Structure-Preserving Image Retargeting With Compression Assessment and Adaptive Registration

R.Nishanthi, Pon L.T.Thai, K.John Peter

***Abstract*: A number of algorithms have been proposed for image retargeting with image content retained as much as possible. But, they usually suffer from some artifacts in the results, such as ridge or structure twists. In this paper, a structure and content preserving image retargeting technique is used that preserves the content and image structure as best as possible. The image content saliency is estimated from the structure of the content using saliency map. A block structure energy is used for structure preservation along x and y directions. Block structure energy uses top-down strategy to constrain the image structure uniformly. However, the flexibilities of retargeting are different for different images. To overcome this problem, compression assessment scheme is used by combining the entropies of image gradient magnitude and orientation distributions. Finally, adaptive registration algorithm is applied. Adaptive registration is used to increase the PSNR ratio. Thus, the resized image is produced to preserve the structure and image content as best as possible. The global image structure is preserved and structure distortions are avoided.**

***Index Terms*: Compressibility estimation, image retargeting, structure-preserving.**

I. INTRODUCTION

Image Processing is a form of signal processing. The input is an image. The output is an image or a set of characteristics or parameters related to image. Image retargeting means changing the size of the image in a content aware manner. The image can be viewed on a large number of display devices such as cell phones, iPods, television, computer etc. Each display device has its own unique aspect ratio. Retargeting the image to fit into the aspect ratio of the display device is called image retargeting.

Converting an image from lower resolution to higher resolution can also be termed as image retargeting. Artifacts in an image means the undesired alterations. Artifacts, ridges, structure twists are errors of an image. Ridges are the non uniform portions of an image.

Image aspect ratio describes how the width of the image in pixels is compared to the height of the image in pixels. A number of algorithms have been proposed to adapt image content to various display settings.

Different approaches may be suitable to retarget images with different content. In this project, the aim is to overcome the common problem of structure distortion. In Human Visual System (HVS), human eyes are very sensitive to certain object shapes, like circles or straight lines. Most of the previous works proposed so far did not consider the preservation of global image structure. Based upon the image saliency measure, most of the previous works constrain the retargeting process in a bottom-up strategy to preserve the structure continuity, which may not well preserve the global image structure.

In this paper, the image structure is preserved with a top-down strategy by using a block saliency map which is adaptive to the sizes of the structured objects. Saliency means the quality of an image. Top-down strategy is a step wise design. Top-down strategy means breaking down of a system into compositional subsystems. Each subsystem is refined in a greater detail. Image content saliency is estimated from the structure of the content. Block structure energy is used to deform the image structure in x and y direction.

Bounding boxes are used. Bounding box means the x-y coordinates of the lower left corner of image followed by x-y coordinates of upper right corner of the image. Compressibility rate for each image is assessed from its image gradient magnitude and orientation distribution. Image gradient is the change in the intensity or color of an image. It is used to extract the information from the image. Orientation distribution can be defined mathematically in any space, appropriate to continuous description of rotations.

The retargeting process is optimized based upon the compressibility assessment in both x and y directions. The resized image frames, therefore, meet the fundamental requirements: retaining salient content and preserving global structure. Different strategies are utilized such that the image content saliency and structure are well preserved.

Based upon the image saliency measure, most of the previous works constrain the retargeting process in a bottom-up strategy to preserve the structure continuity, which may not well preserve the global image structure. In this paper, the image structure is preserved with a top-down strategy by using a block saliency map which is adaptive to the sizes of the structured objects. Saliency means the quality of an image.

Manuscript received on July, 2012.

R.Nishanthi, Computer Science, Vins Christian College of Engineering, Anna University,Nagercoil, India,

Pon L.T.Thai, Computer Science, Vins Christian College of Engineering, Anna University,Nagercoil, India,

K.John Peter, Information Technology, Vins Christian College of Engineering, Anna University, Nagercoil, India.

