows: (1) Our paper complements the resizing task by providing a different perspective for dynamic image sequences; we propose a novel and comprehensive framework that considers the human attention problem and its framework to contain three parts, which are attention region establishment, partition interpolation model construction, and seam carving operation. (2) To capture attention regions, we propose a coarse-to-fine detection approach based on human visual perception. The foundation of the approach combines the frame difference with the latest feature descriptor technique. (3) We propose a partition interpolation model to improve the definitions of attention regions. Specifically, we use the latest image interpolation technique to enhance the quality of specified regions. (4) We propose a region energy protection approach using seam carving. Moreover, we redefine the energy formula of each pixel for two purposes. One is to protect the pixels in attention regions and the other is to accelerate seam carving operation.

The rest of this paper is structured as follows: Section 2 discusses related work. Section 3 illustrates the outline of PDR. Section 4 shows the implementation details of PDR. Section 5 presents experimental work. Finally, Section 6 concludes the paper.

2. Related Work

We complete the resizing task of dynamic image sequences by combining three technologies: the human visual perception technique, the feature descriptor technique, and the image interpolation technique. This section will briefly review the above technologies.

The human visual perception technique uses research results from many subjects [13] (e.g., computer neuroscience, visual neuroscience, cognitive science, and visual physiology) to accomplish visual tasks according to the principles of image perception and information processing used in the human visual system [9]. Humans perceive images through eye retina, which contains two different types of photoreceptors: rods and cones. Rods scatter light in the surface of the retina to determine an overall image, and they respond quickly to weak light. Cones are mainly located in the middle part of the retina [4]. The density of cones is much higher than other visual sensors in the visual central area, and the central area has almost no rods [11]. The visual central area is the most visually sensitive area; perception ability for optical input signals decreases further from the central area of the retina [3]. This is the source of inspiration for PDR.

People hope that computers can identify the objective world and make wise decisions. However, computers do not possess enough memory to distinguish between different complex objects. That is because people have not found an effective way to convert from real objects into a digital form. As a result of 21st century advances in the theory of discrete scale space, local invariant feature techniques have been developed to solve practical problems in the computer vision [15]. Subsequently, the scale-invariant feature transform (SIFT) was developed to effectively describe object features [7]. Recently, the method using compact and real-time descriptors (CARD) has been proposed to establish a quick visual correspondence between two images [1] and the computation time per descriptor is approximately 16 times faster than SIFT. In an attempt to determine the attention region of dynamic image sequences, we adopt CARD to acquire local descriptors caused by temporal variation of the consecutive images.

To maintain image quality in the process of resizing, we use the image interpolation technique [6][10]. The traditional interpolation technique is the nearest interpolation [14] and can quickly obtain high-resolution images [12]. The latest interpolation technique, an artifact-free image upscaling method called ICBI, which was proposed by Giachetti et al. in 2011 [5], can capture image details more clearly. The detailed implementation of ICBI contains two steps. The first step computes the interpolated pixel values based on the gradient direction of

the neighboring pixels. The second step iteratively modifies the obtained pixel values based on energy computation, which is the highlight of ICBI. Many experiments have shown that ICBI obtains good performance in objective and subjective tests. In this paper, we apply ICBI to produce clear resizing results for the attention regions of dynamic image sequences.

3. Overview

Let $DI[1, m] = \{DI_{\{t\}}\}_{t=1}^m$ be a dynamic image sequence DI defined on the 3D space O . The space O is divided into attention regions O_{ar} and general regions O_{gr} ,

$$O = O_{ar} \cup O_{gr}, O_{ar} \cap O_{gr} = \phi. \quad (1)$$

The DI is decomposed using $DI_{\{t\}}^O = (DI_{\{t\}}^{O_{ar}}, DI_{\{t\}}^{O_{gr}})$. In this paper, our goal is to resize $DI_{\{t\}}^O$ from $H_{initial} \times W_{initial}$ to $H_{new} \times W_{new}$. Here our stress is enlarging resizing, i.e., satisfying $H_{new} > H_{initial}$ and $W_{new} > W_{initial}$. To achieve this amount of resizing, we propose a PDR algorithm with an architecture that is illustrated in Figure 2. PDR is roughly divided into the three steps: attention region establishment, partition interpolation model construction, and seam carving operation.

