RESEARCH ARTICLE OPEN

Image Retargetting on Video Based Detection

ALOK THAKUR, NEERAJ TIWARI
Electronics And Communication College-Tit Bhopal Emai-Aloksinghv@Yahoo.Com
Assistant Professor, Electronics And Communication College-Tit Bhopal
Email-nkt.0107@gmail.com

ABSTRACT

We present a algorithm for retargeting large images to small size displays, particularly from video. The sample video is converted in the form of Frames. Now algorithm is apply to the frames of images to filter the background from the video frames. This method adapts large images so that important objects in the image are still recognizable when displayed at a lower target resolution. Existing image manipulation techniques such as cropping works well for images containing a single important object, and down sampling works well for images containing low frequency information. The retargeting algorithm segments an image into regions, identifies important regions, fills the resulting gaps, resizes the remaining image, and re-inserts the important regions. Our approach lies in constructing a topologically constrained epitome of an image based on a visual attention model that is both comprehensible and size varying, making the method suitable for display-critical applications.

Keywords:- Small displays, Image adaptation, Image attention, Image Redirecting, Saliency Detection, Image cropping.

I. INTRODUCTION

The main problem of obtaining objects from the scene and separating them from the background of image is figure– ground separation and redirecting the particular section of the image. The human brain can do this separation very easily and fast [3], but doing same on a machine is one of the major challenge for engineers and scientists. Visual content is becoming more important for sharing, expressing, and exchanging information on devices such as, cell phones and hand-held PCs [1].The main problem is related with image application of machine vision, scene understanding, content-based image retrieval, object recognition and tracking. Image retargeting is also useful for WYSIWYG directory icons for the efficient selection of images from directories and large image databases. Regardless of whether the resolution of the screen is high or whether the bandwidth is high, retargeting addresses the issue of displaying images on screen sizes with limited display real-state.

Simply scaling images reduces the size of important features. If there is a single important feature in the image, the image can be cropped and scaled to fit. Images with multiple, important features present a more challenging case for retargeting. In such cases, valuable image area in the target image may be wasted with unimportant regions between important features. For example, in Figure 1 there are important features on both sides of the image and cropping cannot remove the unimportant area between them. For many images, the key content is a small set of objects. To effectively display such images on small displays, these objects must be displayed at a sufficient size that they can be easily recognized. Other objects in the image, as well as the precise relationships between objects, are less important.

Most of the images are stored in the compressed domain of joint photographic expert group (JPEG). The compressed JPEG images are widely used because of their compress size. In order to obtain features from compressed image, the existing saliency detection have to decompress there JPEG image from the compressed domain into the spatial domain. The full decomposition from saliency detection is very time consuming but computing consuming also.

To assist in generating these increasingly important small images, we introduce a novel method for Automatic Image Retargeting. The goal is to provide *effective* small images by preserving the recognizability of important image features during downsizing. The premise of our method is that if the important objects in a given image can be identified, their size can be exaggerated in the target image such that they are more recognizable. Such exaggeration necessarily comes at the expense of realism: we intentionally distort less important parts of the image to make more important parts clearer. In contrast to traditional resizing and down-sampling techniques, the results of our retargeting method show important objects exaggerated in size, by reducing the spacing between objects, so that they are easier to recognize.

