A Novel 3-D Image Retargeting using Stereo Seam Carving with Semantic Collage of Images

Mr. Mahendra T. Jagtap[1], Prof. Dr. R.C.Tripathi[2]
[1]*Ph.D Scholar, Computer Science and Engineering Department, Lingaya's Vidyapeeth, Faridabad*
[2]*Professor, Computer Science and Engineering Department, Lingaya's Vidyapeeth, Faridabad*

***Abstract*- Preserving the semantic meanings by keeping the contents of the original image as it is without disturbing its nature is the process of 'Image Retargeting'. It's wide aim is to display the image on any arbitrary size of display device irrespective of its aspect ratio or resolution. When the image is to be fitted in the target resolution, certain unimportant regions need to be deleted. The existence method predicts the pixel-wise importance in a bottom-up manner. In contrast, the proposed algorithm uses the top-down approach. This estimation is fused by 'Classification guided Fusion Network (CFN). This feature is widely applied in the 3-D image retargeting which fuses the left and right eye images which are differently designed. Semantic collage the images by fusion.**

Index Terms- Classification fusion network, Inage Retargeting, Semantic collage, Seam carving.

I. INTRODUCTION

In recent years, the rapid development of technologies especially the mobile technology and the popularity of digital display devises imposes the need of effective ways for presenting the images. Nowadays, various types of displays ranging from high resolution computer monitors or multiscreen to low resolution mobile devises are used to view the images. So, the differ in aspect ratio or resolution may affect on the visual quality of the image and it may lead to image degradation. To solve all these consequences, the image retargeting is used as the remedial measure. In order to maintain the desirable visual quality in the various handheld devises and display panels, the appropriate image retargeting is the need of today's era.

Considering a source image is essentially a carrier of visual information, For example, the image in Fig.1 shows a boy kicks a ball on a pitch (sports field), which contains four semantic components including boy, kicking, ball and pitch. Based on the source image, four target images can be generated as shown in Fig.1.

The first three target images are less informative as certain semantic components are missing. The last target image is the only one that preserves all four semantic components. Existing retargeting methods [1]–[4] operate based on an importance map which indicates pixel-wise importance. To generate a target image in Fig.1 that preserves semantics well, the pixels corresponding to semantic components, e.g., boy and ball, should have higher weights in the importance map such that these are preserved in the target image. In other words, an importance map needs to preserve semantics of the original image well.

Existing retargeting methods are based on an importance map which predicts pixel wise importance in a bottom-up manner via eye fixation estimation or saliency detection. The later work estimates a top-down approach where the target image maintains the semantic meanings of the original image. For that, the semantic components are extracted (foreground, background, action context). These semantic component maps are integrated by a classification guided fusion network. Specifically, the deep neural network classifies an original image as – object oriented and scene oriented and fuses the semantic component maps according to classification result. The network output referred to as semantic collage with the same size as the original image.

As Visual comfort is a quite important factor in 3D media service. Few research efforts have been carried out in this area especially in case of 3D content retargeting which may introduce more complicated visual distortions. As a contribution, we can propose a Hybrid Distortion Aggregated Visual Comfort Assessment (HDA-VCA) scheme for stereoscopic

retargeted images (SRI), considering aggregation of hybrid distortions including structure distortion, information loss, binocular incongruity and semantic distortion.

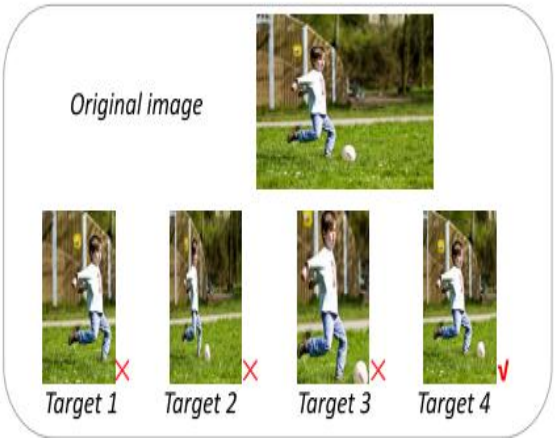

Fig1. : The original image shows a boy kicks a ball on the pitch. The first three target images are less informative as important semantic components are missing, e.g., ball is missing, main foreground object is distorted, and background region is not well retained.

