360° User-Generated Videos: Current Research and Future Trends

S. Priyadharshini and Ansuman Mahapatra

Abstract The 360° video, also known as immersive or spherical video, allows the observer to have a 360° view and an immersive experience of the surroundings. Each direction in this video is recorded at the same time either by an omni-direction camera or by an assembly of cameras synchronized together. The viewing perspectives are controlled by the viewer during playbacks. This article gives an overview of the existing research areas and methods in the user-generated 360° videos for streaming, transcoding, viewport-based projections, video standardization, and summarization. This survey also provides an analysis of the experience estimation in 360° videos. The study of multiple quality evaluation criteria is also reviewed. Moreover, 360° video user experience studies are also focused on this survey. The merits and demerits of each technique are investigated in depth.

1 Introduction

Cameras are affordable these days due to the technology advancements, which leads to a significant utilization of cameras by the users for capturing precious moments in their life. Omni-direction cameras can capture the whole scene using more than one camera and the images captured by these cameras are stitched together to give a 360° view of the scene.

Figure 1 portrays the basic workflow of 360° video. It generally commences with an omni-direction camera capturing 360° frames. Those are organized (i.e., stitched) together and sent to the encoding phase where the spherical video is projected to a 2D plane followed by frame packing and compression. The commonly used two projection formats: Equirectangular Projection (ERP) and Cubemap Projection (CMP) of

S. Priyadharshini · A. Mahapatra (✉)
Department of Computer Science and Engineering, National Institute of Technology, Puducherry, India
e-mail: ansuman.mahapatra@nitpy.ac.in

S. Priyadharshini
e-mail: priya81818@gmail.com

A. Nanda and N. Chaurasia (eds.), *High Performance Vision Intelligence*, Studies in Computational Intelligence 913,
https://doi.org/10.1007/978-981-15-6844-2_9

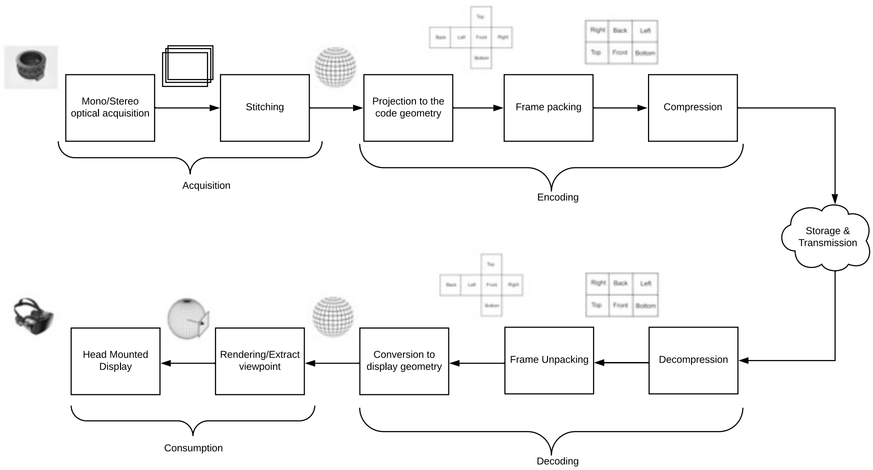


Fig. 1 360° video processing workflow [1]

Fig. 2 Equirectangular projection

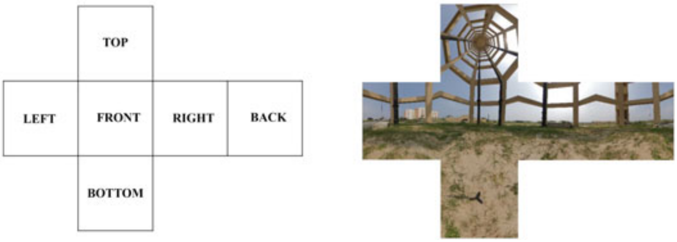
a user-generated 360° video are shown in Figs. 2 and 3, respectively. The encoding phase is followed by the decoding phase where a single video undergoes interactive projection that offers the rendering process inter-relating with the respective input/output technology (such as HMD) at the consumer end.

Figure 4 depicts the different FoVs in traditional viewing mode extracted from the equirectangular projection given in Fig. 2. This gives the content creators flexibility to shoot in 360° and later in the post-processing they can select the FoV that matters the most.

This review article

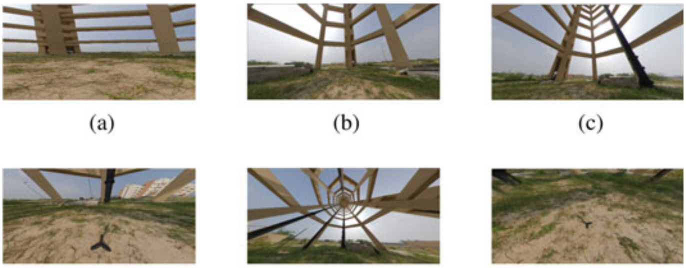
- is the first review on the user-generated 360° video to the best of our knowledge.
- introduces various research areas in the user-generated 360° video.
- investigates recent literature and categorizes based on research areas.
- highlights the pros and cons of each methodology.

The article is organized as follows. Section 2 briefs various research trends in 360° video production, communication, and analysis. The processing techniques applied



(a) Layout (b) An illustration

Fig. 3 Cubemap projection



(a) (b) (c) (d) (e) (f)

Fig. 4 **a–f** Different FoVs from user-generated 360° video

on 360° videos are discussed in Sect. 2.1. Section 2.2 discusses steaming techniques. Video post production methodologies are discussed in Sect. 2.3. The evaluation of the quality of 360° videos are reviewd in Sect. 2.4. Observations are listed in Sect. 3 and Sect. 4 concludes this article.