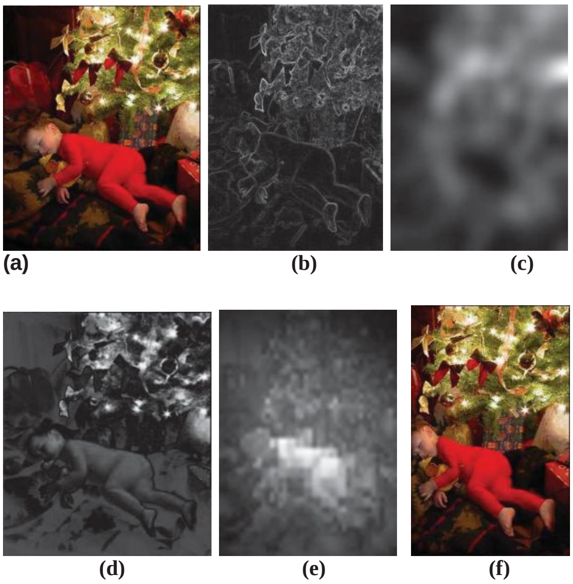


Fig. 1. Comparison of the different image retargeting algorithms (a) original image, (b) gradient map (used in [5] and [7]), (c) Image from Itti's model [3] (used in [6]), (d) saliency map from the model in [6] , (e) image from our proposed model, and (f) retargeted image our proposed algorithm. The width of the retargeted images is 75% of that from the original image.

II. RELETED WORK

Image resizing can be performed manually using standard tools. Commercial products enable the manual resizing of images using cropping and scaling operations. However, this process is often tedious, especially with large data sets. Cropping tends to work well for images containing single objects of importance. Scaling tends to distort important regions. Also, performing retargeting operations other than simple cropping and scaling requires a great deal of skill and effort.

A few researchers have explored automating image retargeting through automatic cropping processes. For example, Suh et al. [2] proposed two techniques for automatic cropping based on using a visual attention model to detect interesting areas in an image. Their first method is based on saliency maps [3], while the second is based on face detection [4]. In both cases the output is a thumbnail, created by cropping and scaling the source image to capture a single object. However, neither method can handle cases where there are multiple important features in an image.

Jojic et al. proposed a model of image representation, called an *'epitome'* that attempts to encode the image's essence [5]. This is similar to the retargeting goal. The epitome of an image is its miniature, condensed version containing most constitutive elements needed to reconstruct the original image. The main idea is to uniquely map every patch in the epitome to a corresponding patch in the original image. This works well when the original image contains several small, repetitive unit patterns. This technique would not be suitable for obtaining a more comprehensible image where the neighborhoods between important regions are required to be maintained. Epitomes do not always give an idea of the overall picture because neighborhood relationships between regions are significantly violated. The topology of an image is characterized by the neighborhood relationships among different elements that comprise the image. Our approach lies in constructing an epitome of an image that is topologically constrained, based on a visual attention model that is both comprehensible and size varying, suitable for display-critical applications.

The use of deformations to exaggerate portions of images and displays has been used in various techniques, as found in presentation literature. See Carpendale and Montagnese [8] for a survey. Our method is the first to combine non-photorealistic deformation to an image retargeting application, and the cut-and-paste algorithm is uniquely suited to maintaining the recognizability of key image objects.

A. Saliency Detection.

The saliency detection model proposed by Itti et al. and it's design is based on the neuronal architecture of the primates early vision system [3]. In our paper, the saliency map is calculated with the help of three features of the image frame of video: intensity color and orientation. The combination of this three quality of images is used to obtain the final saliency map. Based on the Itt's model, Harel et al. proposed the graph-based on visual saliency model by using a graph-based dissimilarity measure [10]. Another author Ma et al. devised a saliency detection model based on the local contrast analysis [11]. In this paper the hyper complex from algorithm and Gaussian filter algorithm is adapted to obtain the salient region of the image. In Gonferman et al. built a contrast – aware saliency detection with consideration of the contrast from both local and global perspective. One another author Liu et al. utilized the machine learning technique to obtain the saliency map for image processing [4]. In this paper we used a Gaussian filter and binary filter to get the image smooth and remove the noise factor from the image frame of video.

Saliency Detection is used to find the Region of interest.
This is used to detect the foreground and background for the image segmentation.