II. LITERATURE REVIEW

Nowadays more efforts have been made to analyze image contents for retargeting. Leu [1] depicts that the image cropping is one of the most important operations performed to enhance photographs. The direct benefit is improved image composition in terms of better subject location, better subject magnification, and reduced background clutter. Indirectly, cropping can also help image and subject rendering because unwanted distractions are eliminated before image enhancement. We present a robust and efficient solution to this challenging problem for consumer snapshot photos.

Issues: Bottom-up method that driven by low-level visual cues, edge and corner in images are detected rather than semantic regions.

Ding [2] et.al. exhibits the Importance filtering for image retargeting. In this paper we present a novel filtering-based technique to tackle this issue, called "importance filtering". Specifically, we first filter the image saliency guided by the image itself to achieve a structure-consistent importance map. We then use the pixel importance as the key constraint to compute the gradient map of pixel shifts from the original resolution to the target. Finally, we integrate the shift gradient across the image using a weighted filter to construct a smooth shift map and render the target image.

The weight is again controlled by the pixel importance. The two filtering processes enforce to maintain the structural consistency and yet preserve the important contents in the target image.

Issues: Filtering operation does not maintain smoothness and consistency of resized image.

Stas Goferman, LihiZelnik-Manor [3] in their literature proposes a new type of saliency—context-aware saliency—which aims at detecting the image regions that represent the scene. This definition differs from previous definitions whose goal is to either identify fixation points or detect the dominant object. In accordance with our saliency definition, we present a detection algorithm which is based on four principles observed in the psychological literature. The benefits of the proposed approach are evaluated in two applications where the context of the dominant objects is just as essential as the objects themselves.

Issues: The background should be essential for understand the context.

In [4], D. Cho et. al. use Weakly and Self-Supervised learning for content aware deep image retargeting. This paper proposes a weakly- and self-supervised deep convolutional neural network (WSSDCNN) for content aware image retargeting. the network takes a source image and a target aspect ratio, and then directly outputs a retargeted image. Retargeting is performed through a shift map, which is a pixel-wise mapping from the source to the target grid. the method implicitly learns an attention map, which leads to a content-aware shift map for image retargeting. As a result, discriminative parts in an image are preserved, while background regions are adjusted seamlessly.

Issues: It is used for high level semantic information.

A. Radford, L. Metz [7] et. al. uses Unsupervised representation learning with deep convolutional generative adversarial networks helps to bridge the gap between the success of CNNs for supervised learning and unsupervised learning. Training on various image datasets showing the convincing evidence that the deep convolutional adversarial pair learns a hierarchy of representations from object parts to scenes in both the generator and discriminator.

Issues: The learned features for novel tasks and their applicability as general image representations is not properly illustrated.

In [13] Y. Chen et. al. helps to Improved seam carving combining with 3D Saliency for Image retargeting. In this paper, the seam carving content aware image retargeting algorithm is used that removes the pixel with less energy values during resizing process to preserve important part.

Issues: difficult problem to determine the energy function for removing task.

Yuming Fang et. al. [14] uses saliency detection for stereoscopic images for 2D multimedia processing applications. It proposes a new stereoscopic saliency detection framework based on the feature contrast of color, luminance, texture and depth are extracted from DCT coefficients to represent the energy for image patches.

Issues: luminance or pixel intensity may differ and extraction may hampered.

III. OUR CONTRIBUTION

As a contribution we can introduce 3D information into seam carving and warping procedure in different ways. In seam caving, 3D saliency works with L-1 norm of gradient to generate energy map for seam searching. In warping, 3D information helps to build structure constraint and 3D saliency adjusts weight of mesh edges in warping. The experimental results demonstrate the advantages of using 3D information both in the seam carving and warping procedure. Viewing 3-D images/ movies are having left and right eye variations while watching, the seam carving uses to fuse the images for better visualization effect.

IV. PROBLEM FORMULATION

Different from existing retargeting methods, we propose to explicitly preserve semantics of the source image by first extracting multiple kinds of semantic components and then combining them automatically.