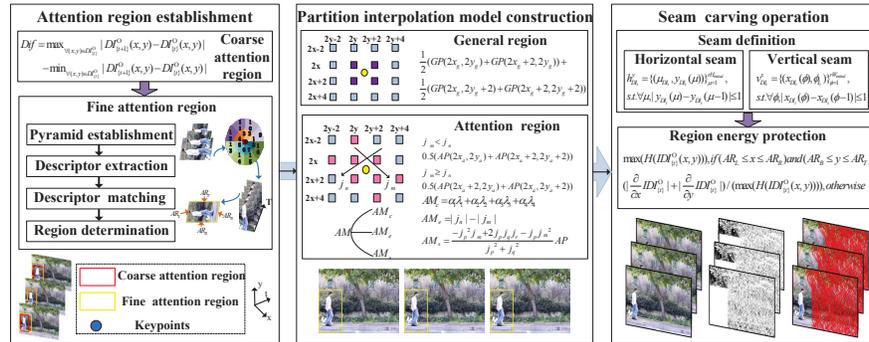


Fig. 2. Architecture of the proposed algorithm

First, we propose the coarse-to-fine detection approach to compute $DI_{\{t\}}^{O_{ar}}$, so that the corresponding $DI_{\{t\}}^{O_{gr}}$ can be obtained. Specifically, the coarse attention region $CDI_{\{t\}}^{O_{ar}}$ is determined using a frame difference method and the fine attention region $DI_{\{t\}}^{O_{ar}}$ is obtained using the feature descriptor technique, whose calculating process includes pyramid establishment, keypoint descriptor extraction, descriptor matching, and attention region determination.

Second, we construct a partition interpolation model to simulate the effects of human visual perception, resulting in interpolated image $IDI_{\{t\}}^O$ with

a scale ratio r , i.e., the size of the $IDI_{\{t\}}^O$ is $rH_{initial} \times rW_{initial}$. Note that this model adopts different calculation approaches for different regions of the original $DI_{\{t\}}^O$. For $DI_{\{t\}}^{O_{ar}}$, we compute the initial value of interpolated pixel, denoted as $AP(2x_a + 1, 2y_a + 1)$, where $x_a \in O_{ar}$ and $y_a \in O_{ar}$. The initial value is modified using energy calculations that include the curvature continuity energy AM_c , the curvature enhancement energy AM_e , and the isolevel curve smoothing energy AM_s . For $DI_{\{t\}}^{O_{gr}}$, we use four adjacent pixels to quickly compute interpolated pixel value, denoted as $GP(2x_g + 1, 2y_g + 1)$, where $x_g \in O_{gr}$ and $y_g \in O_{gr}$.

Finally, we compute the overall energy for each pixel of $IDI_{\{t\}}^O$. Adopting the proposed region protection method, we remove $(H_{new} - rH_{initial})$ vertical seams and $(W_{new} - rW_{initial})$ horizontal seams containing low energy pixels, and obtain the enlarged dynamic image sequence $DI_{\{t\}}^O$ with $H_{new} \times W_{new}$.

4. PDR Implementation

In this section, we will elaborate the details of the PDR in each step. The core part of PDR is the second step, which is the partition interpolation model construction. The first step is the foundation of partition interpolation model construction and the last step is resizing using a constructed partition interpolation model.

4.1. Attention region establishment

First, we demonstrate the calculation approach of $CDI_{\{t\}}^{O_{ar}}$. Given that $\forall(x, y)$, $(x, y) \in DI_{\{t\}}^O$, if $|DI_{\{t+1\}}^O(x, y) - DI_{\{t\}}^O(x, y)| \geq Dif \times c$, then $(x, y) \in CDI_{\{t\}}^{O_{ar}}$. Here, Dif represents the difference value, as shown in Eq.2, and c is a predefined value, for which we set the default value to 0.2.

$$Dif = \max_{\forall(x,y) \in DI_{\{t\}}^O} |DI_{\{t+1\}}^O(x, y) - DI_{\{t\}}^O(x, y)| - \min_{\forall(x,y) \in DI_{\{t\}}^O} |DI_{\{t+1\}}^O(x, y) - DI_{\{t\}}^O(x, y)|. \quad (2)$$

Next, we elaborate the calculation approach of $DI_{\{t\}}^{O_{ar}}$ in detail. The first step is to establish an image pyramid using Eq.3 as follows:

$$L(x_m, y_m, t, \sigma - 1) = L_{2\downarrow}(x_m, y_m, t, \sigma), \quad (3)$$

where $L(x_m, y_m, t, \sigma)$ represents scale function, x_m and y_m represent the values of abscissa and ordinate of the pixel in $CDI_{\{t\}}^{O_{ar}}$, σ represents the level of the image pyramid, and $2 \downarrow$ represents the down sampling factor. Using the corner detectors technique, we find the keypoints of the $CDI_{\{t\}}^{O_{ar}}$ and the corresponding keypoint set denoted as $K_t = \{k_t(x_k, y_k) | (x_k, y_k) \in CDI_{\{t\}}^{O_{ar}}\}$.