2 Research Trends in 360° Video

A brief survey of each research area in a 360° video is discussed in this section. Figure 5 depicts the research trends in 360° video.

2.1 Processing of 360° Video

This section discusses various processing techniques required for 360° videos before transmitting or storing. After capturing a 360° video, they need to be stitched and projected into a suitable representation, and then it will be compressed for transmis-

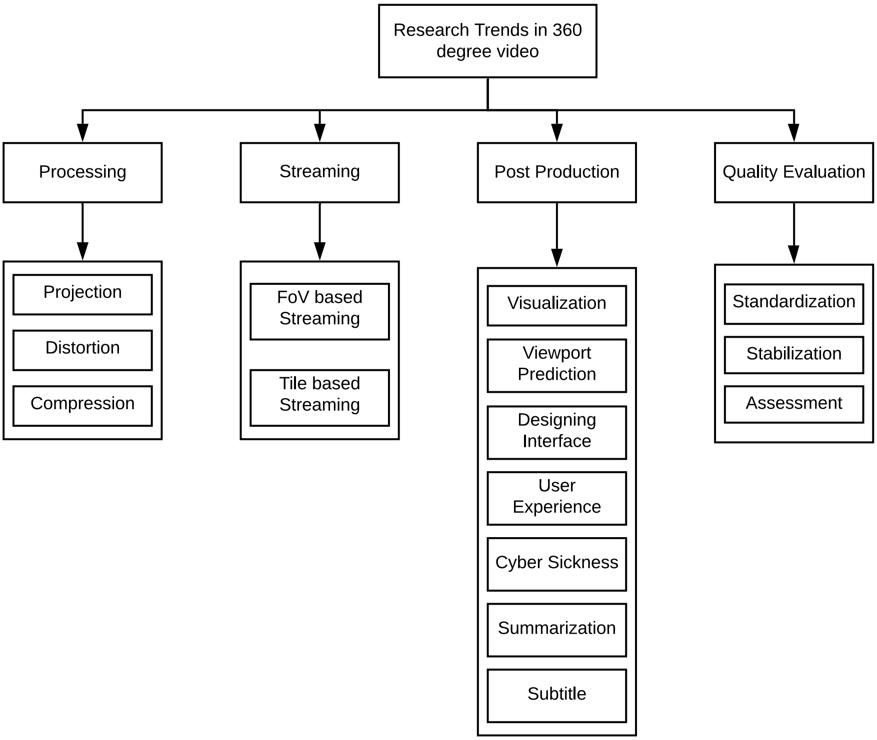


Fig. 5 Areas of research in 360° video

sion or storage. The following subsections present a review of the existing methods in processing 360° video.

2.1.1 Projection

In Sphere Segmented Projection, the visual artifact is caused due to inactive region [2]. In order to enhance coding efficiencies and to minimize visual artifacts, Yoon et al. suggest a scheme of padding inactive region. For panoramic videos, Huang et al. presented a low-complexity prototype scheme and video stitching mechanism [3]. Hanhart et al. recommended a coded approach on the basis of spherical neighboring relationship and projection form adaptation [4]. Su and Grauman proposed a spherical convolutional network used to process 360° imagery straightforward in its equirectangular projection, which is translated from a planar Convolutional Neural Network (CNN) [5]. Lin et al. propose a hybrid equiangular cubemap projection that minimizes seam artifacts [6]. Some characteristic equirectangular projection forms of sequences in the clip are experimented by Wang et al. [7].

It is unfavorable to attain a well-organized compression for storing and transmitting [8]. Hence, Vishwanath et al. recommended a rotational model for identifying the angular motion on the sphere effectively. In 3D space, for an angle α, vector A is rotated around an axis given by a unit vector B. The coordinates of vectors A and B are (p, q, r) and (l, m, n), respectively. The coordinates of the rotated vector $A^{'}$ will be

$$p^{'} = l(B \cdot A)(1 - \cos\alpha) + p\cos\alpha + (-nq + Ar)sin\alpha \tag{1}$$

$$q^{'} = m(B \cdot A)(1 - \cos\alpha) + q\cos\alpha + (np - lr)sin\alpha \tag{2}$$

$$r^{'} = n(B \cdot A)(1 - \cos\alpha) + r\cos\alpha + (-Ap + lq)sin\alpha \tag{3}$$

where $B \cdot A$ is the dot product. Rotation of axis B is the vector right angled to the plane well defined through the origin, vector A, and also rotated vector $A^{'}$. Vector B is computed as follows:

$$B = \frac{A \times A^{'}}{|A \times A^{'}|} \tag{4}$$

Angle of rotation is given as

$$\alpha = \cos^{-1}(A.A^{'}) \tag{5}$$

The summary of techniques, highlights, and challenges of 360° video projections is listed in Table 1.

2.1.2 Distortion

Azevedo et al. provide an extensive analysis of the most popular visual deformity that alters the 360° video signals in immersive applications [1]. Aksu et al. present a scalable multicast live deliver of 360° video with distortion analysis [9]. Yoon et al. recommend an approach of adding inactive regions to lessen deformations [2]. A detailed review of the distortions in 360° video is given in Table 2.

