ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 1, January 2013

Feature Based Image Sequence Retargeting in the Uncompressed Video Domain

Kavitha. S

Abstract:- The system propose a video retargeting algorithm to resize images based on the extracted saliency information from the compressed domain. The system utilizes DCT coefficients in JP2 bit stream to perform saliency detection with the consideration of the human visual sensitivity. Valuable retargeting requires emphasize the main satisfied while retain immediate context with minimal visual deformation. A number of algorithms have been proposed for image retargeting with image substance taken as much as potential. But, they usually suffer from deformation results, such as edge or structure twists. A structure and content preserving image retargeting technique is used that preserves the content and image structure. The image content saliency is estimated from the structure of the content using probability map. A block structure energy is use for structure conservation along both directions. Block structure energy uses top down strategy to constrict the image structure consistently. However, the flexibilities of retargeting are altered for different images. To defeat this problem, the patch transform is introduced, where an image is broken into non-overlapping patches, and modifications or constraints are applied in the "patch domain".. Thus, the resized image is produced to preserve the structure and image content quality.

Keywords: Video Retargeting, Saliency Detection, Structure-Preserving, Block Structure Energy, Patch Transform.

Manuscript received Jan 15, 2013.
KAVITHA.S, *Department of Computer Science and Engineering, Karpagam University, Coimbatore, Tamilnadu, India,*

I. INTRODUCTION

Image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it.

Image processing usually refers to digital image processing, but optical and analog image processing also are possible. This article is about general techniques that apply to all of them. The acquisition of images (producing the input image in the first place) is referred to as imaging. An image defined in the "real world" is considered to be a function of two real variables, for example, a(x,y) with a as the amplitude (e.g. brightness) of the image at the real coordinate position (x,y).

The image/video retargeting becomes more and more important due to the increasing demand for displaying image on a variety of display devices of dissimilar resolutions. A number of algorithms have been proposed to adapt image or video content to various display settings.

For the cropping based methods [1], [2], they proposed to automatically detect and transmit the most important region to the mobile device. Fan *et al.* [12] proposed to retarget video to small display devices by adapting content and creating a zoom-in-like effect. Liu and Gleicher [3] used both cropping and scaling with virtual camera motion to adapt wide screen films to a target resolution. In these algorithms, they applied cropping 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/video. To deal with this problem caused by cropping, Avidan and Shamir [4] 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. However, simply extending seam carving to video retargeting will create jittery artifacts. Therefore, they improved the seam carving algorithm to find minimal-cost seams by computing the forward energy to reduce the effect of artifacts [5]. Seam carving is an effective technique for image/video retargeting, although in some cases it can not avoid the problem of producing artifacts or distorting content structures. Simakov *etal.* [13] proposed to summarize visual data using bidirectional similarity measurement and showed it can be applied to a variety of problems.

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.* [6] segmented the image into several regions, scaled them independently according to the importance and then combined all the scaled regions. Similar idea of segmentation was also used for video retargeting [7]. Tao *et al.* [7] proposed a dynamic video retargeting method that integrated the foreground extraction, active window initialization and optimization processes. 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.

For the warping based methods, Liu and Gleicher [8] tried to find the region-of-interest (ROI) and used a nonlinear fisheye-view warp that emphasizes important parts while shrinking others. Gal *et al.* [9] proposed to warp an image to an arbitrary shape while preserving user specified features by constraining their deformation to a similarity transformation. It is achieved by a Laplacian technique to accommodate local constraints followed by a global optimization procedure. Recently, Wolf *et al.* [10] presented a warping based method that automatically detects the important regions by combining a saliency measure, face detector and motion estimation for video retargeting. They formulated the grid mapping of image resizing as solving a large and sparse linear system. Based upon the similar idea, Wang *et al.* [11] presented a method which allows important regions to scale uniformly and homogeneous regions to be distorted. This method gives more freedom to utilize homogeneous image regions. From the results in [10]and [11], the warping based methods in general show better structure-preserving property than the carving methods and better content-preserving property than the cropping methods. However, they are limited to well preserve the global shapes of prominent objects well in the retargeted image/video. Different approaches may be suitable to retarget images with different content.. Recently, Rubinstein *et al.* [14] proposed to combine cropping, scaling and seam carving to optimize the image similarity between the resized image and the original image.

II. RELATED WORK

To retarget an image with size $W_o \times H_o$ to a new image with size $W_r \times H_r$, traditionally we remove/insert relatively unimportant area 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. Based upon a warping based approach, we will first introduce a novel block structure energy for structure preservation in the following subsection. In addition, we propose to estimate the optimal compressibility rate according to the image content.

