Emergence and Functionality of 3D Videos

Rameez Muneer Shamalik, S. M. Koli

Abstract: Millions of videos are created and watched every day. Quality of desired video plays a significant role in modern day applications. Deep fake videos have alarmed the established institutions while amused others. A Three-dimensional(3D) stereoscopic film or a 3D Video for that instance is a motion picture that enhances the illusion of depth perception, hence adding a third dimension. 3D videos can effortlessly validate and secure the authenticity of communication and broadcast process due to their auxiliary effect. In this research paper we attempt to observe and analyze numerous measures taken towards initiation and rise of 3D videos along with their efficacy from an end user perspective.

Index Terms: 3D Video, 3D video Coding and Processing, 3D-HEVC, Virtual Reality.

I. INTRODUCTION

Through recent technological breakthroughs and ease of Internet access, world is witnessing tremendous amount of data usage. Internet is broadly utilized for either information or communication, usually consisting of images as well as videos. New possibilities and challenges have aroused not only to handle and process data and videos but also to withdraw favorable deductions from them. The basic purpose behind Video processing or coding techniques are for video surveillance of traffic, factories, theft detection, for medical diagnosis, archeological activities and agricultural experiments.

Omni directional (360-degree) video applications are picking up pace in Video processing and communication. It finds its application in interactive shopping and movies, real time broadcast concerts etc. 360-degree is the main video format used in virtual reality which provides application like video games[1]. Depth image-based rendering (DIBR) also plays a vital role in Virtual Reality(VR)[2]. Systematically aligned and synchronized cameras capture scenes from multiple views in a single shot to provide texture based 3D videos[3]. Accordingly an ad-hoc network of cameras which is focused in same direction but located at different angles creates a formidable quality 3D video[4]. In order to achieve perception of 3D video without opting for special glasses a different stereo pair is assigned as approaching angle of spectator can change multiple times[5]. Light field displays and 3D displays can surely benefit edutainment. Also a shopper’s choice based marketing is being done as light field video-on-demand [6]. It’s quite feasible for any individual to get his hands on VR by using various Head Mounted Displays (HMD) such as Oculus Rift[7] HTC VIVE [8], to enjoy real life 3D effects. Subjective monitoring with the help of real observers provides best in class viewing impression but falls short on the test of time and being pocket friendly[9]. On the other hand, its objective counterpart which is purely based on mathematical modeling gives acceptable viewing experience prediction and is repeatable and fast as well[10].

Structure-from-Motion (SIM) is a popular for multiple angle views and 3D reconstruction of scalable areas [11].

Figure 1:[12] shows three key features to obtain 3D video in an effortless manner. Y axis has video as well as depth while x axis attributes the viewing directions namely Left, Center and Right. Authors introduce a brand new codec 3D-High Efficiency Video Coding(3D-HEVC) which works along with multi-layer coding design of HEVC. It gives Multi-view prognosis between the different angles of either video or depth. Interactive holographic byproduct can be seen by multi-component prognosis between the video and depth component of the same view and integrated multi-view or multi-component prognosis between video and depth of contrasting angles.

II. RELATED WORK

Jamali et al.[13] compared the execution of contrasting projections in 3D and multi-view bit rate which utilises disparity estimation that comes out dull when evaluated with simulcast coding as more time spent for inter-view prediction however it gives maximum compression but later stands comfortable not only in quality and coding efficiency but also in Complexity. Kim et. al[14] came up with a hole filling method which ensures global optimization and make use of the information from pre-filled coherent view and previous frame. An economical model for label propagation is extended which serves the purpose of maintaining the spatio-temporal consistency and symmetry between two coaxial frames in synthesized 3D videos at many virtual viewpoints. This model propagates measurable labels available in neighboring view and last frame for the target image to be overwhelmed for desired throughput.

Revised Manuscript Received on February 05, 2020.

Rameez Muneer Shamalik, Research Scholar, G. H. Raisoni College of Engineering and Management, Pune, India & Assistant Professor, Bharati Vidyapeeth’s College of Engineering for Women, Pune, India

Dr. S. M. Koli, Research Supervisor, G. H. Raisoni College of Engineering and Management, Pune India & Professor, Dept. of E&TC, Dr. D. Y. Patil School of Engineering, Charholi(Bk.), Pune, India
Emergence and Functionality of 3D Videos

Figure 2: [15] highlights an array of frames in which Malik et al. for achieving HEVC based spacial coverage scaling mixed resolution coding presents a prototype for multi view videos containing interleaved frames. Real size of basic frames is kept constant while information in neighboring frames of the center and coherent views are down-sampled. Muller et al.[16] came up with new 3D video standard(3D-HEVC). This codec serves as a supplementary addition to the highly used HEVC standard. This standard facilitates stereoscopic multi-view platforms to provide multi-view video plus depth. Concept like disparity-compensated prediction, inter-view motion parameter and residual prediction are further extended and combined for coding of the individual videos. The HEVC extension further inculcates depth coding, new intra coding modes, a modified motion compensation and motion vector coding also motion parameter inheritance.

