SCAN: A Multi-Operator Image Retargeting Scheme

Po-Chyi Su, Zi-Hao Xiang, Hao-Wei Wu
Dept. of Computer Science and Information Engineering
National Central University
Jhongli, Taiwan
pochyisu@csie.ncu.edu.tw

ABSTRACT — This research presents a multi-operator retargeting mechanism termed “SCAN”, in which seam carving, cropping, adding seams and normalization (scaling) are applied on images in an automatic manner. The content-based cropping will first be used to remove insignificant portions on sides. Then a new seam carving algorithm based on both the global saliency and local saliency is proposed to rid of the pixels in the middle of the image. Efficiency is the major advantage of this seam carving algorithm. When the background is not complex, some seams may be inserted in a similar way as the proposed seam carving procedures to make the aspect ratio closer to the target one. Finally, the image is scaled or normalized directly. Experimental results will demonstrate the feasibility and advantages of the proposed method.

KEYWORDS—retargeting; seam carving; cropping; scaling; saliency; multi-operator

I. INTRODUCTION

Digital imagery data are ubiquitous these days and various kinds of facilities to display the visual content have been developed to satisfy the needs of related applications. However, digital image/video files usually have a fixed resolution. In order to adapt to a certain display, directly scaling the images or video frames to the target resolution is a necessary step but may introduce noticeable or unpleasant distortion. Imposing additional boundaries on the sides of images/frames may not always be acceptable. Therefore, content-based retargeting mechanisms to modify the resolution in an adaptive way have drawn a lot of interest in recent years. The basic idea is to process the regions of an image in different ways such that the non-uniform scaling can be achieved to make the resolution fit the target one.

The methodologies of retargeting can be roughly classified into four categories: content-based cropping, warping, seam carving and the use of multi-operator. The content-based cropping can determine the regions of interest in an image or video frame according to the content analysis and remove insignificant sides [3][9][10][11]. The methods of warping impose a set of grids on the image and adjust the grids non-uniformly or in varying ways to achieve the global optimization based on a distortion measurement [4][18]. The gradient, visual saliency, face detection and even motion, etc. are used to evaluate the importance of grids and then different kinds of warping can be applied accordingly. Avidan and Shamir [13] proposed the ideas of seam carving for still images. By successively removing the seams with the minimum energy or saliency in the image, the important areas will be better preserved. Most of the changes are applied on uniform areas so that less visual distortion may be introduced. [5] is an improved seam carving algorithm and extends the proposed technique to videos. The so-called looking-forward energy helps to reduce the discontinuous seams. [8] proposed a discontinuous seam carving method, which helps to prevent the moving objects in consecutive frames from being affected by the seam carving process. [2] proposed the idea of stream carving to remove multiple seams at the same time. The resulting holes will be fixed by image inpainting. The face and line detections are employed to further reduce the possible distortion.

The methodology of multi-operator methods combine different approaches to pursue more satisfactory results. [14] employed Scale Invariant Feature Transform (SIFT) to evaluate the difference between the quality before and the quality after retargeting so that a more suitable termination of seam carving can be determined. [6] defined the so-called Bi-Directional Warping (BDW) to evaluate the quality. The method chooses one operator from cropping, seam carving and scaling based on the evaluation to achieve a good retargeting result. The drawback is the speed since the quality evaluation has to be applied in each step. [15] proposed “Accumulated Energy Seam Carving” to improve the single-direction modification so that the resulting image can maintain a closer resolution to the target. [7] combined the seam carving and scaling based on the difference measurement of images and dominating color descriptors. [16] proposed an energy function to speed up the retargeting process but the complexity will be significantly increased if more operators are used.