The second step is to extract descriptors for K_t . To maintain invariance to rotation, translation, and scaling, we establish an orientation histogram $AH(h_k^t)$

of K_t . The magnitude and orientation of $AH(h_k^t)$ are shown as follows:

$$m_{AH}(x_i, y_i) = ((CDI_{\{t\}}^{Oar}(x_i + 1, y_i) - CDI_{\{t\}}^{Oar}(x_i - 1, y_i))^2 + (CDI_{\{t\}}^{Oar}(x_i, y_i + 1) - CDI_{\{t\}}^{Oar}(x_i, y_i - 1))^2)^{\frac{1}{2}}, \quad (4)$$

$$\omega_{AH}(x_i, y_i) = \arctg \frac{CDI_{\{t\}}^{Oar}(x_i + 1, y_i) - CDI_{\{t\}}^{Oar}(x_i - 1, y_i)}{CDI_{\{t\}}^{Oar}(x_i, y_i + 1) - CDI_{\{t\}}^{Oar}(x_i, y_i - 1)}. \quad (5)$$

Next, we adopt the binning table to extract descriptors based on $AH(h_k^t)$ and obtain the descriptor set $D_k^t = \{d_k^t\}$ of K_t .

The third step is to find the matched descriptor. According to Eq.6, we compute the matching descriptor d_{km}^{t+1} for d_k^t and the corresponding matching descriptor set K_m .

$$d_{km}^{t+1} = \arg \min_{d_{\tau}^{t+1} \in D_{\tau}^{t+1}} \|d_{\tau}^{t+1} - d_k^t\|_1, \quad (6)$$

where $D_k^{t+1} = \{d_k^{t+1}\}$ denotes the descriptor set of K_{t+1} corresponding to $CDI_{\{t+1\}}^{Oar}$. In order to make the computation proceed faster, we adopt a short binary code conversion technique that changes Eq.6 into Eq.7:

$$d_{km}^{t+1} = \arg \min_{p \in K_t} (\sum_{q=1}^L |\zeta_p^{t+1}(q) - \zeta_k^t(q)|), \quad (7)$$

where $\zeta_k^t = (1 + \text{sgn}(W^T d_k^t))/2$, $\zeta_k^t \in \{0, 1\}^L$, in which W represents the weight matrix, and L represents the length of changed binary code.

In the last step, we acquire four boundary values of $DI_{\{t\}}^{Oar}$, including the left border AR_L , right border AR_R , top border AR_T , and bottom border AR_B .

$$\begin{aligned} AR_L &= \min_{k(x_i, y_i) \in K_t \cup K_m} x_i - v_L, AR_R = \max_{k(x_i, y_i) \in K_t \cup K_m} x_i + v_R, \\ AR_T &= \min_{k(x_i, y_i) \in K_t \cup K_m} y_i - v_T, AR_B = \max_{k(x_i, y_i) \in K_t \cup K_m} y_i + v_B, \end{aligned} \quad (8)$$

where v_L , v_R , v_T and v_B denote predefined adjustment values, the default values of which are all 8.

Fig.3 shows some examples of attention regions. We use red lines to mark the coarse attention region, and yellow lines to mark the fine attention region containing feature points denoted by blue dots. It can be seen that our detection approach can perform well in capturing regions of general interest, which provides a good foundation for improving the quality of resizing results.

4.2. Partition interpolation model construction

The computation of interpolated pixels in the general region is simple, as shown in Eq.9.

$$GP(2x_g + 1, 2y_g + 1) = \frac{1}{2}(GP(2x_g, 2y_g) + GP(2x_g + 2, 2y_g)) + \frac{1}{2}(GP(2x_g, 2y_g + 2) + GP(2x_g + 2, 2y_g + 2)). \quad (9)$$



Fig. 3. Attention regions of different dynamic image sequences: (a) a walking man, (b) a jumping woman, (c) a walking man with a dog, (d) space image sequence 1, and (e) space image sequence 2

Inspired by the ICBI algorithm, the computation of interpolated pixels in the attention region includes two steps. In the first step, we compute $AP(2x_a + 1, 2y_a + 1)$ according to Eq.10, Eq.11 and Eq.12.

$$\begin{aligned}
 j_m(2x_a + 1, 2y_a + 1) = & AP(2x_a, 2y_a) + AP(2x_a + 4, 2y_a) + \\
 & AP(2x_a + 2, 2y_a - 2) + AP(2x_a, 2y_a + 4) + AP(2x_a - 2, 2y_a + 2) \\
 & + AP(2x_a + 2, 2y_a + 2) - 3AP(2x_a, 2y_a + 2) - 3AP(2x_a + 2, 2y_a), \quad (10)
 \end{aligned}$$

$$\begin{aligned}
 j_n(2x_a + 1, 2y_a + 1) = & -3AP(2x_a, 2y_a) + AP(2x_a + 4, 2y_a + 2) \\
 & + AP(2x_a + 2, 2y_a + 4) + AP(2x_a, 2y_a - 2) + AP(2x_a - 2, 2y_a) \\
 & + AP(2x_a + 2, 2y_a) + AP(2x_a, 2y_a + 2) - 3AP(2x_a + 2, 2y_a + 2), \quad (11)
 \end{aligned}$$

$$AP(2x_a + 1, 2y_a + 1) = \begin{cases} \frac{1}{2}(AP(2x_a, 2y_a) + AP(2x_a + 2, 2y_a + 2)), \\ \text{if } j_m(2x_a + 1, 2y_a + 1) < j_n(2x_a + 1, 2y_a + 1), \\ \frac{1}{2}(AP(2x_a + 2, 2y_a) + AP(2x_a, 2y_a + 2)), \text{ otherwise.} \end{cases} \quad (12)$$