Top-down strategy is a step wise design. Top-down strategy means breaking down of a system into compositional subsystems. Each subsystem is refined in a greater detail. Image content saliency is estimated from the structure of the content.

II. RELATED WORK

The existing methods largely fall into three categories: one applies the cropping or carving strategies to remove less important regions, one segments the image into foreground and background layers and scales each layer independently, and the third approach is based upon based upon adaptively warping the image image based upon the local image saliency. For the cropping based methods, the most important region is automatically detected and transmitted the to the mobile device.

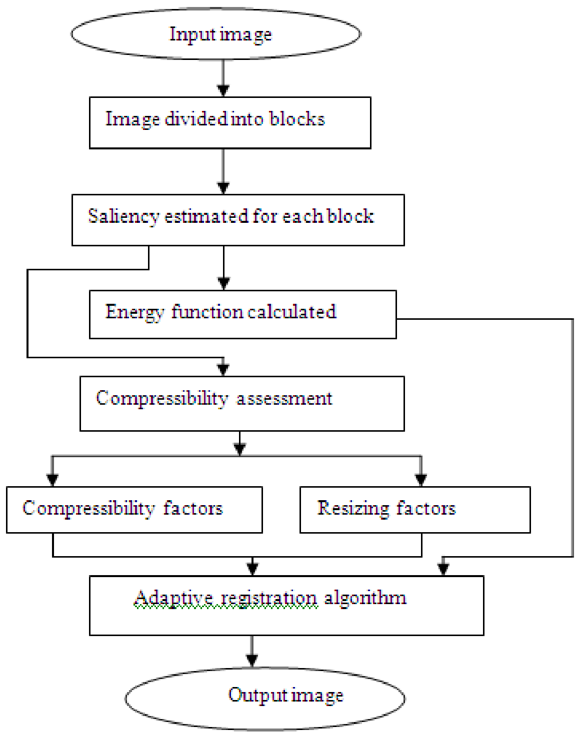
Cropping is applied to remove less important regions from the images which may discard a large amount of information and sometimes fail if the important features are located at distant parts in the image. To deal with this problem caused by cropping, Avidan and Shamir [2] proposed an interesting idea of incrementally removing or inserting regions, called seam carving. An 8-connected path of least importance pixels is incrementally removed or inserted to resize an image frame. Image adaptation, one of the essential problems in adaptive content delivery for universal access, has been actively explored for some time. Most existing approaches have focused on generic adaptation with a view to saving file size under constraints in client environment and have hardly paid attention to user perceptions of the adapted result.

Seam carving is an effective technique for image retargeting, although in some cases it cannot avoid the problem of producing artifacts or distorting content structures. Seams are computed as the optimal paths on a single image and are either removed or inserted from an image. This operator can be used for a variety of image manipulations including: aspect ratio change, image retargeting, content amplification and object removal. The operator can be easily integrated with various saliency measures, as well as user input, to guide the resizing process. Without cropping or carving the image content, image segmentation can provide an alternative way to separate the image regions of different importance. Setlur et al. segmented the image into several regions, scaled them independently according to the importance and then combined all the scaled regions. However, the results of these segmentation-based methods heavily depend upon the robustness and accuracy of the image segmentation results. In many cases, it is difficult to segment the image frame accurately and consistently.

III. PROPOSED SYSTEM

In Human Visual System (HVS), human eyes are very sensitive to certain object shapes, like circles or straight lines. However, most previous works proposed so far did not consider the preservation of global image structure. Based upon the image saliency measure, most of the previous works constrain the retargeting process in a bottom-up strategy to preserve the structure continuity, which may not well preserve the global image structure. We propose to preserve the image structure with a top-down strategy by using a block saliency map which is adaptive to the sizes of the structured objects. Compressibility rate for each image is assessed from its image gradient magnitude and orientation distribution.

