COMPRESSED TRANSMISSION OF DEPTH MAPS IN 3D STREAM SERVICES FOR ROBOTICS AND SURVEILLANCE

ANKUSH RAI*, JAGADEESH KANNAN R

School of Computing Science & Engineering, VIT University, Chennai, Tamil Nadu, India. Email: ankushressci@gmail.com

ABSTRACT

Building high end processing hardware for depth mapping in mobile robotics is a major drawback. The problem could be addressed by processing the scene through one end and then streaming it to the other robotic mobile platforms or actuators to perform physical operations; thereby rendering global depth map for all the arbitrary viewpoints of the robots. In this study, we present the algorithm for compressed transmission of depth maps over a network and provide a synthetic viewpoint with low geometric distortions.

Keywords: Network transmission, Video sequences, Depth determination.

INTRODUCTION

Recent developments in 3D display technologies allow incorporation of depth cues. Furthermore, in free viewpoint video, the viewer can display the viewpoint as per the dynamic view of viewer's, rendered into the video scene. To enable such dynamic synthesis, the depth information is required to be added in the color video. Therefore, the new compression algorithm is required to be investigated. The process of segmenting several feature parts in a digital image is known as image segmentation. However, these varying features of the image may represent an object with certain patterns, colors or textures which in the scenario of image segmentation are required to extract all these features of the object to in the form of a computable property to avail the position, depth and geometrical characteristics of the object in the form of contours. Edge detection and region detections are the frequent research topic in computer vision and image processing. The problem in this scenario lies with the changing boundaries of the object leading to bestowing varying illumination of light intensities and thereby making the processing for object detection fail miserably.

The previous work on depth encoding mechanism involves various classical transformation based encoding mechanism like in H.264/AVC [1]. Later, the HPEG-2000 approach is considered which guarantees transformation based algorithms [2,3]. However, these methods are often prone to the ringing effects on the artifacts along the edges, leading to incorrectly wrapped pixels [4]. Another category of such depth compression algorithm involves dividing the images into triangular meshes or platelets [5,6]; this is followed by modeling the segments by using piecewise linear function and partitioned it in correspondence to tree structure. The drawback of these approaches is that it cannot give pixel accurate boundaries for precise object representation [4,7]. The following approach combines the advantages of the two and is integrated into a newly modeled algorithm into pooled contour networks.

METHODS

Experimental setup

The proposed model is prototyped over MATLAB R2012a under Windows platform. The experiments are conducted over the machine with hardware configurations of Intel's seventh generation 8-core microprocessor, 8 GB RAM giving a fine clocking speed of 2.7 GHz. For evaluation purpose, we have used the publicly available Global

Pb system [8] which is widely used by the researchers for contour detection. For benchmark, the datasets that we used in this study is BSDS5000 benchmark [8], which is just an extension of the previous BSDS300 benchmark used commonly for evaluation of contours. It consists of more than 500 natural images with a pixel resolution of around 321 × 481. These 500 natural images comprise 200 frames for training purposes and the rest 100 and 200 for validation and testing purposes, respectively. In addition, we have also used MSRC2 and PASCAL2008 segmentation datasets [9,10] which comprise 1023 images, characterized roughly with the resolution of 350 × 500. Finally, we used NYU Depth dataset (v2) which includes 1449 pairs of color RGB and depth frames of resolution 480 × 640 [11].

Algorithm for compressed depth map video streaming over local network

The proposed algorithm ensures the representation of the accurate depth maps continued across the several frames of the given video sequences compressed by ABLATA algorithm [12-14]. As the contours offers, an advantage over the information extracted by gradient-based edge detection technique to give closed boundaries and describe the complete segment [12,13]. Therefore, the algorithm fills those segments with networks of contours and hence ideally describes perfectly the segment boundaries in dynamic environment. The idea here is to normalize the given RGB color scheme, thereupon the contours are segmented and finally pooled in together to form a network of contours which is carried over to the next frames and so on. The result and the workflow process of the proposed algorithm

Fig. 1: Sample result of the compressed depth map created by the proposed algorithm from the given input video sequence and facilitating sharing the mapped data over a network

are shown in Figs. 1 and 2, respectively, also the algorithm for the same is given as:

Algorithm: Compressed Depth Map Streaming

Input: Video sequence with N frames, each frame is characterized with m ×n RGB standard image, where m and n are the row and column of the given image.

Output: Coded contour networks.

Step 1: Normalize RGN channels

Loop: for i to m

Loop: for j: n//RGB normalization to rg scheme

$$r_{i,j} = \frac{R}{(R+G+B)}$$

$$g_{i,j} = \frac{G}{(R+G+B)}$$

end

end

Step 2: Segment block patches:

for $r_{i,j}$ to m'

for $g_{i,j}$ to n'

if $(B^3_M \le \Sigma_M P(r_{i,j}, g_{i,j}))$ //To sample pixels P

{

$Y_{edges} = P(edge|rg, \min_{DX} Y - DX^{\gamma_M}_F$ //for creating sparse matrix

else

$Y_{edges} = P(edge|rg, \min_{DX} Y - DX^{\gamma_M}_F$ //for Matching Contours

}

Step 3: Create pooling of contour entries and coded into contour networks

$$CCN(N) = \left[\sum_i^N Y_{edges} \sum_i^N X - \sum_i^N \min_{DX} DX^{\gamma_M}_F \right]$$

Step 4: End process.

Where, the value M represents the color model of the texture, which is embarked as low-intensity pixels after the preprocessing. γ_M and Σ_M are the mean and covariance of the pixel distribution based on intensities in rg color scheme after preprocessing. $Y_{edge} = [y1,y2,...yn]$ are the imagery patches from sampled block set $B_{\gamma M}$, $\|.\|_F$ represents Frobenius norm, D is the dictionary of all blocks D=[d1,d2,...dn] with size of 5 × 5 pixels and X is the sparse matrix X=[x1,x2,...xn]. Here, X is used to couple the several of the pooled contours.