In the second step, we continually modify $AP(2x_a + 1, 2y_a + 1)$ using iterative energy computation to accurately predict the values of unknown pixels. The formula of energy computation is

$$\begin{aligned}
 AM(2x_a + 1, 2y_a + 1) = & m_c AM_c(2x_a + 1, 2y_a + 1) \\
 & + m_e AM_e(2x_a + 1, 2y_a + 1) + m_s AM_s(2x_a + 1, 2y_a + 1), \quad (13)
 \end{aligned}$$

where m_c, m_e and m_s are modification coefficients that satisfy $m_c + m_e + m_s = 1$. The first energy term AM_c can be computed using Eq.14.

$$\begin{aligned}
 AM_c(2x_a + 1, 2y_a + 1) = & \alpha_1 \lambda_1 + \alpha_2 \lambda_2 + \alpha_3 \lambda_3 + \alpha_4 \lambda_4, \\
 \text{s.t. } \lambda_{1,2} = & |j_m(2x_a, 2y_a) - j_m(2x_a + 1, 2y_a \pm 1)| + |j_n(2x_a, 2y_a) - j_n(2x_a + 1, 2y_a \pm 1)|, \\
 \lambda_{3,4} = & |j_m(2x_a, 2y_a) - j_m(2x_a - 1, 2y_a \pm 1)| + |j_n(2x_a, 2y_a) - j_n(2x_a - 1, 2y_a \pm 1)|, \quad (14)
 \end{aligned}$$

where $\alpha_i (i = 1 \dots 4)$ is a weight coefficient. The second energy term AM_e can be computed using Eq.15.

$$AM_e(2x_a + 1, 2y_a + 1) = |j_n(2x_a + 1, 2y_a + 1)| - |j_m(2x_a + 1, 2y_a + 1)|. \quad (15)$$

The third energy term AM_s can be computed using Eq.16.

$$\begin{aligned}
 AM_s(2x_a + 1, 2y_a + 1) &= \frac{-T^2Z + 2TUB - TZ^2}{T^2 + U^2} AP(2x_a + 1, 2y_a + 1), \\
 s.t. \quad T &= j_p(2x_a + 1, 2y_a + 1) = \frac{1}{2}(AP(2x_a, 2y_a) - AP(2x_a + 2, 2y_a + 2)), \\
 U &= j_q(2x_a + 1, 2y_a + 1) = \frac{1}{2}(AP(2x_a, 2y_a + 2) - AP(2x_a + 2, 2y_a)), \\
 B &= j_r(2x_a + 1, 2y_a + 1) = \frac{1}{2}(AP(2x_a + 1, 2y_a - 1) + AP(2x_a + 1, 2y_a + 3)) \\
 &\quad - \frac{1}{2}(AP(2x_a - 1, 2y_a + 1) + AP(2x_a + 3, 2y_a + 1)), \\
 Z &= j_m(2x_a + 1, 2y_a + 1).
 \end{aligned} \tag{16}$$

4.3. Seam carving operation

Aiming to achieve regional energy protection, we introduce definitions for the horizontal seam and the vertical seam to PDR, which are shown as follows:

The horizontal seam $h_{DI_t}^y$ of $IDI_{\{t\}}^O$ is given by

$$h_{DI_t}^y = \{h_{\mu}^y\}_{\mu=1}^{rH_{initial}} = \{(\mu_{DI_t}, y_{DI_t}(\mu))\}_{\mu=1}^{rH_{initial}}, s.t. \forall \mu, |y_{DI_t}(\mu) - y_{DI_t}(\mu-1)| \leq 1, \tag{17}$$

and the vertical seam $v_{DI_t}^y$ of $IDI_{\{t\}}^O$ is given by

$$v_{DI_t}^x = \{v_{\phi}^x\}_{\phi=1}^{rW_{initial}} = \{(x_{DI_t}(\phi), \phi_{DI_t})\}_{\phi=1}^{rW_{initial}}, s.t. \forall \phi, |x_{DI_t}(\phi) - x_{DI_t}(\phi-1)| \leq 1. \tag{18}$$

On this basis, we redefine the energy function (see Eq.19) to preserve the pixels in the attention region.

$$e_{HoG}(IDI_{\{t\}}^O) = \begin{cases} \max(H(IDI_{\{t\}}^O(x, y))), \\ \text{if } (AR_L \leq x \leq AR_R) \text{ and } (AR_B \leq y \leq AR_T), \\ \frac{|\frac{\partial}{\partial x} IDI_{\{t\}}^O| + |\frac{\partial}{\partial y} IDI_{\{t\}}^O|}{\max(H(IDI_{\{t\}}^O(x, y)))}, \\ \text{otherwise.} \end{cases} \tag{19}$$

Where $H(IDI_{\{t\}}^O(x, y))$ denotes the oriented gradient histogram of the pixels in $IDI_{\{t\}}^O$.