A.System Design



B.Inter frame interaction

To retarget an image with size W_0 x H_0 to a new image with size W_r x H_r, traditionally the relatively unimportant area is removed/inserted to fit the image to the target resolution based upon the image saliency energy distribution. In many cases, however, it is not easy but important to retain the important content and also preserve the structure. In this paper, the problem of image resizing is dealt from a new perspective, i.e., retargeting the image aspect ratio. Based upon a warping based approach, novel block structure energy is introduced for structure preservation in the following subsection.

On the other hand, over compression is another reason that may distort the content structure. Therefore, how to well distribute the compression/stretching along and directions is essential to minimize the distortions. In addition, the optimal compressibility rate is estimated according to the image content. In this paper, the problem of image resizing is dealt from a new perspective, i.e., retargeting the image aspect ratio. Without considering the actual target size, first determine the optimal width and height that best fit the original image to the target aspect ratio and then uniformly scale it to the target image size.

Let S_x and S_y denote the resizing factors in x and y directions respectively. The linear relation of image resolutions between the input and retargeted images can be written as:

$$\frac{S_X.W_0}{S_Y.H_0} = \frac{W_r}{H_r} \qquad (1)$$

According to different image content, the best S_x and S_y might be different even though the target aspect ratio keeps unchanged. All these issues will be discussed in the following subsections. First of all, the content structure in an image is defined.

C. Regularization

To estimate the significance of an image, one can measure the magnitude of gradient as local energy or combine color, intensity and orientation for content saliency. Different high level image saliency measures, e.g., face detector, can also be included to improve the saliency measure. For general-purpose image retargeting, high-level image saliency measures are not used and simply the gradient and content saliency is used. First, the input image is decomposed into a set of multiscale feature maps which extract the information of color, intensity and orientation.

Most of the previous works resize the images according to the estimated energy, and also minimize the distortion of neighboring pixels, or quads. Based upon these bottom-up schemes, the distortion will accumulate to generate undesirable image structure distortion or noticeable artifacts. In the warping-based methods, most distortions are caused by non uniform warping on vulnerable objects. In this paper, the object structure is protected and the distortions are minimized with a global approach.

In order to extract the contours in colorful images, color tensor is applied to detect the salient edges. The extracted contour may connect to other object contours and spread over the whole image. Therefore, the extracted edges are decomposed into several pieces by simply cutting at the corners detected by the corner detector. The structure segments consisting of not enough pixels are treated as noise and then removed. In order to maintain the prominent structures, each structure piece should be considered as a single unit and protected.

Take a straight line for example; the pixels on the line segment should have the same slope after retargeting. Therefore, all pixels lie on the line are constrained to have the same slope. However, the slope (or position) of the line segment is unknown after retargeting and there are many prominent structure pieces of various shapes to be protected. The block structure energy is to be estimated. Therefore, all pixels lie on the line are constrained to have the same slope. However, the slope (or position) of the line The structure complexity is the combination of the image gradient magnitude and orientation.

There are many prominent structure pieces of various shapes to be protected. A simple and effective way to protect the extracted structure pieces is the block structure energy. The basic idea is that bounding boxes are used.

w= c.exp (- H_{all})

Where $H_{all} = (H_{\|g\|} + Hg)$

Where $H_{\|g\|}$ is the entropy of image gradient magnitudes and it stands for the complexity of gradient variations. H_g denotes the entropy of gradient orientations, which is related to the consistence of content structure, and is a constant used to rescale to the range between 0 and c. Note that H_{all} is positive related to the complexity of the image gradient magnitudes and orientations.

the extracted structure pieces to build block structure energy and all pixels inside the block energy should be stretched or compressed as uniformly as possible. The deformation is enforced to be uniform in each direction. The flexibility of each structure piece is defined separately for x and y directional retargeting.