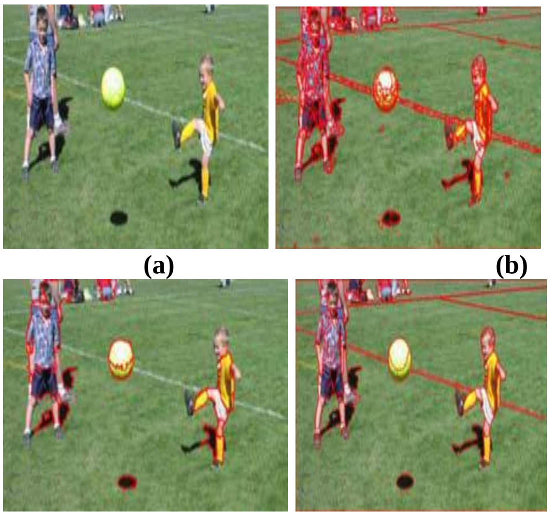
B. Regions of Importance (ROI)

The saliency and face detection algorithms take color images as input return gray-scale images whose pixel values represent the importance of the corresponding pixel in the input image. The importance map, which is the attention model for the image, is built up by combining a series of importance measures. This allows the system to be adapted to differing image creation goals, and to be easily extensible.

III. IMAGE SEGMENTATION

In order to identify important regions in the image, we must first segment the image. We use mean-shift image segmentation [9] to decompose the given image into homogeneous regions. The advantages of this approach include flexible modeling of the image and noise processes and consequent robustness in segmentation. The segmentation routine takes as input, the parameters.

Spatial radius *hs*, color radius *hr*, and the minimum number of pixels *M* that constitute a region. As with other segmentation methods, choosing optimal parameter values is often difficult. Therefore we over-segment the image using lower values of *hr* and *M* and merge adjacent regions based on color and intensity distributions in a perceptually uniform color space, CIE-Luv. In practice, values of $hs = 7$, $hr = 6$, and $M = 50$, tends to work well for over segmentation for most images.



(a) **(b)**

(c) **(d)**

Fig 2: Image segmentation. *a)* The original image. *b)* Applying mean-shift with *parameters hs* = 7, *hr* = 6, *and M* = 50. *c)* Applying mean-shift with *parameters hs* = 32, *hr* = 30, *and M* = 150. d) Performing region simplification on (b).

We then compute a color similarity measure called histogram intersection [Swain and Ballard 1991] to determine color similarity between regions, and perform region simplification by merging adjacent regions. Figure 2 illustrates an example of this technique. Histogram intersection matches the image color histogram of a given segmented region with histograms of each of the adjacent regions. Given a pair of histograms, *Q* and *T*, each containing n buckets, the intersection of the histograms is defined to be:

$$\sum_{j=1}^{n} min(Q_j, T_j).$$

Where j ranges over each color in the histograms.

Our system creates a *DualGraph*, defined by nodes and edges to store the spatial region information of the segmented image. A node in the dual graph corresponds to a region in the segmented image, and an edge between two nodes indicates that two regions are adjacent to one another. Each node also contains a histogram of the RGB color information of the region, and is later used in the retargeting process.

IV. RETARGETING PROCESS

Our algorithm takes as input, a source image and a specification for the size of the output image. Figure 4 summarizes the algorithm. We first segment the source image into regions. We then use an importance map to select a set of Regions *of* Importance (*ROI*) to exaggerate in the result. Alternatively, the algorithm can be applied in a semi-automatic fashion by having the user specify the *ROI*. In previous section. We discuss the techniques involved in segmenting the image, and combining adjacent regions based on their spatial distribution of color/intensity. In order to identify important regions, we generate an importance map of the source image using saliency and face detection. If the specified size contains all the important regions, we simply crop the source image. Otherwise, we remove the important regions from the image, and fill the resulting "holes" using a background creation technique as described. We then resize the updated background to fit the input specification. Regions of importance are then "pasted" back onto the updated background based on their importance, and relative topology within the scene. If all the important regions are not able to fit within the new image, we resize these regions inversely proportional to their importance.

V. WRAPPING BASED METHOD

Liu and Gleicher introduce a technique that warps the image in a way similar to a fisheye lens, by employing a piecewise linear warping scheme that has more distortion in uninteresting areas and less distortion in the ROI [12]. This method assumes that there is only one ROI per image, and the importance map is computed using a contrast-based method. Gal et al. propose a method for mapping textures into different surfaces, which avoids distortion of important features. The user manually species important areas, and their deformation is constrained to be a similarity transformation, within a Laplacian image editing optimization framework. Wolf et al. retarget videos to a smaller width or height (images come as a particular case). A sparse system of linear equations is solved to determine the new pixel locations. It is built from constraints that specify where output pixels must be with respect to their neighbors, weighted according to an importance map that combines the L2-norm of the gradient, face detection, and motion detection. In the video case, additional constraints enforce smoothness across adjacent frames.