Fig.4 shows the examples of seam removal. We used red curves to denote the removed seams from interpolated images and from this figure, it can be seen that image content in the attention region can be protected on the premise of ensuring a global visual effect.

5. Experimental results and discussion

In this section we compare our method with previous work that studied the resizing of dynamic image sequences, and we demonstrate the effectiveness of our method. In all our experiments, we used the MATLAB platform and a PC with a 2.60 GHz Intel(R) Pentium(R) Dual-Core CPU processor with 1.96 GB of main memory.

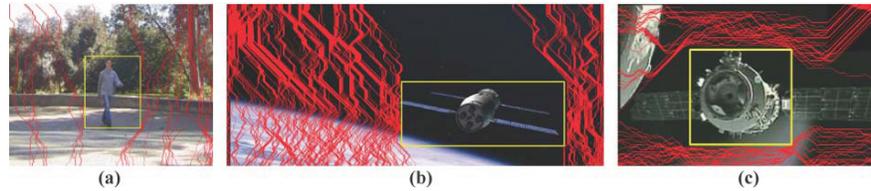


Fig. 4. Seam removal: finding and removing the seams in the general region

First, we demonstrate the performance of our proposed partition interpolation model. The dynamic image sequences used in Figures 5 and 6 are publically available video sequences (obtained from <http://www.wisdom.weizmann.ac.il/vision/SpaceTimeActions.html>). Fig.5(a) shows the original frame in the sequence of the walking man, and Fig.5(b) and 5(c) show resizing results from 180×144 to 345×287 using our method and scaling respectively. Similarly, Fig.5(d) and 5(e) are the resizing results from 180×144 to 700×575 . In the same manner, Fig.6 shows the resizing examples of jumping. In these figures, we use a yellow-bordered region to mark attention regions. The scaling method appears to introduce jagged effects, however, our method generates clearer image details, which results in a better resizing quality for the attention region.

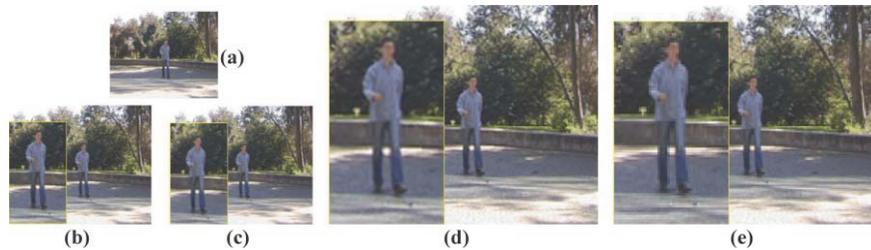


Fig. 5. Attention region comparison of resizing results of a walking man. (a) the original frame(180×144), (b) and (c) show results (345×287) using PDR and scaling, (d) and (e) show results (700×575) using PDR and scaling

To measure the resizing quality of the attention region in the dynamic sequences, five evaluation indicators (average gradient (AG), edge intensity (EI), information entropy (IE), spatial frequency (SF), and image definition(ID)) are adopted and the quantitative results are shown in Table 1. This table shows that the results using our method are higher than those of the scaling method, proving that our method can effectively improve the resizing quality of attention region.

We evaluate the performance of PDR by observing the resizing results. Fig.7 shows the comparison results of the sequence of a walking man with a dog us-

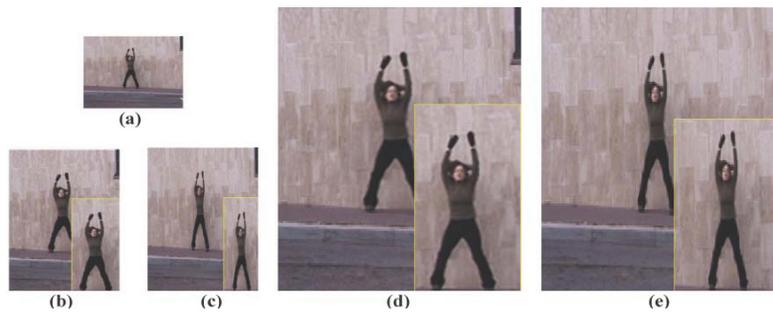


Fig. 6. Attention region comparison of resizing results of a jumping woman. (a) the original frame(180×144), (b) and (c) show results(200×287) using PDR and scaling. (d) and (e) show results (440×575) using PDR and scaling