Figure 3:[17] shows two variants of the same frame proposed by Ploshkin et. al for a temporal-shift estimation algorithm which highlights geometric distortions between stereoscopic views with sub frame accuracy. As it is assumed constant throughout each scene, it makes the process quite strong. Motion and Disparity maps are obtained using Block-Matching algorithm. Tamboli et. al[18] assessed the quality of visuals with respect to light field video frames. Same distortions to the source images were given while retaining the default parameter settings of the objective metric to evaluate the quality; they concluded that the distortions were separated in metric space decently.

A database is created by Wu et. al[19] with various formats of projections for Efficient VR quality assessment. subjective scores are given based on the assessment quality of self-built.

Table 1 Various techniques for 3D-HEVC[24]

Table 1 provides a bird’s eye view of the 3D-HEVC texture and depth coding techniques. enhanced compression operation is produced due to novel IV prediction schemes and texture coding tools.Many of the texture coding measures are deducted disparity for IV prognosis from different available samples of a pre-decoded depth layer.
III. RESULT AND DISCUSSION

We can work upon a non-linear regression model to produce a variety of parameters with extra coefficients. We can work on coefficients like gain, offset, and gamma using the given formula. 

\[ P_{\text{ref}} = C_{\text{gain}} \left( \frac{P_{\text{par}}}{2^{\text{bitdepth}} - 1} \right) + \gamma \left( \frac{2^{\text{bitdepth}} - 1}{2^{\text{bitdepth}}} \right) + C_{\text{offset}} \]

The results of which are shown in Table II[25]

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Red</th>
<th>Green</th>
<th>Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gain</td>
<td>0.5456</td>
<td>0.1136</td>
<td>0.3702</td>
</tr>
<tr>
<td>Offset</td>
<td>13.9680</td>
<td>0.46</td>
<td>0.37</td>
</tr>
<tr>
<td>Gamma</td>
<td>0.42</td>
<td>0.39</td>
<td>0.38</td>
</tr>
</tbody>
</table>

As shown above all three parameters are maximum for color red and in a reverse way minimum for color blue. Almost every screen which is used on daily basis gives us a 2D video; however 3D video adds depth and validity to the video processing and coding techniques. 3D glasses were compulsory to experience a 3D video but now state of the art cameras recording same frame with different angle is capable of delivering the same 3D effect to the end user. Available Video Processing techniques are mostly binary as in outcomes are in terms of either positive or negative. Techniques are complex including multidisciplinary fields, can be replaced by a simple model with clear objectives. Some of the techniques include training a Neural Network which can be a lengthy process. Datasets created for Processing can be divided according to application. Estimations are limited to certain human facial expression and gestures.

IV. CONCLUSION

360 degrees screen also facilitate the same effect with DIBR to give a Virtual reality experience. Deep learning algorithms efficiently takes care of depth in a video frame especially when a Neural network is trained for it. Neural networks can be bit complex at times and does not guarantee reliability on a repetitive basis. Block partitioning contribute to geometric realisation of 3D videos while prediction adds pace to the whole process. Both interview and inter-component prediction gives satisfactory results. Guassian Mean Models and its types are extensively used for proper estimation of videos. Video Processing is done on database available from various sources. Histograms also play a vital role for edge estimation of videos. New standards have been introduced for 3D video processing. In 3D Video Processing, major work is done on gesture and human activity detection. Depth and Geometry have added new paradigms for 3D video processing whereas prediction adds value to the whole process. For future scope real time processing based on direct application of 3D video processing can be proposed.

REFERENCES


Emergence and Functionality of 3D Videos


AUTHORS PROFILE

Rameez Muneer Shamalik, received his B. E. & M. E. degree from the Department of Electronics and Telecommunications Engineering of Savitribai Phule Pune University in 2008 & 2013 respectively. He has about ten years’ experience in teaching including research. He is currently pursuing Ph. D. with the department of Electronics and Telecommunication Engineering, G. H. Raisoni College of Engineering and Management, Pune India. He works as assistant professor in Bharati Vidyapeeth’s College of Engineering for Women, Pune India. His areas of interest are Signal processing and Wireless devices and communication.

Dr. S. M. Koli has received his B.E. degree from the Department of Electronics Engineering from Shivaji University, Kolhapur, Maharashtra, India in 1997. M. Tech. degree from the Department of Electronics and Telecommunication Engineering from Dr. Babasaheb Ambedkar Technological University, Lonere, Maharashtra, India in 2006 and Ph.D. degree in the Department of Engineering and Technology, from Sant Gadage Baba Amravati University, Amravati, Maharashtra, India in 2015. He is currently working as Professor in E&TC at the Department of Electronics and Telecommunication Engineering at Dr. D. Y. Patil School of Engineering, Pune, India and research supervisor at G. H. Raisoni college of Engineering and Management, Pune, India. His research interests are mainly in the wireless domain and video signal processing.