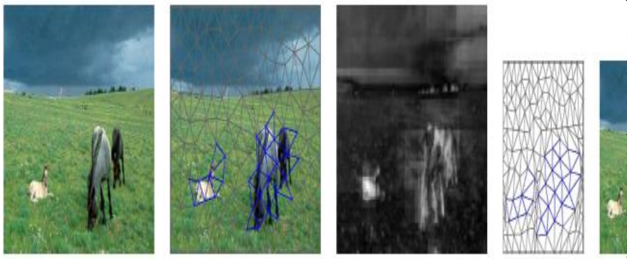


Fig3: A mesh is built from the input and associated with saliency information (blue edges). The mesh is then resized, and the output image is rendered using texture mapping.

VI. PROPOSD WORK

The overall diagrammatic representation of the proposed work is as shown below:-

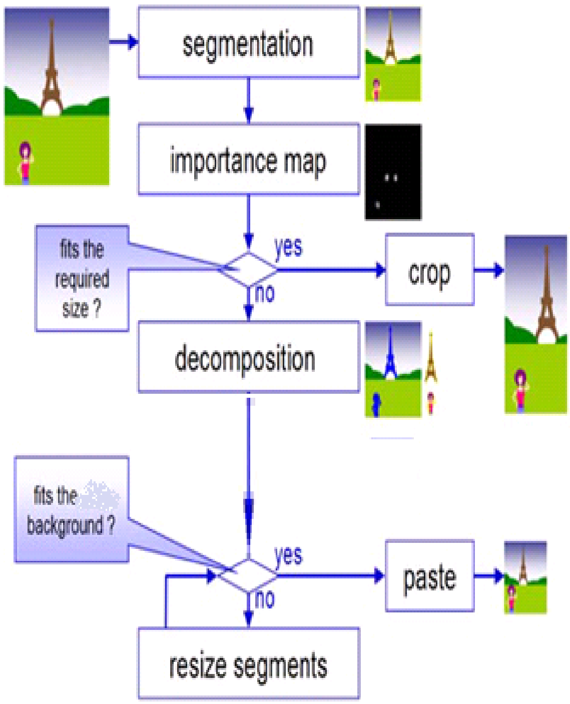
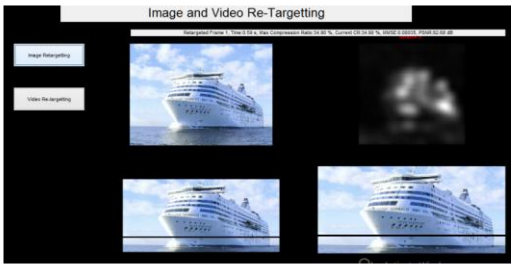


Fig. 4 Flowchart of our retargeting algorithm.

Above flowchart represents the overall proposed work for the image redirecting. First the video is converted into series of images, to apply the redirecting algorithm. Then segmentation is apply to image for obtaining the segments from the image. After getting the segments from the image importance map is apply for the finding the ROI. Saliency detection technique is used to find the ROI. If the required size is obtained then crop the image for redirecting, otherwise decomposed the image by finding RIO from image to obtain the background of image. Reduce the size of the background. Now if the background is fits for redirecting the image then paste to obtain it, if not then repeat the procedure again to obtain the desired output.