Table 1. Quality comparison of the attention region

Test sequence	target size	Method	AG	EI	IE	SF	ID
A walking man	345 × 287	Ours	10.80	100.64	7.27	33.10	10.51
A walking man	345 × 287	Scaling	6.91	72.49	7.23	17.92	8.09
A walking man	700 × 575	Ours	5.79	55.90	7.30	22.29	5.46
A walking man	700 × 575	Scaling	3.99	43.71	7.29	12.40	4.55
A jumping woman	200 × 287	Ours	8.57	79.28	7.25	29.77	8.20
A jumping woman	200 × 287	Scaling	4.83	51.17	7.24	14.40	5.45
A jumping woman	440 × 575	Ours	4.51	42.87	7.26	20.12	4.24
A jumping woman	440 × 575	Scaling	2.82	31.03	7.24	10.43	3.20

ing four methods, including scaling, the best cropping, seam carving, and our method. Fig.7(b) shows that using scaling method, the walking man and his dog all become vaguer than before. Fig.7(c) shows that using the best result from the cropping method, the walking man is only partly displayed, resulting in the original information becoming missing. Fig.7(d) shows that using seam carving, the prominent part of this image sequence becomes smaller than before, indicating that the seam carving method is not suitable for image enlarging. Fig.7(e) shows that our method clearly displays the prominent objects from the original frames and ensures a global visual effect when the image sequence resolutions are changed. Similarly, resizing results of the sequence of a man with waist bent are shown in Fig.8.

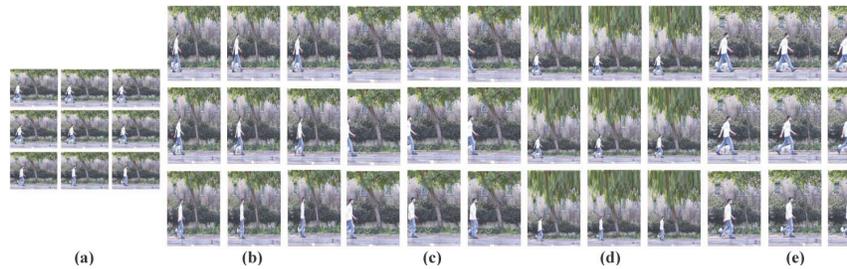


Fig. 7. Comparison results of a walking man with a dog for the four resizing methods when the resolution is resized from 180×144 to 200×288 : (a) the original frames, (b) scaling, (c) the best cropping, (d) seam carving, and (e) our algorithm

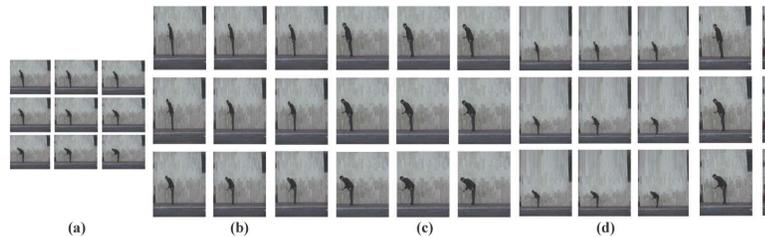


Fig. 8. Comparison results of a man with waist bent for the four resizing methods when the resolution is resized from 180×144 to 220×277 : (a) the original frames, (b) scaling, (c) the best cropping, (d) seam carving, and (e) our algorithm

Fig.9 shows the resizing comparison results of space image sequences. In this figure it is quite clear that the proposed algorithm shows comparatively

better performance in terms of visual quality. For example, the airship's flight action can be distinctly displayed in Fig.9(e).

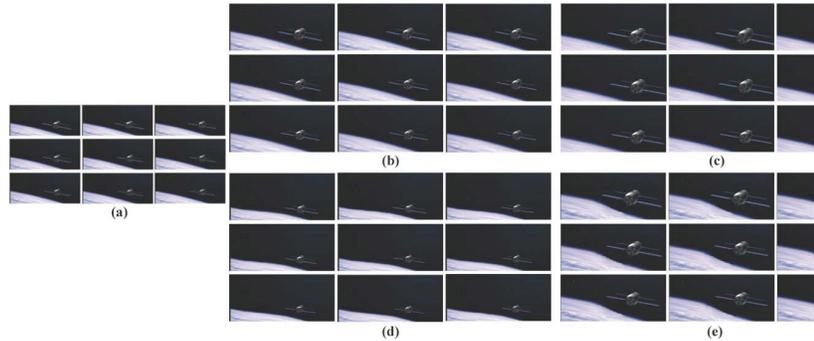


Fig.9. Comparison results of a space image sequence for the four resizing methods when the resolution is resized from 592×256 to 884×392 : (a) the original frames, (b) scaling, (c) the best cropping, (d) seam carving, and (e) our algorithm