2.1.3 Compression

Le et al. designed a transcoding system with ARIA block cipher for encoding purpose [10]. To gain high steady sampling, Lin et al. offer 360° specific coding tools [6]. The mapping function is given as follows:

Table 1 Summary on projection of 360° video

Author	Technique	Highlights	Challenges
Yoon et al. [2]	Method of padding inactive Regions	Significantly lessens the visual artifact	–
Huang et al. [3]	Procedure for video stitching	Minimizes computation	Finding the best seam has a single constraint
Hanhart et al. [4]	Approach on projection forms	Reduces the face seam artifacts	Dependent on the frame packing configuration
Su and Grauman [5]	Spherical convolutional network	Yields accurate results	Further exploration of spherical convolution is required
Lin et al. [6]	Hybrid equiangular cubemap projection scheme	Achieves increased uniform sampling	–
Wang et al. [7]	Sphere-shaped Coordinates Transform Placed Mobility Method (SCTMM)	Saves significant bits for video sequence	Complexity computation is very high
Vishwanath et al. [8]	Prototype intended for rotational movement	Globally suitable for all the projection geometries	–

Table 2 Summary on distortion of 360° video

Author	Technique	Highlights	Challenges
Azevedo et al. [1]	Investigation on visual distortions	Detects the reasons for deformations	–
Aksu et al. [9]	Distortion analysis	Aggregate distortion is minimized	Decrease in encoding gain is not focussed
Yoon et al. [2]	Method of padding inactive regions	Visual artifacts was minimized	–

2D (Cube-Map) to Sphere:

$$h_b(a, b) = \frac{b}{1 + 0.4(1 - a^2)(1 - b^2)} \tag{6}$$

Sphere to 2D (Cube-Map):

$$k_b(a, b) = \begin{cases} b, & \text{if s=0} \\ \frac{1-\sqrt{1-4s(b-s)}}{2s}, & \text{otherwise} \end{cases} \tag{7}$$

$$\text{where } s = 0.4b(k_b(a)^2 - 1).$$

To improve the quantity of storing and compressing video based on perception, an efficient compression mechanism called Vignette was suggested by Mazumdar et al. [11]. Xiu et al. recommended that gain for the paired categories of video such as HDR and SDR attains considerable efficiency on coding [12]. Aimed at the spherical environment, Wang et al. propose an algorithm for compensation and estimation based on the motion prototype [7].

In order to enhance efficiency and minimize encoding time, Zhang et al. present an optimization procedure on compression [13]. Choi et al. offer an inventive video compression approach for video service accompanied by high quality and video coding schemes using HDR [14]. Lin et al. propose a subject labeled database with a massive scale, which comprises compressed H.265/HEVC videos consisting of miscellaneous PEAs [15]. In order to have an enhanced performance, Le et al. designed a transcoding system that plays a vital role in modifying bit rates and changing the resolution of 360° videos [10]. Various 360° video compression techniques, highlights, and challenges are summarized in Table 3.

2.2 *Streaming of 360° Video*

This section presents various mechanisms required for 360° video streaming. Streaming can be done based on FoV or Tiles. The following subsection gives a detailed description of the techniques involved in FoV-based and Tile-based streaming.

2.2.1 FoV-Based Streaming

Duanmu et al. established a two-tier framework to intensify the utilization of bandwidth for 360° video streaming [16]. Skupin et al. propose an optimal way of streaming based on the FoV [17]. Sun et al. propose a two-tier solution to deliver the entire 360° span video at a truncated quality base tier and a higher quality enhancement tier [18]. Jiang et al. recommended Plato for viewport adaptive streaming using reinforcement learning [19].

Qian et al. introduce a cellular-friendly streaming methodology which conveys only 360° video viewport created on the prediction of head actions [20]. The 360° video stream has greater bandwidth requirements and needs quicker responsiveness to viewers' inputs [21]. In this aspect, Zhou et al. perform an analysis of oculus 360° video streaming. Among the future and past viewpoints, in order to capture the long-term dependent and nonlinear relation Yang et al. presented a single viewpoint prediction model built on CNN [22]. Corbillon et al. give a viewpoint adaptive approach that allows the streaming video to have a lower bit rate on comparison with the original video [23]. Table 4 gives the review on FoV-based streaming of 360° video.

Table 3 Summary on compression of 360° video

Author	Technique	Highlights	Challenges
Le et al. [10]	Transcoding and encryption methodology	Security is enhanced	–
Lin et al. [6]	360° certain coding tools	Highly effective in representing 360° videos	–
Mazumdar et al. [11]	Integrated perceptual compression approach	Maintaining the perceptual quality	Cost of compression is high
Xiu et al. [12]	Compression based on video coding methodology	Gains coding efficiency	Significant decoding complexity is increased
Wang et al. [7]	SCTMM with JEM scheme	SCTMM takes less time for decoding	Cost of SCTMM encoding is high
Zhang et al. [13]	Optimized compression algorithm	Improves efficiency	CU partition algorithm can be underestimated
Choi et al. [14]	High-quality video compression technique	Provides ordered representation of coding in a flexible manner	–
Lin et al. [15]	Containing various PEAs in large-scale database	Motivates perceptual video encoding mechanisms	–
Le et al. [10]	Real-time transcoding system	The two 4 K sessions and six 1080p sessions are optimized	–

2.2.2 Tile-Based Streaming

Sanchez et al. illustrate the streaming established by means of tile tactics followed in the Moving Picture Expert Group OMAF requirement [24]. Xie et al. presented a compatible streaming model for probabilistic tiles referred as 360ProbDASH [25]. Graf et al. propose adaptive tile-based streaming over HTTP to present the solution for the problems faced in video delivery infrastructures [26].