For x-directional retargeting,the block structure energy values of all the pixels is defined as:

$$E_i^x(P_j)= \begin{cases} \frac{1}{e} \sum_{k\epsilon e} G\, x(P) & \text{if } P_j \epsilon B_i \\ 0 \end{cases} \qquad (2)$$

For x-directional retargeting,the block structure energy values of all the pixels is defined as:

$$E^y_{all}(P_j)=\max\left\{Gy(Pj), Esal(Pj), \max\{E_i^y(Pj)\}\right\}$$

(3)

D. Interpolation

Resizing is utilized in both directions. Images with different content may have different flexibility for retargeting. the compressibility rate is estimated in each direction and its definition is given as follows:

$$r_x=\frac{1}{W}\sum max\left\{(G_{\max\{G\}} - \max\{G(x,y)\}),0\right\}$$

$$r_y=\frac{1}{H}\sum max\left\{(G_{\max\{G\}} - \max\{G(x,y)\}),0\right\}$$

(4)

In addition, the resizing factor (S_X , S_y) is made close to the safe line defined by the estimated compressibility rate. For example, if the image width is changed to be half, compress along x direction and stretch in y direction to fit the target aspect ratio. Thus, the resizing factors are represented by (1- r_x , 1+ r_y). On the other hand, if image height is changed to be half of the original size, rotate the image by 90 degree before applying the algorithm and rotate it back after the retargeting.

Minimize $(S_x - 1)^2 + (S_y - 1)^2 + w.\ ((1 - r_x - S_x)^2 + (1 + r_y - S_y)^2)$

Subject to $S_x = \beta . S_y$

$S^*_x = \beta .\ S_y$

On the other hand, if image height is changed to be half of the original size rotate the image by 90 degree before applying the algorithm and rotate it back after the retargeting. w can be user defined or just a constant. Since it may depend upon image content, w is automatically determined by structure complexity estimation. If the content structure is clear, it weights heavier on the compressibility rate and vice versa.

Fig 1. Input image of low resolution

E. Adaptive Registration

Adaptive registration algorithm is used. Low resolution image is mapped to the uniform high resolution grid, and then the adaptive registration is used to form a high resolution image. Wiener filter is applied to reduce the effects of blurring and noise. PSNR ratio is increased. The relation of stretching/shrinking in x and y directions is defined by the resolution of source and target image: $S_x = \beta . S_y$. The optimal S^*_x and S^*_y corresponding to different image content can be determined.

If β<1, compress the image along x-direction and stretch along y-direction. If β>1, compress a little along y-direction and stretch a little along x-direction, and just rotate the image and do the same thing. Utilizing the image resizing in both x and y directions is better than that along one direction only. PSNR(Peak Signal to Noise Ratio) and MSE(Mean Square Error) values are estimated. PSNR is the ratio of the power of signal to the power of noise. MSE is the difference between the expected values and the true values obtained.

The bottom-up neighboring smoothness will accumulate the structure bending and distortion. In contrast to carving-based methods, the warping-based approach is used with the combination of compressibility and block structure energy for image retargeting. The formulation for the image retargeting problem is cast as solving a constrained linear system. To retarget a given image of size (W_0 x H_0) to size ($S_x\ W_0$ x $S_y\ H_0$) (This target size may not be equal to the exact target resolution (W_r x H_r)), the task is to recover the new coordinate (X_P, Y_P) of pixels p=(x,y) under three types of constraints. Take the calculation of X_P for instance. First, each pixel is supposed to be at a fixed distance from its left and right neighbors: $X_{x,y}-X_{x-1,y}=1$ and $X_{x+1,y}-X_{x,y}=1$.

Fig 2. Output image of high resolution

The second constraint is to map each pixel to a location similar to the one of its upper and lower neighbors: $X_{x,y}-X_{x,y+1}=0$. The third constraint fits the warped image to the dimensions of the resized image size: $X_{1,y}=1$ and $X_{w,y}=S_xW_0$. For content-aware image retargeting, an important pixel should be warped to a distance approximately one pixel from its neighbors while less important ones can be blended with their neighbors or be carved off.