In Fig.10, for measurements of the resizing quality of the above image sequences, we show the average gradient, edge intensity, and information entropy comparison results. The original images of Fig.10(a), Fig.10(d) and Fig.10(g) are shown in Fig.7(a). Similarly, the original images of Fig.10(b), Fig.10(e) and Fig.10(h) are shown in Fig.8(a), and the original images of Fig.10(c), Fig.10(f) and Fig.10(i) are shown in Fig.9(a). In Fig.10, the cyan curve stands for scaling, the blue curve for the best cropping, the green curve for seam carving, and the red curve for our method. It is obvious that the red curve is always above the others, which shows that our algorithm can generally achieve the highest AG, EI, and IE. This indicates that our algorithm obtains an encouraging performance in image quantization measure.

Table 2 summarizes the average indicator values for the three image sequences corresponding to Fig.7, Fig.8 and Fig.9. This table shows that our algorithm always obtains the highest average values. Because of the encouraging performance in terms of image visualization and quantitative quality assessment, our proposed algorithm is very competitive.

6. Conclusions

In this paper, we propose a perception-driven resizing method of dynamic image sequences, which can acquire high quality resizing results in accordance with human visual perception. Our method stresses the most important image content perceived by human beings, i.e., attention regions, and improves their resizing quality. The implementation process of our method involves three

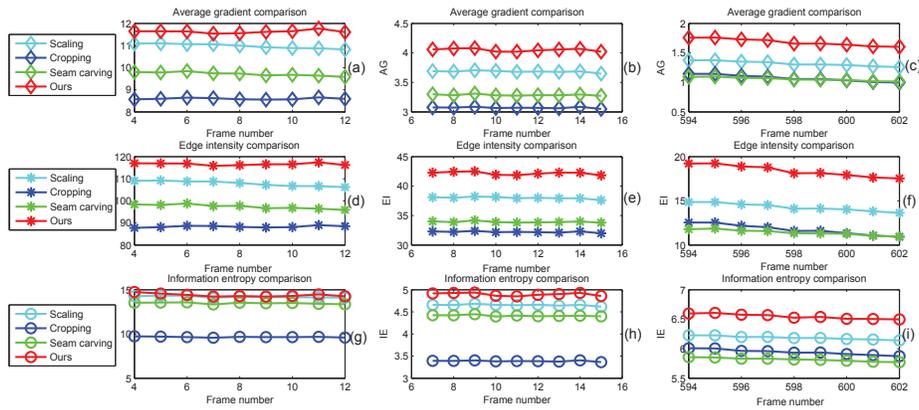


Fig. 10. Comparison of indicators using four resizing methods for a walking man with a dog, a man with waist bent and space image respectively. Subplots (a)-(c) are the average gradient comparison results, (d)-(f) are the edge intensity comparison results, and (g)-(i) are information entropy comparison results

Table 2. Comparison of the average indicator values for the four image resizing methods

Test sequence	Method	AG	EI	IE
A walking man with a dog	Scaling	10.98	107.85	14.20
A walking man with a dog	The best cropping	8.60	88.33	9.66
A walking man with a dog	Seam carving	9.72	97.39	13.48
A walking man with a dog	Ours	11.63	116.61	14.39
A man with waist bent	Scaling	3.68	37.99	4.65
A man with waist bent	The best cropping	3.07	32.22	3.39
A man with waist bent	Seam carving	3.29	33.92	4.41
A man with waist bent	Ours	4.05	42.17	4.89
Space image Sequence	Scaling	1.32	14.29	6.19
Space image Sequence	The best cropping	1.07	11.76	5.94
Space image Sequence	Seam carving	1.06	11.42	5.82
Space image Sequence	Ours	1.68	18.39	6.55

steps: attention region establishment using a coarse-to-fine detection approach based on a feature descriptor, the construction of a partition interpolation model based on real-time artifact-free image upscaling, and a seam carving operation using a region energy protection approach. Our experiments demonstrate that our proposed method can obtain more satisfactory results than the representative algorithms previously used in image resizing. In the future, we will improve the perception ability of our method by adding the salient detection technique to our resizing framework.

Acknowledgments. This work is supported by the National Basic Research Program of China (973 Program) 2012CB821200 (2012CB821206), the National Natural Science Foundation of China (No.91024001, No.61070142) and the Beijing Natural Science Foundation (No.4111002).

