Special Issue Published in International Journal of Trend in Research and Development (IJTRD), ISSN: 2394-9333, www.ijtrd.com

Performance Pattern Metrics Evaluation in Three Dimension Visual Data

[1]Donthukurthi Mourya karthik and [2]R Bala Dinakar,
[1]PG Student, [2]Assistant Professor,
[1,2]Department of Computer Applications, Godavari Institute of Engineering and Technology, Rajahmundry, India

Abstract—Quality assessment I is a key role for the development for multimedia applications there are many challenges in 3d quality assessment depth perception and virtual view and 2d counterparts these are most tedious parts in the quality metrics due to low correlation scores in this paper we implement binocular integration behaviors—the binocular combination and the binocular frequency integration, in the proposed method we find that the performance metrics and synthesized color-plus depth are evaluated and we find that this approach is more efficient than the other approaches with our comparison results

Keywords-- *Stress detection, factor graph model, social media, healthcare.*

I. INTRODUCTION

The process of creating S3D videos usually involves one of three approaches: capturing footage using a stereoscopic camera system, conversion from 2D videos, and computer graphics. In this work we address automatic quality control of stereoscopic videos converted from a 2D source. The International Telecommunication Union defines the quality of experience that a stereoscopic system provides as comprising three perceptual factors: picture quality, depth quality, and visual discomfort.In this paper we mostly examine the last two, as stereoscopic conversion rarely affects picture quality. The quality of the resulting S3D video, however, depends on both the depth maps and the warping algorithms. Processing of object boundaries is a principal concern; it may require filling of holes and handling of semitransparent edges. Low-quality depth maps may cause numerous problems ranging from annoying jitter on object boundaries to a complete inability of viewers to perceive the scene. A number of artifacts may arise during stereo conversion; these artifacts not only cause visual discomfort to viewers watching these videos, but they also impede popularization of S3D. It should be noted that automatic quality measurement of 2D-to-3D conversion is a significantly more difficult task thanquality evaluation of natively captured S3D. Numerous issues may arise during the conversion process: structural deformations, inconsistent occlusion filling, low-quality edge processing, and so on. Thus, owing to the problem’s complexity, we restrict our present coverage to the two most common and serious issues arising for 2D-to-3D conversion: edge-sharpness mismatch and the cardboard effect. To evaluate the quality of objectboundaries, we define the term edge-sharpness mismatch to describe defective stereo pairs with particular asymmetric impairments. This termrefers to anyinconsistencies in the appearance of object edges between the stereoscopic views (edge-sharpness variation, edge doubling, ghosting, and so on). Under the viewing conditions of a real environment, such situations rarely occur. In the case of 2D-to-3D conversion, however, the likelihood of edge-sharpness-mismatch can be rather highII.

II. LITERATURE REVIEW

Emerging research works of 3D IQA model for 3D stereoscopic images/videos.Based on the utilized information, we can classify the 3D IQA models into two categories: Color Information only computed quality scores on the SIFT matched feature points. A multiple channel model is adopted to estimate the stereo image quality [34]. The perceptual distortion of a stereo video is computed in discrete wavelet transform (DWT) domain quality assessment of stereo images based on the extracted edge information. The Gabor response of binocular vision was modeled the quality of color images and the disparity maps are integrated to evaluate the overall 3D perceptual quality.

In this presented a number of metrics for estimating the quality of captured S3D videos; these metrics enable detection of color mismatch, sharpness mismatch, and vertical disparity. Moreover, we described a method for analyzing horizontal disparity in S3D videos, which is relevant to converted S3D videos as well. However, color mismatch and vertical disparity rarely occur in converted S3D owing to the nature of a common conversion pipeline, and sharpness mismatch in converted S3D has nothing to do with the camera system, but happens because of the warping and interpolation algorithms. Even though all of these artifacts can arise in converted S3D videos, in [3] we avoided addressing artifacts specific to converted S3D. All the existing models for 3D-video quality estimation can be generally divided into two categories: • Metrics that aim to achievegood correlation with mean opinion score (MOS) for the whole video sequence, and may require some additional information (e.g., the original 2D video, if the input S3D sequence was converted from 2D). • Metrics that retrieve artifacts of concrete types from the video sequence (e.g., methods proposed in this paper). One model from the first category for objectively estimating 3D-video quality appears in. Although the authors took picture quality and visual discomfort into account, analyzing the quality difference between the left and right views, this model disregards depth-map quality, which plays a significant role in stereoscopic conversion. For example, the cardboard effect isunlikely to result into quality difference between the stereoscopic views, yet it does affect the quality of experience.