In this paper, we deal with the problem of image resizing from a new perspective, i.e., retargeting the image aspect ratio. Without considering the actual target size, we firstly 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.

A. Image Extraction

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.

To estimate the importance of an image, one can measure the magnitude of gradient as local energy or combine color, intensity and orientation for content saliency. Dissimilar image saliency measures, e.g., face detector, can also be included to improve the saliency measure. First, the input image is decayed into a set of multiscale feature maps which extract the information of color, intensity and orientation. Second, with nonlinear spatially competitive dynamics, all feature maps are combined into a unique scalar saliency map

In order to maintain the well-known 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. Here we provide a more simple and effective way to protect the extracted structure pieces described in the previous paragraph; namely, the block structure energy. The basic idea is that we use bounding boxes for 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. Please see Fig. 1 for an example, and assume that the straight line and curve are the extracted structure pieces and we enforce the deformation to be uniform in each direction and the curve also bend gracefully without producing artifacts.

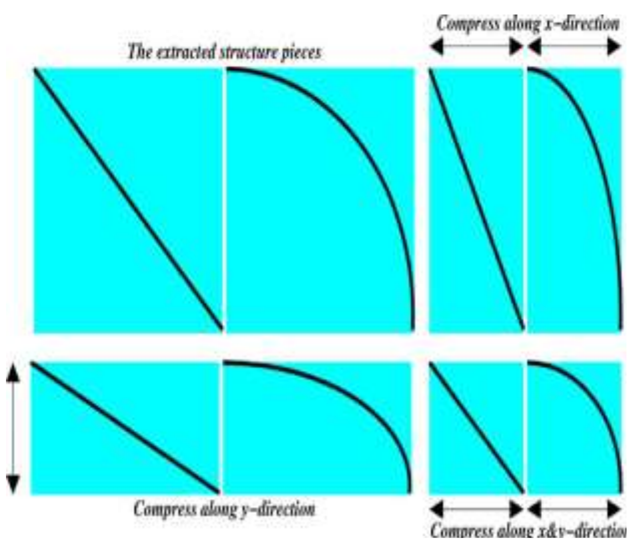

Fig. 1. Example of Extracted Structure pieces along both X & Y direction

Since we only retarget image along X and Y direction, some other bounding shapes are not suitable for the structure preservation, e.g., oriented bounding box. Take the straight line for instance, although an oriented bounding box better fits the structure piece, but it will degenerate to the line segment itself and cannot protect the structure in a global manner. The retargeting process is optimized based upon the compressibility estimation 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 .

B. Patch Transformation

The patch transform is used in order to break an image into non-overlapping patches, and modifications or constraints are applied in the patches. A modified image is then reconstructed from the patches, subject to those constraints. Constraints the user may specify include the spatial locations of patches, the size of the output image, or the pool of patches from which an image is reconstructed. The user can select regions of the image and move them to new locations. The patch transform reconstruction will then try to complete the rest of the image with remaining patches in a visually pleasing manner, allowing the user to generate images with different layout but similar content as the original image. The user specifies both the position of some patches and the size of the target image. The user can also mix patches from multiple images to generate a collection combining elements of the various source images.

C. Compressibility Estimation

The retargeting process is optimized based upon the compressibility estimation 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 .The idea of utilizing homogeneous image regions for shrinking and stretching along x and y directions to compress the image frames.

D. Video Retargeting

Video retargeting aims to adapt images to displays of small sizes and different aspect ratios. Effective retargeting requires emphasizing the important content while retaining surrounding context with minimal visual distortion. In video retargeting becomes more and more reasonable, users can easily create their own media content. To fit it into different types of display devices, the video frames should be resized adaptively based upon their content. However, naively resizing each frame individually may lead to sparkling artifacts. Since the streaming video is basically endless while the amount of memory available for our program is relatively small, a technique for processing endless video is needed. It is the limited extension of a ribbon that makes the sliding-window algorithm possible to process streaming video.

The basic idea is: load a block of frames from the original video into the buffer and condense the buffer until the stopping criterion is met, which will free some space in the buffer. As discussed above, because ribbons have limited extension along time axis, it is possible that there are some frames in the condensed video which won't be affected by any new ribbons. Therefore, these frames can be written out to a file and then some space in the buffer is freed up. Read new frames from the original video into the buffer until it is full. Repeat the procedure to condense a long video as desired.