As the complexity of the 360° video increases with the essential to accomplish bitrate adaptation for a varying network [27], Le Feuvre and Concolato recommended MPEG DASH (Dynamic Adaptive-Streaming over HTTP) standard to designate by what means spatial accessing can be attained. Kammachi-Sreedhar and Curcio described an optimal way of streaming technology [28].

Nguyen et al. suggest a flexible method for tiling-based viewpoint streaming [29]. Due to the latency in the network, 360° video streaming is a difficult task [30]. Hence, Mahzari et al. recommended a tile-based caching policy. In real life using cellular networks, tiled video develops a probable solution for violently minimizing the essential bandwidth for 360° video transmission [31]. As a result, Lo et al. give

Table 4 Summary on FoV-based streaming of 360° video

Author	Technique	Highlights	Challenges
Duanmu et al. [16]	Two-tier mechanism	Achieves 25% gain	Bandwidth analysis is not focussed
Skupin et al. [17]	FoV-dependent optimal streaming	Enables ranking the FoV quality	Concurrent decoding for particular FoV is not done
Sun et al. [18]	Two-tier-based streaming scheme	Predicts FoV with high accuracy	–
Jiang et al. [19]	Viewport adaptive streaming	Qualities are adjusted based on FoV	–
Qian et al. [20]	Cellular-friendly 360° video delivery strategy	The client only yields the portions that are observable	To assess for a huge scale FoV, user study is challenging
Zhou et al. [21]	Analysis of Oculus 360° video streaming	Provides improved FoV visual quality	Improved performance is not analyzed
Yang et al. [22]	Single viewpoint prediction model	Enhances accuracy of FoV prediction	Spherical CNN is not explored in depth
Corbillon et al. [23]	Viewport-based adaptation algorithm	Viewport-dependent QER selection is done precisely	To predict the head movement is tedious

the performance over a cellular network of tile-based streaming. For high-quality streaming, there is a limitation of power consumption and bandwidth effectiveness [32]. Hence, Son et al. offer a tiling-based streaming approach. Summary on tile-based streaming of 360° video is shown in Table 5.

2.3 *Post-production of 360° Video*

At the user end, post-processing of the stored or streamed content is done. It provides consumer ease in comprehension, seamless visualization, and user experience. Several methods for post-production have been discussed under the following subsections.

2.3.1 Visualization

On live broadcasting, the broadcaster may not be aware of the user's FoV [33]. In this aspect, Takada et al. propose a visualization method based on users' Points of View (PoV) making use of a spherical heat map allowing the broadcaster to grip users' FoV easily and exchange information with users evenly. Azevedo et al. alter the 360° video signals for better visualization in immersive applications [1]. Existing

Table 5 Summary on tile-based streaming of 360° video

Author	Technique	Highlights	Challenges
Sanchez et al. [24]	Tile-based streaming approach	Observed fidelity was reduced	End-to End delay is critical
Xie et al. [25]	Probabilistic scheme for tiles	Provides contiguous playback	Saliency model needs to be precise
Graf et al. [26]	Adaptive tile-based streaming over HTTP	Evaluation of streaming is performed	–
Le Feuvre and Concolato [27]	MPEG DASH standard	Highly Interactive spatial navigation is achieved	Collaborative tile selection is not concentrated
Kammachi Sreedhar and Curcio [28]	Streaming technology based on adaptive bit rate	Multiple tiles synchronization	–
Nguyen et al. [29]	Tile-based viewpoint streaming	Tiles quality is improved	Optimality is not checked
Mahzari et al. [30]	Tile-based caching policy	Performance of cache is better	–
Lo et al. [31]	Tile-based streaming	Gives improved awareness	Transfer time depends on tile size
Son et al. [32]	Tile-based streaming	Transmit tiles autonomously	Lesser efficient

Table 6 Summary on visualization of 360° video

Author	Technique	Highlights	Challenges
Takada et al. [33]	Users PoV visualization method	Improves communication accuracy between users and senders	Functional improvement is required
Azevedo et al. [1]	Investigation on visualization	Proper psycho-visual examine on immersive applications is achieved	–

techniques, highlights, and challenges of visualization in 360° video are summarized in Table 6.

2.3.2 Viewport Prediction

User head movements result in user interaction and modifications in the spatial parts of the video allowing them to view only essential portions in the video for a specified time [9]. To achieve this, Aksu et al. offered a novel adaptable framework for the

Table 7 Summary on viewport prediction of 360° video

Author	Technique	Highlights	Challenges
Aksu et al. [9]	Viewport prediction with multicast live streaming scheme	Quality of viewport is maximized	–
Heyse et al. [34]	CB-based learning approach	Enhances Viewport quality	Further Enhancement is required
Jiang et al. [19]	Long short-term memory model	Future PoV is predicted early	–
Hu et al. [35]	Automated 360° piloting	Gives supreme performance	–
Sanchez et al. [24]	Viewpoint-dependent approach	Proves better visual resolutions	Significant gains can still be achieved
Li et al. [36]	Two groups prediction model	SD and mean of future PoV are predicted	–

prediction of the viewport. Heyse et al. offered an approach for contextual bandit based on reinforcement learning [34]. The tiles which map the field of view, provided with high resolution by using viewpoint adaptive streaming, was proposed by Jiang et al. [19]. Hu et al. recommended a mechanism of agent-based deep learning “deep 360° pilot” for viewers to pilot the 360° sports video spontaneously and develops an agent-specified domain to have a clear definition about the objects in the video [35].