This research aims at developing a multi-operator retargeting scheme called “SCAN”, which employs cropping, seam carving, adding seams and direct scaling/normalization to obtain the resulting image that has a closer aspect ratio to the target one so that the important content can be better preserved. Based on a saliency map, the cropping is first applied to remove insignificant sides of images. The process is efficient so the time spent on the computationally expensive seam carving can be reduced because of the smaller image size. The seam carving according to both the global saliency and local saliency is designed to remove pixels in the middle of the image without affecting the content match. Some seams can be further added in the uniform background. Finally, the image is normalized to the target resolution. The paper is organized as follows. Sec. II will detail the proposed algorithm. Some experimental results are shown in Sec. III, followed by the conclusion and future work in Sec. IV.
II. THE PROPOSED SCHEME

A. Visual Saliency

The method proposed by Montabone et al. [12] is used to calculate the visual saliency, which will help to determine the locations where we should stop the cropping and to find a suitable seam to carve. This method makes use of the property that the human perceive the visual information by the on-center and off-center ganglion cells in eyes; on-center ganglion cells respond to bright areas surrounded by a dark background while off-center ganglion cells respond to dark areas surrounded by bright areas. The saliency map is then calculated as follows. After converting the image into grayscale values, the 3 by 3 Gaussian filter is employed to smooth the grayscale image twice and then the integral image is calculated. The integral image \( G(m,n) \) represents the sum of the values above and to the left of \((m, n)\), i.e.,

\[
G(m,n) = \sum_{m' \leq m, n' \leq n} g(m',n'),
\]

(1)

where \( g \) is the filtered pixel. The sum of any rectangular area defined by two points, \((m1,n1)\) and \((m2,n2)\), can be calculated by \( G(m,n) \) in constant time, i.e.,

\[
\text{Sum}(m1,n1,m2,n2) = G(m2,n2) - G(m1,n2) - G(m2,n1) + G(m1,n1)
\]

(2)

Then, the on-center and off-center differences with two surrounding values, 3 and 7, and three scales, 2, 3 and 4 can be computed efficiently to acquire six intensity submaps: \(3 \cdot 2^2, 3 \cdot 2^3, 3 \cdot 2^4, 7 \cdot 2^2, 7 \cdot 2^3\) and \(7 \cdot 2^4\) so the set of scale \( q \) is \{12, 24, 48, 28, 56, 112\}. The surrounding pixels are defined as

\[
surround(m,n,q) = \frac{\text{Sum}(m,q,n,q,m+q,n+q,q) - g(m,n)}{(2q+1)^2-1}
\]

(3)

The intensity submaps are then calculated by

\[
\text{Int}_{(on,q)}(m,n) = \max(g(m,n) - \text{surround}(m,n,q),0),
\]

(4)

\[
\text{Int}_{(off,q)}(m,n) = \max(\text{surround}(m,n,q) - g(m,n),0),
\]

(5)

The saliency image \( P(m,n) \) is formed by

\[
P(m,n) = \max(\sum_q \text{Int}_{(on,q)}(m,n), \sum_q \text{Int}_{(off,q)}(m,n))
\]

(6)

Fig. 1 shows one example. The butterfly is apparently the main object of the image and higher saliency values are assigned correctly to the area where the butterfly is located.

\[\]

B. Automatic Cropping

In the following, we will show the adjustment of the width as an example to illustrate the retargeting procedure. The automatic cropping is the first step, which will be applied on the background on both sides of the image until a meaningful foreground is touched. The foreground will thus be extracted according to the saliency map. To be more specific, we calculate the average values of the saliency pixels in 8 by 8 blocks. If the average value is larger than the threshold \( \epsilon \) (set as 72), then this 8 by 8 block is viewed as the foreground block. All the pixels in this block will be set as 255 and as 0 otherwise. Then, the foreground image will be processed by the four-connected-component labeling and we will decide whether a component is important based on the area or the number of foreground blocks. Next, we employ the property of “Rule of Third” in digital photography [1], which states that the central 1/3 of the image is usually the region of interest. Two constraints are imposed:

- The foreground objects located around the two sides of the image are seldom important so they should be viewed as the background and can be cropped off. We therefore trace the significant components starting from the two sides toward the center. When there are three consecutive vertical background blocks become foreground during the tracing, an important object has been encountered and should be kept.

- Since the top/bottom 1/3 of images may not be important according to the property of “Rule of Third”, we conservatively ignore top/bottom 1/9 of the image, which will not be traced to find the foreground area.