References

1. Ambai, M., Yoshida, Y.: Card: Compact and real-time descriptors. In: Proceedings of 2011 International Conference on Computer Vision. pp. 97–104. Barcelona, Spain (2011)
2. Avidan, S., Shamir, A.: Seam carving for content-aware image resizing. *ACM Transactions on Graphics* 26(3), Article 10 (2007)
3. Chen, L.Q., Xie, X., Fan, X., Ma, W.Y., Zhang, H.J., Zhou, H.: A visual attention model for adapting images on small displays. *Multimedia Systems* 9(4), 353–364 (2003)
4. Fang, F., Boyaci, H., Kersten, D.: Border ownership selectivity in human early visual cortex and its modulation by attention. *Journal of Neuroscience* 29(2), 460–465 (2009)
5. Giachetti, A., Asuni, N.: Real-time artifact-free image upscaling. *IEEE Transactions on Image Processing* 20(10), 2760–2768 (2011)
6. Hou, Z., Zhang, Q.: Interpolation algorithm for recovering the missing phase values in 3d measurement. *Optik* 121(14), 1324–1329 (2010)
7. Huang, F.C., Huang, S.Y., Ker, J.W., Chen, Y.C.: High-performance sift hardware accelerator for real-time image feature extraction. *IEEE Transactions on Circuits and Systems for Video Technology* 22(3), 340–351 (2012)
8. Kumar, M., Conger, D.D., Miller, R.L., Luo, J., Radha, H.: A distortion-sensitive seam carving algorithm for content-aware image resizing. *Signal Processing Systems* 65(2), 159–169 (2011)
9. Liu, H.: Generative 3d images in a visual evolutionary computing system. *Computer Science and Information Systems* 7(1), 111–125 (2010)
10. Lucilio, C.G., Gonzalo, V.S.F., Pablo, C.d.l.H., Carlos, A.L.: A markov random field approach for topology-preserving registration: Application to object-based tomographic image interpolation. *IEEE Transactions on Image Processing* 21(4), 2047–2061 (2012)
11. Mayhew, S.D., Li, S., Kourtzi, Z.: Learning acts on distinct processes for visual form perception in the human brain. *Journal of Neuroscience* 32(3), 775–786 (2012)
12. Ni, K.S., Nguyen, T.Q.: An adaptable k-nearest neighbors algorithm for mmse image interpolation. *IEEE Transactions on Image Processing* 18(9), 1976–1987 (2009)

13. Paltoglou, A.E., Neri, P.: Attentional control of sensory tuning in human visual perception. *Journal of Neurophysiology* 107(5), 1260–1274 (2012)
14. Pan, M.s., Yang, X.l., Tang, J.t.: Research on interpolation methods in medical image processing. *Journal of Medical Systems* 36(2), 777–807 (2012)
15. Pedro, N., Ricardo, V.M., Antonio, B.: Visual odometry based on structural matching of local invariant features using stereo camera sensor. *Sensors* 11(7), 7262–7284 (2011)
16. Wang, Y.S., Hsiao, J.H., Sorkine, O., Lee, T.Y.: Scalable and coherent video resizing with per-frame optimization. *ACM Transactions on Graphics* 30(4), Article 4 (2011)
17. Wang, Y.S., Tai, C.L., Sorkine, O., Lee, T.Y.: Optimized scale-and-stretch for image resizing. *ACM Transactions on Graphics* 27(5), Article 118 (2008)
18. Zhang, G.X., Cheng, M.M., Hu, S.M., Martin, R.R.: A shape-preserving approach to image resizing. *Computer Graphics Forum* 28(7), 1897–1906 (2009)

Lingling Zi is a Ph.D. candidate in School of Computer Science Beijing University of Posts and Telecommunications. Also she is Lecturer in School of Electronic and Information Engineering Liaoning Technical University. She was born in 1981 and obtained her Master Degree in 2007. Her research interests include multimedia systems and intelligent information systems.

Junping Du is a full professor and Ph.D. tutor with the School of Computer Science and Technology, Beijing University of Posts and Telecommunications. She received the Ph.D. degree in computer science from the University of Science and Technology, then held a post-doctoral fellowship with the Department of Computer Science at Tsinghua University, Beijing, China. She was a visiting professor with the Department of Computer Science at Aarhus University, Denmark, from September 1996 until September 1997. Now, she is the Director of the Computer Applications Center in Beijing University of Posts and Telecommunications University. Her current research interests include Multimedia systems, artificial intelligence and intelligent management systems.

Lisha Hou is a master degree candidate at Beijing University of Posts and Telecommunications, China. Her research interests include multimedia systems and intelligent information systems.

Xiangda Sun is a master degree candidate at Beijing University of Posts and Telecommunications, China. His research interests include multimedia systems and intelligent information systems.

Jangmyung Lee is a Professor with the Intelligent Robot Laboratory, Pusan National University, Busan, Korea. He received the B.S. and M.S. degrees in electronic engineering from Seoul National University, Seoul, Korea, in 1980 and 1982, respectively, and the Ph.D. degree in computer engineering from the University of Southern California, Los Angeles, in 1990. His current research interests include intelligent robotic systems, ubiquitous ports, and intelligent sensors.

Received: July 31, 2012; Accepted: December 6, 2012.