To analyze visual quality at the viewport based on end-to-end delay, a viewpoint-dependent scheme was proposed by Sanchez et al. with the gain of 46% when compared with viewpoint-independent scheme [24]. Foreseeing the future PoV in a long time horizon can help in saving bandwidth incomes for on-request streaming of a video in which pausing of the video is diminished with noteworthy bandwidth variations in network [36]. To support this Li et al. introduced a two clusters point of view prediction models. Table 7 summarizes the viewport prediction of 360° video.

2.3.3 Designing Interface

Pavel et al. presented a technique based on the interactive orientation of shots enabling users to view all the significant content in the film [37]. Poblete et al. proposed a scalable appeal of design on crowdsourced technique [38]. Tang and Fakourfar supported collaborative perspective and interaction through proper awareness on gaze and technique on gesture for 360° videos [39]. The designing interfaces of 360° video are completely reviewed in Table 8.

Table 8 Summary on designing interface of 360° video

Author	Technique	Highlights	Challenges
Pavel et al. [37]	Interactive orientation of shots	Allows to choose significant points	Fatigue effect is not investigated
Poblete et al. [38]	Scalable appeal on crowdsource	Examines multiple fields of view	Lack of stitching multiple videos
Tang and Fakourfar [39]	Demonstrates the current view of interfaces	Powerful mechanism to be familiar with one another's view	Challenges faced are awareness of gaze and displaying in HMD

Table 9 Summary on user experience of 360° video

Author	Technique	Highlights	Challenges
Broeck et al. [40]	Numerous interaction methodologies	Highly ranking visualized experience	Results in motion sickness
Lin et al. [41]	Two focus guidance mechanism	Focus ease is improved	Multiple targets are not intensified
Nasrabadi et al. [42]	Taxonomy on 360° videos	Analysis with the varied clusters of users	Consequence of objects in motion is not studied in depth

Table 10 Summary on cybersickness of 360° video

Author	Technique	Highlights	Challenges
Bala et al. [43]	Study on existing methodologies	Guards from visually induced motion	Statistical implication is not encountered

2.3.4 User Experience

Broeck et al. proposed a numerous interaction methodology [40]. One task of looking at 360° videos is endlessly focusing and refocusing intentional targets [41]. To overcome this, Lin et al. addressed an approach on two focus guidance such as Automatic Piloting (directly taking audiences to the goal) and Visual Supervision (representing track of the goal). Nasrabadi et al. proposed taxonomy on 360° videos and classified them based on the motion of the camera and object [42]. Existing 360° video user experience techniques, highlights, and challenges are reviewed in Table 9.

2.3.5 Cybersickness

Bala et al. proposed an investigational study toward comparing and joining numerous available methodologies in 360° video to minimize cybersickness [43]. Cybersickness of 360° video is summarized in Table 10.

Table 11 Summary on summarization of 360° video

Author	Technique	Highlights	Challenges
Sung et al. [44]	Prototype based on memory network	Discourses narrative time-related summarization	–

Table 12 Summary on subtitle of 360° video

Author	Technique	Highlights	Challenges
Brown et al. [45]	Behaviors of subtitle	Answers each behavior usage	Diverse styles are not experimented

2.3.6 Summarization

For a long 360° videos, Sung et al. addressed the issue of story-based time-oriented summarization [44]. An innovative prototype based on memory network (Past Future Memory Network) was proposed. Available techniques, highlights, and challenges about summarization of 360° video are listed in Table 11.

2.3.7 Subtitle

Brown et al. designate behaviors of four subtitle (120-degree, static-follow, lag-follow, appear) in order to accomplish user testing in 360° video experience [45]. A detailed review of the subtitle of 360° video is illustrated in Table 12.

2.4 Quality Evaluation of 360° Video

This section gives the literature review on assessing the quality of the user-generated 360° videos. Some of the existing works have been listed in the following subsections.

2.4.1 Standardization

Wien et al. addressed the current status of standardization on focus with scientific aspects associated with the video [46]. Hannuksela et al. give an outline of the foremost edition of the standards in OMAF [47]. Skupin et al. presented the details regarding up-to-date status of precise efforts available in standardization [17]. Azevedo et al. offered some standardization techniques [1]. Domanski et al. proposed different kinds of visual media that are highly immersive [48]. Table 13 describes the standardization of 360° video techniques, highlights, and challenges.

Table 13 Summary on standardization of 360° video

Author	Technique	Highlights	Challenges
Wien et al. [46]	Outlines standardization efforts	Provides standards for coding and transmission	Standards for 6DoF remain unexplored
Hannuksela et al. [47]	Summary of the initial issue in OMAF	Embraces the representation of video setups, OMAF video profiles	Includes no abundant facts on image, audio, and text
Skupin et al. [17]	Describes the standardization status	Reports on the current status	–
Azevedo et al. [1]	Investigation of the most popular visual distortions	Acts as the basis for standardization	–
Domanski et al. [48]	Immersive standardization of visual media	Identify the immersive levels that are attained	Standardizing free perspectives are absent

Table 14 Summary on stabilization of 360° video

Author	Technique	Highlights	Challenges
Kopf [49]	Hybrid 2D-3D procedure	Gives improved smoothness	–
Tang et al. [50]	Combined stabilization approach	Demonstrates the stabilization of observing experience	–

2.4.2 Stabilization

Kopf offers a hybrid 2D-3D procedure for 360° video stabilizing by means of a deformed rotationally moving model [49]. Tang et al. introduce an approach for combined stabilization with the direction of 360° videos [50]. It includes a precisely designed new motion determination technique for 360° videos. Stabilization of 360° video is summarized in Table 14.