Therefore, the first constraint mentioned previously should be weighted by the corresponding structure energy value E^x_{all} . Since the shrinking and stretching of pixels inside each block structure energy should be as uniform as possible, the maximum importance value 1 can be assigned to the second constraint (smoothness term) if $(x,y) \in B$ where $B = \bigcup_i B_i$ is the set of all pixels covered by the structure blocks. The equations in the constrained linear system are written as

$$E^x_{all}(x,y)(X_{x,y}-X_{x-1,y}) = E^x_{all}(x,y)$$
$$E^x_{all}(x,y)(X_{x+1,y}-X_{x,y}) = E^x_{all}(x,y)$$
$$E^x_{ud}(x,y)(X_{x,y}-X_{x,y+1}) = 0$$
$$X_{1,y}=1,\ X_{w,y}= S_x.W_0 \qquad (5)$$

Where

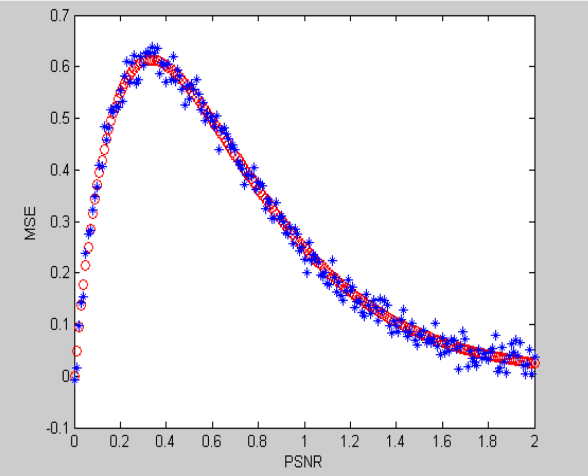
$$E^x_{ud}(x,y) = \begin{cases} 1, & \text{if } (x,y) \in \mathbf{B} \\ E^x_{all}(x,y), & \text{if } (x,y) \notin \mathbf{B} \end{cases}$$

The weight E^{x}_{ud} denotes the importance between up-and-down pixels which controls the smoothness over pixels. All the equations in (5) form an over-determined constrained sparse linear system. The optimized new coordinates of the pixels can be obtained by minimizing the sum of squared errors of the previous equations, which is equivalent to finding the least-squares solution of the sparse linear system: $Ax \approx b$. By solving the linear system formed by (5), the coordinates of all pixels in the target domain can be obtained which are constrained by the target width. On the other hand, the coordinates must be computed which are constrained by the target height to exactly determine the target location in the resized image.

IV. RESULT

In the previous methods, each method has its unique features and advantages. In this paper, the aim is to overcome the common problem in most of the previous work, i.e., structure distortion. In human visual system (HVS), human eyes are very sensitive to certain object shapes, like circles or straight lines. However, most previous works proposed so far did not consider the preservation of global image structure. Based upon the image saliency measure, most of the previous works constrain the retargeting process in a bottom-up strategy to preserve the structure continuity, which may not well preserve the global image structure.

For instance, in the seam carving approach, an eight-connected seam is removed/inserted o resize the images. A controlled mesh is deformed to guide image resizing and smoothly scales the neighboring mesh quads during the optimization process. In the previous work, the preservation of line structure in an image is focused, but it is not sufficient for different kinds of object structures. In contrast to the previous works, the image structure is preserved with a top-down strategy by using a block saliency map which is adaptive to the sizes of the structured objects.



On the other hand, the compressibility rate is computed for each image from its image gradient magnitude and orientation distribution. The retargeting process is optimized based upon the compressibility assessment in both x and y directions. The resized image frames, therefore, meet the fundamental requirements: retaining salient content and preserving global structure. In this paper, different strategies are utilized to adaptively resize images such that the salient image content and structures are well preserved. The proposed method is implemented in Mat lab.

Traditional scaling protects the structure well but changes the aspect ratios of prominent objects severely. The seam carving algorithm using backward energy to remove a connected seam and this sometimes produces structure twist. Compared to seam carving, the warping-based methods in general produce more smooth results. The image is partitioned into grid mesh and the bending of the grid lines is minimized, the results are produced smoothly between neighboring quads. However, it does not avoid the distortion accumulation and propagation along the structures. Without high-level saliency detection (e.g., face detection), content-aware algorithms may fail to preserve the prominent content.