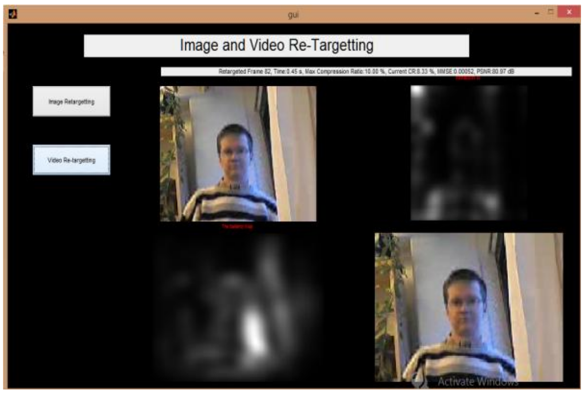
VII. EXPERIMENTAL RESULT

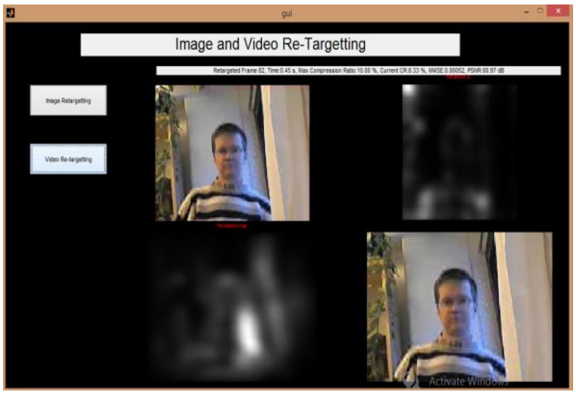
Here we check the performance of our research on various videos, and the following outputs are obtained,



Here the compression ratio after re-targetting is about 35%, and we can see that the output image is a retargetted form of the input image (ship is being retargetted here)

Secondly we tested the same on videos, here are some results for the same,





As we can see for videos the compression ratio goes from 10% to 30%, and the quality is also very good

VIII. CONCLUSION

Image and video retargetting gives perfect results with proper image quality at the output. Static images give a compression ratio or re-targeting factor between 0.35 and 0.5, while videos give a compression ratio of 0.1 to 0.3. Thus we can conclude that the saliency map retargeting technique is suited for both the applications, but works at a high efficiency for images

REFERENCES

[1] LIU, H., XIE, X., MA, W.-Y., AND ZHANG, H.-J. 2003. Automatic browsing of large pictures on mobile devices. In Proceedings of the eleventh ACM international *conference on Multimedia*, ACM, 148–155.

[2] SUH, B., LING, H., BEDERSON, B. B., AND JACOBS, D. W. 2003. Automatic thumbnail cropping and its effectiveness. In *User interface software and technology*,ACM, 11–99.

[3] H. Zhou, H. Friedman, and R. von der Heydt, "Coding of Border Ownership in Monkey Visual Cortex," J. Neuroscience, vol. 20, no. 17, pp. 6594-6611, 2000.

[4] ROWLEY, H. A., BALUJA, S., AND KANADE, T. 1996. Human face Detection in visual scenes. In *Advances in Neural Information Processing Systems*, The MIT Press, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds., vol. 8, 875–881

[5] JOJIC, N., FREY, B., AND KANNAN, A. 2003. Epitomic analysis of Appearance and shape. In *9th IEEE International Conference on Computer Vision*, 34–41.

[6] T. Ren, Y. Liu, and G. Wu, "Image retargeting based on global energy optimization," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jun.– Jul. 2009,pp. 406–409.

[7] M. Rubinstein, A. Shamir, and S. Avidan, "Improved seam carving for video retargeting," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–9,2008.

[8] CARPENDALE,M. S. T., AND MONTAGNESE, C. 2001. A framework for unifying presentation space. In *UIST '01*, 82–92...

[9] MEER, P., AND GEORGESCU, B. 2001. Edge detection with embedded confidence. *IEEE Transactions on Pattern Analysis and Machine Intelligence 23*, 12

[10] A. Oppenheim and J. Lim, "The Importance of Phase in Signals," Proc. IEEE, vol. 69, no. 5, pp. 529-541, May 1981.

[11] M. Hayes, J. Lim, and A. Oppenheim, "Signal Reconstruction from Phase or Magnitude," IEEE Trans. Acoustics, Speech, and Signal Processing, vol. 28, no. 6, pp. 672-680, 1980.

[12] Liu, F. and Gleicher, M., \Automatic image retargeting with _sheye-view w arping," in [ACM Symposium on User Interface Software and Technology], 153{162 (2005).