2.4.3 Assessment

Huang et al. support evaluation of video quality and propose a visual attention model for latitude-based 360-degree videos [51]. Hanhart et al. aim at the quality evaluation scheme recognized by JVET of ITU-T VCEG and ISO/IEC MPEG [52]. Zakharchenko et al. discussed the immersion media delivery format and quality assessment process [53]. Tran et al. investigated the quality benchmark of both subjective and objective for 360° videos [54]. For 360° video communication, Tran et

Table 15 Summary on assessment of 360° Video

Author	Technique	Highlights	Challenges
Huang et al. [51]	Visual attention model	Exploits the mean attention	Quality estimation in a large database is not investigated
Hanhart et al. [52]	Outlines the quality assessment framework	Gives an overview on quality assessment	–
Zakharchenko et al. [53]	Position-invariant quality metrics	Accurate and reliable	–
Tran et al. [54]	Quality benchmark	Assess the perceived quality	–
Tran et al. [55]	Study on the quality relationship	To detect suitable objective quality	Quality assessment on adaptive-based transmission is not focussed
Xie et al. [25]	QoE-based adaptation system	Variance of quality is minimized	–
Jiang et al. [19]	QoE metrics	Improvement on QoE	–
Corbillon et al. [23]	Multiple QER-based representation of a clip	Measurement of QoE only at the extracted PoV	–

al. aid to recognize suitable objective quality benchmark [55]. Xie et al. presented a QoE-based optimization framework [25]. Jiang et al. suggested Plato that outperforms existing strategy in numerous QoE metrics [19]. Corbillon et al. recommended an interactive high-quality mechanism for QoE measurement in Head-Mount-Device audience with small supervision [23]. Table 15 gives the quality assessment of 360° video.

3 Observation

The following are the observations made through this study:

- The 360° video is gaining interest among the consumers due to its simplicity.
- During projection there may be a chance of occurring visual artifacts which are also termed as distortion. Hence, extra caution has to be taken during the process of projection to learn the contents of the clip fruitfully.
- Once the 360° video is projected they undergo coding in order to have efficient storage and transmission where the 360° videos are compressed by preserving the quality of the video.
- As the viewpoint increases, the concept of streaming becomes difficult. Hence, an efficient approach of streaming the 360° video has to be done with better visual qualities.

- The high immersive nature of the video should not lead to motion sickness.
- 360° video can be delivered to the user optimally by summing up all the significant informations that are available in the clip.
- In order to have a clear understanding of the information available in the 360° clip, the video can be streamed with closed caption (i.e., text form).
- On the aspect of maximizing the smoothness of visual quality, the video can be stabilized.
- At the user end, the quality of the video can be checked by using the quality metrics.

4 Conclusion

360° video can offer an immersive experience for the users. As the FoV in 360° video increases in comparison with the standard videos, they encompass a huge amount of information. Due to the high resolution, 360° video processing, transmission, and displaying have to be done efficiently. This article presents the various techniques, highlights, and challenges involved in processing, transmission, and displaying the 360° video. At the viewer end, the decoded video has to be checked for its standardization, stabilization, and the quality of experience, to analyze the video for high standards, increased immersion, and improved QoE, respectively. The various techniques involved in the mechanism of standardization, stabilization, and QoE are listed in this survey with its highlights and challenges. The overall challenges faced in 360° videos are the high rate of compression and improvement in the quality-based viewport prediction.

On the aspect of future trends, the 360° videos are growing at a faster pace. In the near future, this technology will experience a huge leap. The major role of 360° video is storytelling with an immersive environment. Further improvement in terms of the cost may be possible in the coming years to give users an immersive experience. Faster improvement in the 360° technology and inexpensiveness of the equipment makes the 360° video to spread swiftly across many industries in the near future. In the upcoming years, for high-end performance, the 360° technology will provide a high level of video capturing with a High Dynamic Range (HDR).

References

1. R.G. de A. Azevedo, N. Birkbeck, F. De Simone, I. Janatra, B. Adsumilli, P. Frossard, Visual distortions in 360-degree videos, arXiv:1901.01848 (2019)
2. Y.-U. Yoon, D.-H. Park, J.-G. Kim, A method of padding inactive region for sphere segmented projection of 360° video, in *2018 International Workshop on Advanced Image Technology (IWAIT)*. IEEE (2018), pp. 1–3