The proposed approach encodes depth maps and is compared to other classical transformation based depth map compression algorithm like JPEG-2000 and platelet-based approaches. Here, the quality of encoded map is evaluated by comparing the reference image with the synthesized results and is determined as:
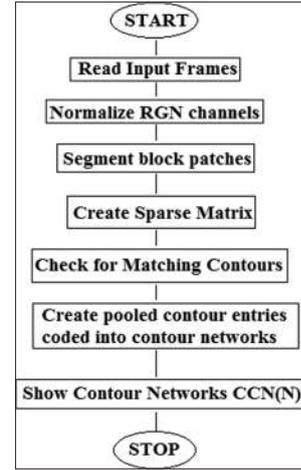

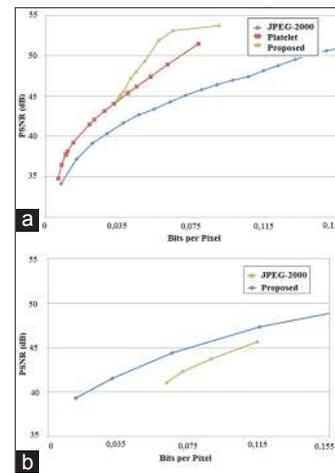**Fig. 2: Flow chart of the proposed algorithm depicting the workflow process**

Fig. 3: (a) Comparison plot of depth map compression rate-distortion curves for the three types of the compression schemes, (b) compression rate-distortion curves between the JPEDG - 2000 and the proposed depth map compression schemes for the same synthesized virtual viewpoint

a. Peak signal to noise ratio (PSNR) is used to gauge a rough idea of encoding quality. As the depth map resort to synthesize virtual viewpoints. As shown in Fig. 3a, the rate distortion curve during the compression for encoding the depth maps by the proposed method and that of cited in the literature reveals efficiency of 78% at about 0.1 bpp. At o.13 bpp, it outperforms the platelet-based technique by approximately 3.3dB PSNR value. However, the proposed approach yields a good result for medium bit rates.

b. Analyzing the synthesis result of compressed depth maps as shown in Fig. 3b. The difference in visual quality is noted in PSNR value. Here, the simple distortions in coded depth maps are marginally lower in comparison to the JPEG-2000 method.

CONCLUSIONS

In this study, we presented a novel algorithm for depth map compression for a video sequence which can be shared over the local network for performing global scale actuation in mobile robots. It is compared with the major depth map compression schemes to fulfill the major requirements for coding the contour networks. Upon compression, the efficiency of segmenting and pooling the network of contours gives a better compression rate than the previous methods. In addition, it encodes network of pooled contours which was not possible in previous techniques. We believe that the results put forth in this study

will serve latter stages of perceptual computing and organization for object recognition.

REFERENCES

1. Fehn C, Schuur K, Kauff P, Smolic A. Coding Results for EE4 in MPEG 3DAV. Vol. 9561. Pattaya: ISO/IEC JTC1/SC29/WG11 M; 2003.
2. Howard P, Kossentini F, Martins B, Forchhammer S, Rucklidge W. The emerging JBIG2 standard. IEEE Trans Circuits Syst Video Technol 2002;8(7):838-48.
3. Krishnamurthy R, Chai B, Tao H, Sethuraman S. Compression and transmission of depth maps for image-based rendering. In: Proceedings of International Conference on Image Processing. Vol. 3. IEEE; 2001. p. 828-31.
4. Liu S, Lai P, Tian D, Gomila C, Chen C. Sparse dyadic mode for depth map compression. In: 17th IEEE International Conference on Image Processing (ICIP). IEEE; 2010. p. 3421-4.
5. Morvan Y, de With P, Farin D. Platelet-based coding of depth maps for the transmission of multiview images. In: Proceedings of SPIE, Stereoscopic Displays and Applications. Vol. 6055; 2006. p. 93-100.
6. Sarkis M, Zia W, Diepold K. Fast depth map compression and meshing with compressed tritree. Computer Vision-ACCV 2009. Berlin: Springer; 2010. p. 44-55.
7. Shen G, Kim W, Ortega A, Lee J, Wey H. Edge-aware intra prediction for depth-map coding. In: Image Processing (ICIP), 2010 17th IEEE International Conference on IEEE; 2010. p. 3393-6.
8. Arbelaez P, Maire M, Fowlkes C, Malik J. Contour detection and hierarchical image segmentation. IEEE Trans Pattern Anal Mach Intell 2011;33(5):898-916.
9. Everingham M, Van Gool L, Williams CK, Winn J, Zisserman A. The PASCAL Visual Object Classes Challenge 2008 (VOC2008). Available from: http://www.pascalnetwork.org/challenges/VOC/voc2008.
10. Shotton J, Winn J, Rother C, Criminisi A. Textonboost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation. In: ECCV. Heidelberg: Springer; 2006.
11. Silberman N, Fergus R. Indoor scene segmentation using a structured light sensor. In: IEEE Workshop on 3D Representation and Recognition (3dRR); 2011.
12. Rai A. Shell implementation of neural net over the UNIX environment for file management: A step towards automated operating system. J Oper Syst Dev Trends 2014;1(2):10-4.
13. Rai A. Dynamic pagination for efficient memory management over distributed computational architecture for swarm robotics. J Adv Shell Program 2014;1(2):1-4.
14. Rai A. Attribute based level adaptive thresholding algorithm (ABLATA) for image compression and transmission. J Math Comput Sci 2014;12:211-8.