Fig. 2 shows an example. The white pixels in Fig. 2(b) are viewed as the foreground area while the gray pixels are removed because the associated area is just too small to be considered as the foreground. The small red blocks in Fig. 2(b) generated during the tracing help to determine the locations of edges so the cropping can be applied efficiently.
C. Seam Carving

Basically, we modify the seam carving scheme in [13] to remove the pixels in the middle of an image. Given an image with the resolution \( M \times N \). The vertical seam is

\[
S^* = \{x(i), i\}_{i=1}^{N}, |x(i) - x(i-1)| \leq 1. \tag{7}
\]

where \( x \) is in the range of 1~\( M \) so there are at most \( M \) seams. The path of seam can be expressed as

\[
I_s = [x, (x(i), i)]_{i=1}^{N} \tag{8}
\]

When a seam is carved, only a small portion in an image will be affected. The pixels on its right will be shifted to left by one sample to fill in the carved seam. The selection of seams to carve is based on the saliency. Given the saliency of a seam is \( E(I_s) \), we will select the best seam \( S^* \) such that the penalty from such a seam carving will be the minimum,

\[
S^* = \min_s \{E(I_s)\} \tag{9}
\]

The calculation is performed via the dynamic programming. From the second row to the last row of an image, all the saliency values are accumulated by

\[
Q(i, j) = P(i, j) + \min \{Q(i-1, j-1), Q(i-1, j), Q(i-1, j+1)\} \tag{10}
\]

Again, \( P(i,j) \) is the saliency of the point. After calculating all the saliency values according to (10), we trace back from the last row with the minimal saliency and then select the recorded path. The resulting seam is chosen and removed.

One unique idea in the proposed seam carving scheme is that we consider not only the global saliency but also the local saliency, which will significantly reduce the negative effects of seam carving. We claim that the local saliency has to be maintained in a reasonable value to avoid generating large variations. To be more specific, although the global saliency of a seam is not large, some large local saliency values may indicate that some obvious distortions have happened in the image. The sum of four continuous pixels in a seam is considered as the local saliency. If it exceeds a threshold, it is likely that the important portion will be affected. We need to calculate the local energy of every pixel to check whether a connected direction of a seam is allowed. Since the local energy involves four continuous pixels in the current implementation, there are 27 kinds of paths as shown in Fig. 3 and 27 bits are used to record whether a path is allowed to be removed. The calculation of 27 bits in each pixel is as follows:

- Each pixel is added to the \( R \) values of its upper-left, upper and upper-right pixels. Nine values are computed for each pixel as

\[
\begin{align*}
R_{UL}(i,j) &= P(i,j) + P(i-1,j-1) \\
R_{UP}(i,j) &= P(i,j) + P(i-1,j) \\
R_{UR}(i,j) &= P(i,j) + P(i-1,j+1)
\end{align*}
\]

- Each pixel is added to the \( U \) values of its upper-left, upper and upper-right pixel to create 27 values for each pixel.

\[
\begin{align*}
U_1(i,j) &= P(i,j) + R_{UL}(i-1,j-1) \\
U_2(i,j) &= P(i,j) + R_{UP}(i-1,j-1) \\
U_3(i,j) &= P(i,j) + R_{UR}(i-1,j-1) \\
U_4(i,j) &= P(i,j) + R_{UL}(i-1,j) \\
U_5(i,j) &= P(i,j) + R_{UP}(i-1,j) \\
U_6(i,j) &= P(i,j) + R_{UR}(i-1,j+1) \\
U_7(i,j) &= P(i,j) + R_{UL}(i-1,j+1) \\
U_8(i,j) &= P(i,j) + R_{UP}(i-1,j+1) \\
U_9(i,j) &= P(i,j) + R_{UR}(i-1,j+1)
\end{align*}
\]

- Each pixel is added to the \( U \) values of its upper-left, upper and upper-right pixel to create 27 values for each pixel.

- Each pixel is added to the \( U \) values of its upper-left, upper and upper-right pixel to create 27 values for each pixel.

After carving one seam, we need to update the 27 bits of each pixel. Because we only know whether the path formed by the 4 continuous pixels can be removed or not, we should check the 27 bits of pixel along the path to make sure a path is allowed from the bottom to the top of the image. The updating process is as follows:

- For each pixel, we copy the 9 bits that are related to the upper-left pixels 3 times to create 27 bits. Then, apply the AND operation with the 27 bits of upper-left pixel. If any one of three values that are copied is nonzero, we set the bit as 1, and set it as 0 otherwise. After the processing, each pixel will have 27 bits that can represent which path of seams can be removed.