3. K.-C. Huang, P.-Y. Chien, C.-A. Chien, H.-C. Chang, J.-I. Guo, A 360-degree panoramic video system design. Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test, IEEE (2014), pp. 1–4
4. P. Hanhart, X. Xiu, Y. He, Y. Ye, 360-degree video coding based on projection format adaptation and spherical neighboring relationship, in *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, IEEE (2018)
5. Y.-C. Su, K. Grauman, Learning spherical convolution for fast features from 360 imagery, in *Advances in Neural Information Processing Systems* (2017), pp. 529–539
6. J.-L. Lin, Y.-H. Lee, C.-H. Shih, S.-Y Lin, H.-C. Lin, S.-K. Chang, P. Wang, L. Lin C.-C. Ju, Efficient Projection and coding tools for 360° Video. IEEE J. Emerging Selected Topics Circuits Syst. IEEE (2019)
7. Y. Wang, D. Liu, S. Ma, F. Wu, W. Gao, Spherical coordinates transform- based motion model for panoramic video coding. IEEE J. Emerg. Selected Topics Circuits Syst. IEEE (2019)
8. B. Vishwanath, T. Nanjundaswamy, K. Rose, Rotational motion model for temporal prediction in 360 video coding, in *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*. IEEE (2017), pp. 1–6
9. R. Aksu, J. Chakareski, V. Swaminathan, Viewport-driven rate-distortion optimized scalable live 360° video network multicast, in *2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE (2018), pp. 1–6
10. T. Thanh Le, J.B. Jeong, E.-S. Ryu, Efficient transcoding and encryption for live 360 CCTV system. Applied Sciences, 9, 4, 760, Multidisciplinary Digital Publishing Institute (2019)
11. A. Mazumdar, B. Haynes, M. Balazinska, L. Ceze, A. Cheung, M. Oskin, Vignette: perceptual compression for video storage and processing systems, arXiv:1902.01372 (2019)
12. X. Xiu, Y. He, Y. Ye, R. Vanam, H. Philippe, T. Lu, E. Pu, P. Yin, W. Husak, T. Tao., Improved video coding techniques for next generation video coding standard, in *2019 Data Compression Conference (DCC)*. IEEE (2019), pp. 290–299
13. M. Zhang, J. Zhang, Z. Liu, C. An, An efficient coding algorithm for 360-degree video based on improved adaptive qp compensation and early cu partition termination, in *Multimedia Tools and Applications*, 78, 1. Springer (2019), pp. 1081–1101
14. K. Choi, J. Chen, A. Tamse, H. Yang, M.W. Park, S. Ikonin, W. Choi, S. Esenlik, New video codec for high- quality video service and emerging applications, in *2019 Data Compression Conference (DCC)*. IEEE (2019), pp. 310–319
15. L. Lin, S. Yu, T. Zhao, Z. Wang, others, PEA265: Perceptual assessment of video compression artifacts, arXiv:1903.00473 (2019)
16. F. Duanmu, E. Kurdoglu, Y. Liu, Y. Wang, View direction and bandwidth adaptive 360 degree video streaming using a two-tier system, in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE (2017), pp. 1–4
17. R. Skupin, Y. Sanchez, Y.-K. Wang, M.M. Hannuksela, J. Boyce, M. Wien, Standardization status of 360 degree video coding and delivery, in *2017 IEEE Visual Communications and Image Processing (VCIP)*. IEEE (2017), pp. 1–4
18. L. Sun, F. Duanmu, Y. Liu, Y. Wang, Y. Ye, H. Shi, D. Dai, A two- tier system for on- demand streaming of 360 degree video over dynamic networks. IEEE J. Emerg. Selected Topics Circuits Syst. IEEE (2019)
19. X. Jiang, Y.-H. Chiang, Y. Zhao, Y. Ji, Plato: learning- based adaptive streaming of 360-degree videos, in *2018 IEEE 43rd Conference on Local Computer Networks (LCN)*. IEEE (2018), pp. 393–400
20. F. Qian, L. Ji, B. Han, V. Gopalakrishnan, Optimizing 360 video delivery over cellular networks, in *Proceedings of the 5th Workshop on All Things Cellular: Operations, Applications and Challenges*. ACM (2016), pp. 1–6
21. C. Zhou, Z. Li, Y. Liu, A measurement study of oculus 360 degree video streaming, in *Proceedings of the 8th ACM on Multimedia Systems Conference*. ACM (2017), pp. 27–37
22. Q. Yang, J. Zou, K. Tang, C. Li, H. Xiong, Single and sequential viewports prediction for 360-degree video streaming, in *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE (2019), pp. 1–5