We propose a classification guided fusion network to fuse the semantic component maps into a semantic collage with pixel-wise importance measures. In addition, object and scene images are considered differently.

V. PROPOSED SOLUTION

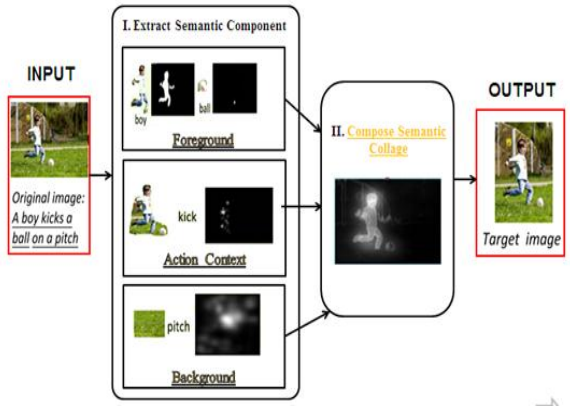

Fig.2 Semantic component Extraction

- Foreground Extraction (boy and ball Fig.1)
  To extract the foreground components image parsing and image classification is used.

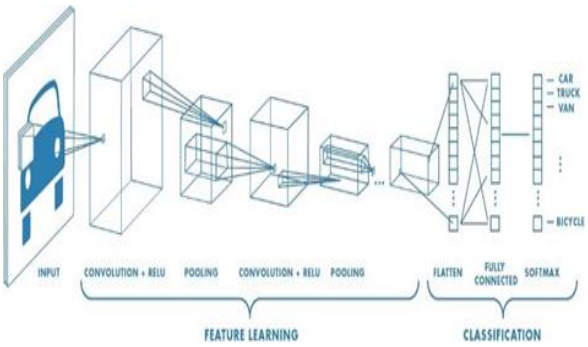

Fig.2 Image parsing and classification steps

Semantic component is denoted by M_p for image parsing and for image classification M_c.

Action Context Extraction (Kicking the ball- Fig. 1)

- The action context surrounding the foreground objects for image retargeting.
- The action recognition process carried out on the detected bounding box surrounding object by RCNN method.

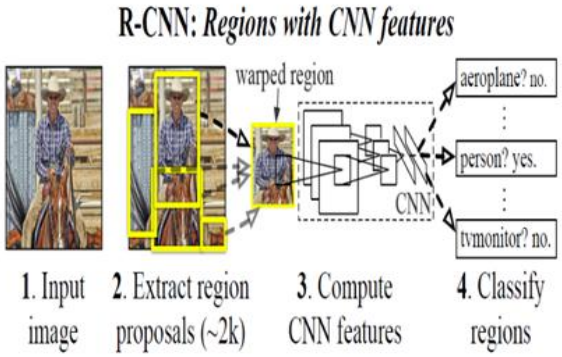

Fig.3 Steps of RCNN method

– The semantic component map derived from the action context is denoted as M_a.

• *Background Extraction* (Pitch is a background component- Fig.1)
– Scene classification is used to extract the background components which is same as image classification.
– The semantic component map is constructed from scene classification M_s which explains the scene labels.

• *Compose Semantic Collage*
– The semantic collage preserves the semantics and integrates multiple semantic component maps based on different cues.

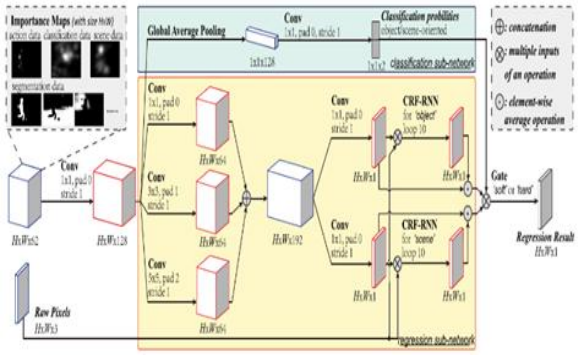

Fig.4 Classification Fusion Network

• Compose Semantic Collage

The semantic collage M_g is obtained by:

$M_g = c(o|M).\, r_o(M) + c(s|M).\, r_s(M)$

Where, $M = \{M_p, M_c, M_s, M_a\}$ is the concatenation of all semantic component maps to be fused. r_o and r_s are regression function for object-oriented and scene-oriented. c(o) and c(s) are the confidences that the image belongs to object-oriented or scene-oriented.