In order to meet the requirement of video retargeting, we make some modifications and extend the image retargeting technique presented previously. First of all, the most salient area that draw a viewer's concentration in a video clip is the moving objects. The magnitude of motion field computed from the neighboring image frame and is also normalized to range between 0 and 1. For efficiency consideration, we only estimate the motion vectors on the detected corners. We approximately take the average of the motion vectors as the camera motion and the actual object motion can be estimated by eliminating the camera motion. This approximation is applicable for most of the videos with distant scene or small camera motion. For videos with more complex scene or large depth variations, more complex motion model would be a better alternative.

III. CONCLUSION

In this paper, we presented an adaptive video resizing algorithm that well preserves prominent structure in the video. Instead of minimizing the distortion of neighboring pixels or grids, we define a block structure energy 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, we further propose to estimate the compressibility rate in each direction. The compressibility rate and the total entropy helps to determine the optimal scaling factors which are used to resize the image to the optimal resolution with the same aspect ratio of the target image size. Experimental comparisons with previous image retargeting methods showed superior structure preservation in the image resizing results by using the proposed algorithm.

IV. REFERENCES

[1] **L. Q. Chen, X. Xie, X. Fan, W. Y. Ma, H.JZhang, 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.

[2] **H. Liu, X. Xie, W. Y. Ma, and H. J. Zhang,** "Automatic Thumbnail Cropping and its effectiveness," ACM Multimedia, pp. 148–155, 2003.

[3] **F. Liu and M. Gleicher,** "Video retargeting: Automating pan and scan," ACM kmultimedia, pp. 241–250, 2006.

[4] **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.

[5] **M. Rubinstein, A. Shamir, and S. Avidan,** "Improved seam carving for video retargeting," ACM Trans. Graph., pp. 1–9, 2008.

[6] **V. Setlur, S. Takagi, R. Raskar,M. Gleicher, and B. Gooch,** "Automatic image retargeting," ACM Trans. Graph., vol. 154, p. 4, 2004.

[7] **C. Tao, J. Jia, and H. Sun,** "Active window oriented dynamic video retargeting," in Proc. Workshop Dynam. Vis. ICCV, 2007.

[8] **R.Gal, O. Sorkine, and D. Cohen-Or,** "Feature-aware texturing," in Proc. Eurograph. Symp. Rendering, 2006, pp. 297–303.

[9] **L.Wolf, M. Guttmann, and D. Cohen-Or**, "Non-homogeneous content driven video- retargeting," in Proc. Int. Conf. Comput. Vis., 2007, pp. 1–6.

[10] **Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee,** "Optimized scaleand- stretch for image resizing," ACM Trans. Graph., pp. 1–8, 2008.

[11] **X. Fan, X. Xie, H.-Q. Zhou, and W.-Y. Ma,** "Looking into video frames on small displays," ACM Multimedia, pp. 247–250, 2003.

[12] **D. Simakov, Y. Caspi, E. Shechtman, and M. Irani**, "Summarizing visual data using bidirectional similarity," in Proc. IEEE Conf. Comput. is. Pattern Recognit., 2008, pp1–8.

[13] **M. Rubinstein, A. Shamir, and S. Avidan,** "Multi-operator media retargeting," ACM Trans. Graph., vol. 28, no. 3, pp. 1–11, 2009.

[14] **B.Chen and P. Sen,** "Video carving," in Proc. Eurograph., 2008.

[15] **T. S. Cho, M. Butman, S. Avidan,** and W. T. Freeman, "The patch transform and its applications to image editing," in Proc. IEEE Conf. Comput. Vis. attern Recognit., 2008, pp. 1–8.

[16] **Y. Pritch, E. Kav-Venaki, and S. Peleg,** "Shift-map image editing," in Proc. Int. Conf. Comput. Vis., Sep.–Oct. 2009, pp. 151–158.

[17] **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.

[18] **L. Itti, C.Koch, and E. Niebur,** "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Pattern Anal. Mach. Intell. vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[19] **J. V. D. Weijer, T. Gevers, and A. W. M. Smeulders,** "Robust photometric invariant features from the color tensor," IEEE Trans. Image Process., vol. 15, no. 1, pp. 118–127, Jan. 2006.

[20] **C. C. Paige and M. A. Saunders,** "Semi-automatic Approach to Home Video editing," SIAM J. Numer. Anal., vol. 12, pp. 617–629, 1975.