23. X. Corbillon, G. Simon, A. Devlic, J. Chakareski, Viewport- adaptive navigable 360-degree video delivery, in *2017 IEEE International Conference on Communications (ICC)*. IEEE (2017), pp. 1–7
24. Y. Sanchez, G.S. Bhullar, R. Skupin, C. Hellge, T. Schierl, Delay Impact on MPEG OMAF's tile-based viewport- dependent 360° video streaming. IEEE J. Emerg. Selected Topics Circuits Syst. IEEE (2019)
25. L. Xie, Z. Xu, Y. Ban, X. Zhang, Z. Guo, 360probdash: improving qoe of 360 video streaming using tile-based Http adaptive streaming, in *Proceedings of the 25th ACM international conference on Multimedia*. ACM (2017), pp. 315–323
26. M. Graf, C. Timmerer, C. Mueller, Towards bandwidth efficient adaptive streaming of omnidirectional video over Http: design, implementation, and evaluation, in *Proceedings of the 8th ACM on Multimedia Systems Conference*. ACM (2017), pp. 261–271
27. J. Le Feuvre, C. Concolato, Tiled- based adaptive streaming using MPEG-DASH, in *Proceedings of the 7th International Conference on Multimedia Systems*. ACM (2016), p. 41
28. K. Kammachi-Sreedhar, I.D.D. Curcio, Omnidirectional video delivery with decoder instance reduction, in *Internet Technology Letters*, vol. 2, 1, e79, Wiley Online Library (2019)
29. D.V. Nguyen, H.T.T. Tran, T.C. Thang, Adaptive tiling selection for viewport adaptive streaming of 360-degree video, in *IEICE Transactions on Information and Systems*, vol. 102, 1. The Institute of Electronics, Information and Communication Engineers (2019), pp. 48–51
30. A. Mahzari, A. Taghavi Nasrabadi, A. Samiei, R. Prakash, Fov- aware edge caching for adaptive 360 video streaming, in *2018 ACM Multimedia Conference on Multimedia Conference*. ACM (2018), pp. 173–181
31. W.-C. Lo, C.-L. Fan, S.-C. Yen, C.-H. Hsu, Performance measurements of 360° video streaming to head-mounted displays over live 4G cellular networks, in *2017 19th Asia-Pacific Network Operations and Management Symposium (APNOMS)*. IEEE (2017), pp. 205–210
32. J. Son, D. Jang, E.-S. Ryu, Implementing 360 video tiled streaming system, in *Proceedings of the 9th ACM Multimedia Systems Conference*. ACM (2018), pp. 521–524
33. M. Takada, D. Nishioka, Y. Saito, Proposal of a spherical heat map in 360-degree internet live broadcasting using viewers' POV, in *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE (2019), pp. 596–600
34. J. Heyse, M.T. Vega, F. De Backere, F. De Turck, Contextual bandit learning-based viewport prediction for 360 video, in *IEEE Virtual Reality (VR)* (2019)
35. H.-N. Hu, Y.-C. Lin, M.-Y. Liu, H.-T. Cheng, Y.-J. Chang, M. Sun, Deep 360 pilot: learning a deep agent for piloting through 360 sports videos, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE (2017), pp. 1396–1405
36. C. Li, W. Zhang, Y. Liu, Y. Wang, Very long term field of view prediction for 360-degree video streaming, in *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE (2019), pp. 297–302
37. A. Pavel, B. Hartmann, M. Agrawala, Shot orientation controls for interactive cinematography with 360 video, in *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM (2017), pp. 289–297
38. B.M. Poblete, E.C. Mendoza, J.P. De Castro, J.A. Deja, G. Nodalo, A research through design (Rtd) approach in the design of a 360-video platform interface, in *Proceedings of the 5th International ACM In-Cooperation HCI and UX Conference*. ACM (2019), pp. 166–171
39. A. Tang, O. Fakourfar, Watching 360 videos together, in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM (2017), pp. 4501–4506
40. M.V. den Broeck, F. Kawsar, J. Schöning, It's all around you: exploring 360 video viewing experiences on mobile devices, in *Proceedings of the 25th ACM international conference on Multimedia*. ACM (2017), pp. 762–768
41. Y.C. Lin, Y.-J. Chang, H.-N. Hu, H.-T. Cheng, C.-W. Huang, M. Sun, Tell me where to look: investigating ways for assisting focus in 360 video, in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM (2017), pp. 2535–2545
42. A.T. Nasrabadi, A. Samiei, A. Mahzari, R.P. McMahan, R. Prakash, M.C.Q. Farias, M.M. Carvalho, A taxonomy and dataset for 360° videos, in *Proceedings of the 10th ACM Multimedia Systems Conference*. ACM (2019), pp. 273–278

43. P. Bala, D. Dionísio, V. Nisi, N. Nunes, Visually induced motion sickness in 360° videos: comparing and combining visual optimization techniques, in *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE (2018), pp. 244–249
44. S. Lee, J. Sung, Y. Yu, G. Kim, A memory network approach for story-based temporal summarization of 360 videos, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 1410–1419
45. A. Brown, J. Turner, J. Patterson, A. Schmitz, M. Armstrong, M. Glancy, Subtitles in 360-degree video, in *Adjunct Publication of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*. ACM (2017), pp. 3–8
46. M. Wien, J.M. Boyce, T. Stockhammer, W.H. Peng, Standardization status of immersive video coding. IEEE J. Emerg. Selected Topics Circuits Syst., IEEE (2019)
47. M.M. Hannuksela, Y.K. Wang, A. Hourunranta, An overview of the OMAF standard for 360° video, in *2019 Data Compression Conference (DCC)*. IEEE (2019), pp. 418–427
48. M. Domański, O. Stankiewicz, K. Wegner, T. Grajek, Immersive visual media- mpeg-i: 360 video, virtual navigation and beyond, in *2017 International Conference on Systems, Signals and Image Processing (IWSSIP)*. IEEE (2017), pp. 1–9
49. J. Kopf, 360 video stabilization. ACM Trans. Graph. (TOG) **35**(6), 195 (2016), ACM
50. C. Tang, O. Wang, F. Liu, P. Tan, Joint stabilization and direction of 360° videos. ACM Trans. Graph. (TOG) **38**(2), 18 (2019), ACM
51. H. Huang, Y. Xu, J. Chen, S. Song, T. Zhao, Latitude-based visual attention in 360-degree video display, in *Pacific Rim Conference on Multimedia*. Springer (2018), pp. 282–290
52. P. Hanhart, Y. He, Y. Ye, J. Boyce, Z. Deng, L. Xu, 360-degree video quality evaluation, in *2018 Picture Coding Symposium (PCS)*. IEEE (2018), pp. 328–332
53. V. Zakharchenko, K.P. Choi, J.H. Park, Quality metric for spherical panoramic video, optics and photonics for information processing X, 9970, 99700C. International Society for Optics and Photonics (2016)
54. H.T.T. Tran, C.T. Pham, N.P. Ngoc, C.M. Bui, M.H. Pham, T.C. Thang, An evaluation of quality metrics for 360 videos, in 2017 *Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE (2017), pp. 7–11
55. H.T.T. Tran, C.T. Pham, N.P. Ngoc, A.T. Pham, T.C. Thang, A study on quality metrics for 360 video communications, in *IEICE TRANSACTIONS on Information and Systems*, vol. 101, 1. The Institute of Electronics, Information and Communication Engineers (2018), pp. 28–36