III. PROPOSED APPROACH

Compression plays a vital role for promoting the 3D image/video since its huge data volume makes 3D image/video requiring more storage space and higher transmission bandwidth than that of its 2D counterpart. That we can classify the types of stereo image compression proposed FI-metrics and the reported MOS, for a given symmetrically/asymmetrically compressed stereo image, as compared with the normal averaged quality metrics of the left-eye and the right-eye images. Since the success of objective 3D IQA can bring new thoughts for the existing 3D visual applications and the development of 3D video compression standards the significance and/or contribution of our work is quite clear. the left-eye image (QPL) and the right-eye image (QPR). That is, the types of stereo image compression can be classified as:

1. Symmetric-Stereo compression: QPL ~= QPR.
2. Asymmetric-Stereo compression: QPL = QPR

A. Advantages

1. The quality of the final synthesized stereo image is close to symmetric
2. The image encoded at lower quality in the asymmetric compressed stereo image
3. we focus on the binocular visual behaviors that describe the visual inputs integration process
4. Frequency integration behaviors. In order to overcome the challenges of 3D image

IV. METHODOLOGY

A. Edge-sharpness mismatches detection:

Consider a na¨ıve solution to the problem presented above. Such a solution would perform edge matching between stereoscopic views and would use a state-of-the-art technique for edge-sharpness estimation to assess edge sharpness mismatch. Owing to the high variability of possible edge inconsistencies between stereoscopic views, however, such na¨ıve approaches often fail. Another major problem is dealing with complex backgrounds and large occlusion areas. Under such conditions background change often occurs, defined as a significant deviation in background appearance along a given object edge between stereoscopic views. Thearbitrary nature of possible background changes makes estimating differences in edge appearance quite challenging in such cases. To overcome this problem we make the following assumption: in regions where backgroundchange occurs, edge-sharpness mismatch is absent. Obviously, this assumption doesn't always hold, but itgreatly simplifies the problem, essentially transforming it into the problem of edge detection with background change—circumstances in which we cannot reliably estimate edge-sharpness mismatch.

B. Edge-sharpness mismatch detection involves four steps

1. Edge detection and matching between stereoscopic views with outlier rejection
2. Estimation of a raw edge-sharpness mismatch map
3. Estimation of a background change penalty map
4. Refinement of the edge-sharpness mismatch map

C. Cardboard-Effect Estimation

The problem of estimating the cardboard effect can be regarded as a special case of detecting inconsistencies between the depth map used in 2D-to-3D conversion and the "real" depth map of the scene. This problem is insoluble for the general case owing to the absence of both a reference depth map and a reliable way to estimate it. Furthermore, depth maps can vary significantly because of creative choices made during the 2D-to3D conversion process. For example, some cardboarded characters can be introduced intentionally to draw the viewer's attention to other, more volumetric characters.

D. Flat Foreground Objects Detection

Detection of flat foreground objects involves the following:

1. Stereo matching
2. Disparity map segmentation and foreground-object detection
3. Flatness estimation for each detected foreground object To perform stereo matching we use an algorithm identical to that of the edge-sharpness mismatch metric, resulting in a dense disparity map. The mean-shift algorithm segments the disparity map and extracts as foreground object masks M_i any sufficiently large segments with negative disparity.
4. Flatness estimation of a foreground object considers three main aspects: the strength of the object's texture, the size of the object, and the ability to reconstruct the object in one view through a simple uniform shift of the other view.

E. Flat scenes detection

To detect flat scenes, we use an algorithm similar to the one mentioned above. The main difference between flat-object detection and flat-scene detection is the definition of the term flat. We consider a scene to be flat if its disparity map is a linear function of pixel position (x,y). Therefore, we perform a linear least-squares approximation of the disparity map to estimate how well a linear function represents the scene's disparity map.