After carving one seam, we need to update the 27 bits of each pixel. Because we only know whether the path formed by the 4 continuous pixels can be removed or not, we should check the 27 bits of pixel along the path to make sure a path is allowed from the bottom to the top of the image. The updating process is as follows:

- For each pixel, we copy the 9 bits that are related to the upper-left pixels 3 times to create 27 bits. Then, apply the AND operation with the 27 bits of upper-left pixel. If any one of three values that are copied is nonzero, we set the bit as 1, and set it as 0 otherwise.

- Apply the same calculation on the upper and upper-right pixels.

After updating the 27 bits, we can ensure that, if the bit representing a direction is 1, the pixel can be removed or included in a seam that can be carved. Again, the trace is going from the bottom to the top of image. If the 27 bits of the bottom pixel is not 0, then there is at least one path of
seam that can be removed since the local energy of this seam is surely smaller than the threshold.

Fig. 3. The 27 examined paths in considered local saliency.

The seam carving is carried out as follows. First, we calculate the minimum value of seam path for each bottom pixel whose 27 bits are not all equal to 0. Then we choose the one with the minimum energy and record the seam. After determining a path of seam, we set the 27 bits of the associated pixels as 0 and update the 27 bits of three pixels below. After updating the 27 bits of the pixels below, if the values of 27 bits have been changed, we should update the three pixels below the changed pixels and continue. This process only updates the changed pixels to save some computing time. The seam carving will stop when all the 27 bits of the bottom pixels are 0, which means that there is no seam that can be removed anymore. Fig. 4 shows an example of the proposed seam carving.

Fig. 4. (a) The seams that can be removed and (b) the resulting image.

D. Adding Seams

For the process of seam adding, the procedure is very similar to the propose seam carving algorithm. Basically, we determine the seam based on the global and local saliency and then copy the seams. As long as the threshold is set lower, the copying is similar to the interpolation since the considered areas are usually flat. Fig. 5 shows an example, where the red lines on Fig. 5(a) are the locations of seams that can be duplicated and Fig. 5(b) is the resulting image.

Fig. 5. An example of adding horizontal seams: (a) the determination of locations of seams and (b) the resulting image.

E. Normalization

The final step is the normalization, which simply scales the image directly to the target resolution. More advanced interpolation surely helps to maintain the image quality. We can see that the above mentioned cropping, seam carving and adding seams basically try to preserve the aspect ratio of the image so that the normalization will not cause serious distortions.

III. EXPERIMENTAL RESULTS

We compare the results with Seam Carving [5], SNS [18] and Multi-Op [6]. The test images and target resolutions are set according to [7]. The original images are shown in Fig. 6. Figs. 7 to 12 demonstrate the comparisons. We also list the reduced numbers of columns of each step by the proposed method. The proposed scheme tries to maintain the aspect ratio to preserve the content better, such as the cyclist in Fig. 7, the vehicle in Fig. 8, the butterfly in Fig. 9, and the circle formed by the pencils in Fig. 10. The chairs are well maintained in Fig. 11 as the chair feet are easy to be cropped. In Fig. 12, the wings are preserved and some seams are added in the area of sky so that the eagle is similar to its original shape.

Fig. 6. Test images.

Fig. 7. Comparison of Bicycle1: (a) Seam Carving (b) SNS (c) Multi-Opt (d) Ours C:128 SC:25 SCL:73.
IV. CONCLUSION AND FUTURE WORK

This research presents a multi-operator retargeting scheme that combines the automatic cropping, seam carving/adding and scaling. The design principle is to reduce the computational load, as the saliency is calculated only once and used in all the operators. The experiments show that, the proposed method outperforms [6] in several images. The proposed method may still be improved by considering more advanced saliency detection, such as the inclusion of face detection, to avoid generating perceptible distortions when the human face occupies a larger area. Besides, the efficiency needs further enhancement as the seam operation is still the bottleneck of the execution. Finally, the algorithm can be extended to the video processing, e.g., on the key frame of a video segment without obvious camera motions.

ACKNOWLEDGMENT

This research is supported by the Ministry of Science and Technology in Taiwan, R.O.C., under Grant MOST 103-2221-E-008-080.

REFERENCES