VI. AIM

To design, develop and test a system to display an image on various devices with respect to its aspect ratio and the semantic collage of 3-D images.

VII. OBJECTIVES

- To preserve important contents and structure of an image
- To fit an image with the target resolution.
- To test the system performance with respect to various data size and dimension.
- To compare the system performance with existing methods.
- To incorporate 3D information into both seam carving and mesh warping for image retargeting.

VIII. MOTIVATIONS

It is observed that conventional and semantic based image retargeting focuses on low-level features. Edges and corners in images are detected rather than semantic regions.

There is a need to improve the system performance under various circumstance such as:

The pixel-wise importance based high-level feature where the target image maintains the semantic meaning of original image.

If image contain multiple objects, then it should preserve important information.

IX. CONCLUSION

The proposed semantic collage algorithm performs favorably on the S-Retarget dataset using all metrics. System shows the semantic collage generated by the proposed SP-DIR (**S**emantic **P**reserving **D**eep **I**mage **R**etargeting) algorithm for extracting semantic components and compose semantic collage and the importance maps using a typical indoor image. Saliency detection on 3-D image in both seam carving and mesh warping for image retargeting will be used.

ACKNOWLEDGMENT

I acknowledge my sincere thanks to my Guide Prof. Dr. R.C. Tripathi for timely guidance in the contents. my sincere thanks to Dr. Amit Tyagi for support and thanks to HOD, CSE and All the faculties of CSE department.

REFERENCES

[1] J. Luo, "Subject content-based intelligent cropping of digital photos," in Proc. ICME, Jul. 2007, pp. 2218–2221.

[2] Y. Ding, J. Xiao, and J. Yu, "Importance filtering for image retargeting," in Proc. CVPR, Jun. 2011, pp. 89–96.
[3] S. goferman, Z. Lihi, "Content aware saliency detection", IEEE Trans.,Pattern anal.,vol.34,otc-2012.
[4] D. Cho, J. Park, T.-H. Oh, Y.-W. Tai, and I.-S. Kweon, "Weakly- and self-supervised learning for content-aware deep image retargeting," in Proc. ICCV, 2017, pp. 4568–4577.
[5] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proc. CVPR, 2015, pp. 3431–3440.
[6] A. van den Oord, N. Kalchbrenner, and K. Kavukcuoglu, "Pixel recurrent neural networks," in Proc. CVPR, 2016, pp. 1–11.
[7] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in Proc. CVPR, 2015, pp. 1–16.
[8] I. Goodfellowet al., "Generative adversarial nets," in Proc. NIPS, 2014, pp. 2672–2680.
[9] M. Mirza and S. Osindero, "Conditional generative adversarial nets," in Proc. CVPR, 2014, pp. 1–7.
[10] L. Itti, C. Koch, E. Niebur, "A model of saliency-based visualattention for rapid scene analysis", IEEE Trans. Pattern Anal. Mach.Intell. 20 (11) (1998) 1254–1259.
[11] M.A. Hasan, C. Kim, "An automatic image browsing technique for small display users", in: 11th International Conference on AdvancedCommunication Technology, 2009. ICACT 2009, vol. 3, IEEE, 2009,pp. 2044–2049.
[12] Y. Liang, Z. Su, C. Wang, D. Wang, X. Luo, "Optimized image retargeting using aesthetic-based cropping and scaling", IET ImageProcess. 7 (1) (2013) 61–69.
[13] Y. Chen, Y. Pen, Minglong Song, Meng Wang, "Improved seam carving combining with 3D Saliency for Image retargeting", Neuro Computing 151 (2015) Elsevier.
[14] Yuming Fang, Junle Wang, Manish Narwaria, Patrick Le Callet, Weisi Lin, "saliency detection for stereoscopic images" School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China.
[15] Si Liu, Zhen Wei, "Composing Semantic Collage for Image Retargeting", IEEE Trans. On Image Proc, Vol. 27, No. 10, Oct. 2018.