Compared to the improved seam carving, the warping-based methods can produce smoother results in most cases. However, local smoothness constraint sometimes can not prevent accumulated distortion. Compared to other methods, the proposed results show better image content and object structure preservation without over compressing or stretching.

Although scale-and-stretch is a warping-based and 2-D resizing technique, without top-down global structure protection, it still suffers from accumulated structure distortion problem. Different types of retargeting approaches may need to combined together to obtain appropriate resized media. Except protecting the object structure, the proposed block structure energy facilitates the algorithm to operate like multioperator method. The final operations include cropping, uniform scaling and no homogeneous compression or stretching. Since the computation only takes the target aspect ratio and the image compressibility into account, the proposed method can also be used to enlarge the image size in either direction. The proposed method can better preserve the content structure and object aspect ratio.

Adaptive registration algorithm is used, in which low resolution image is mapped to the uniform high resolution grid. Then adaptive registration is used to form high resolution image. Weiner filter is used to reduce the effects of blurring and noise. PSNR can be defined as the ratio of the maximum possible power of signal to the power of noise. Using adaptive registration algorithm, PSNR ratio can be increased. As the PSNR value increases, resolution of the image also increases.

V. CONCLUSION

In this paper, an adaptive image resizing algorithm that well preserves prominent structure in the image is presented. Instead of minimizing the distortion of neighboring pixels or grids, block structure energy is defined that uniformly distributes the energy of local structure over the pixels inside the bounding box of the detected structure segment. Based upon this energy, the proposed algorithm enforces the deformation of each block area to be as uniform as possible. According to the image content, compressibility rate in each direction is estimated and the optimal scaling factors are determined which are used to resize the image to the optimal resolution with the same aspect ratio of the target image size. The salient contents of the image and the image structure are retained. In the graph, the previous image retargeting method is denoted by red dot whereas the proposed algorithm is denoted by blue dot. Experimental comparisons with the previous image retargeting methods showed superior structure preservation in the image resizing results by using the proposed algorithm.

REFERENCES

1. V. Setlur, S. Takagi, R. Raskar,M. Gleicher, and B. Gooch, "Automatic image retargeting," ACM Trans. Graph., vol. 154, p. 4, 2004.
2. S. Avidan and A. Shamir, "Seam carving for content-aware image resizing, ACM Trans. Graph., vol. 26, no. 3, pp. 10.1–10.9, 2007.
3. L. Q. Chen, X. Xie, X. Fan, W. Y. Ma, H. J. Zhang, and H. Q. Zhou, "A visual attention model for adapting images on small displays," ACM Multimedia Syst. J., vol. 9, no. 4, pp. 353–364,2003.
4. M. Rubinstein, A. Shamir, and S. Avidan, "Multi-operator media retargeting, ACM Trans. Graph., vol. 28, no. 3, pp. 1–11, 2009.
5. S.-F. Wang and S.-H. Lai, "Fast structure-preserving image retargeting," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., 2009, pp. 1049 – 1052
6. H. Liu, X. Xie, W. Y. Ma, and H. J. Zhang, "Automatic browsing of large pictures on mobile devices," ACM Multimedia, pp. 148–155,2003.
7. F. Liu and M. Gleicher, "Automatic image retargeting with fisheye-view warping," in Proc. ACM Symp. UIST, 2005, pp. 153–162.
8. R. Gal, O. Sorkine, and D. Cohen-Or, "Feature- aware texturing," in Proc. Eurograph. Symp. Rendering, 2006, pp. 297–303.
9. Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee, "Optimized scale-and-stretch for image resizing," ACM Trans. Graph., pp. 1–8, 2008.
10. Y. Pritch, E. Kav-Venaki, and S. Peleg, "Shift-map image editing," in Proc. Int. Conf. Computing.