V. ALGORITHM

A. Multimodal Shot Pruning (MSP)

Algorithm:

*t*wo V x S shot matrices, S and F, are initially constructed:

S is the total number of movie shots

V is the total number of visible speakers, i.e., different actors that

Typically S >> V

Most likely, the basis vector sets for S and F are the standard basis, with one basis vector corresponding to each visible speaker. S and F are modified, through a Gaussian expansion process, in order to extend each speech / face appearance to neighboring shots:

–Binary matrices are converted to real ones –For each Si j = 1 / Fi j = 1, a discrete approximation of a Gaussian distribution, with its mean at Si j / Fi j , is locally assigned to the entries of the i-th row around Si j / Fi j.

- Shot matrix values derived from different speech / face appearances and corresponding to the same shot matrix entry are added –Neighboring speech / face appearances are temporally diffused, to achieve rudimentary scene modeling
- –Most likely, the basis vector sets of the modified shot matrices include vectors corresponding to the most prominent actors and vectors corresponding to combinations of more and less prominent actors.
- –Intuition: lead actors can appear alone, while supporting actors appear mainly along with leads
- Approach: Cast the problem as a joint matrix Column Subset Selection Problem (CSSP) on S and F, where the desired solution is a vector c of matrix column indices (corresponding to retained movie shots).
- Solve it with a genetic algorithm [2], using the following joint-CSSP fitness function: f(c) = (||S – (CS CS+)S||F + ||F – (CF CF+)F||F)
  1 – CS / CF are sub-matrices of S / F, respectively, containing only the columns indicated by c.
- Key-segments contained within the same shot and separated by less than a second of video duration are merged. Too short key-segments are eliminated. Purpose: eliminate abrupt temporal jump cuts.

- Visually annoying depth jump cuts, i.e., severe mean disparity mismatches among consecutive video frames induces by the skim construction process, are detected and fixed by applying the method in.

B. Frame Moments Descriptor (FMoD)

Alogritham

- The frame (M x N) is iteratively partitioned in small blocks (m x n), under a spatial pyramid scheme.
- For each block, profile histograms are computed for the horizontal dimension and the vertical dimension, by averaging pixel values across block columns/rows, respectively.
- The result is an n-dimensional and an m-dimensional vector. Each one is summarized by its first 4 statistical moments.
- The process is repeated multiple times, for larger values of m and n, resulting in different local frame descriptions in different scales.
- The inclusion of disparity-derived information leads to more representative key-frames.
- Key-frame selection per shot via intra-shot K-Means++ video frame clustering and extraction of cluster medoids.
- Adaptive number of clusters (2 – 5): regulates the number of key-frames per shot, based on an internal clustering evaluation metric (Silhouette Coefficient) within each shot.
- The computed key-frames, derived from all movie shots, are subsequently partitioned in an inter-shot KMeans++ clustering
- stop with fixed number of clusters (percentage of movie duration).
- All remaining key-frames are temporally extended to key-segments, using p neighboring frames.
- Each key-segment is extended so as to completely include any temporally overlapping speech segment appearances. These are pre-computed using speaker diarization and speaker clustering algorithms. Thus, in the final skim, no speech instance will be abruptly interrupted.
- Any temporally overlapping key-segments are concatenated.

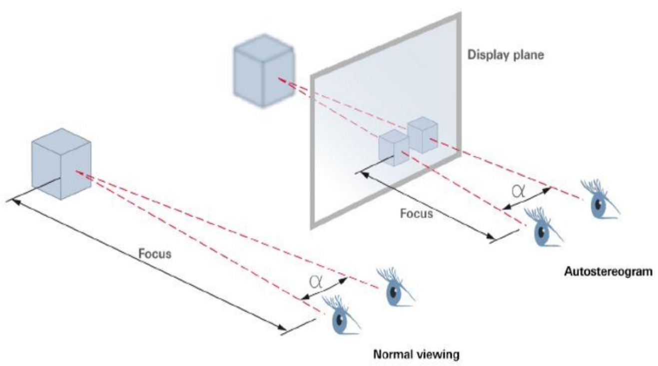
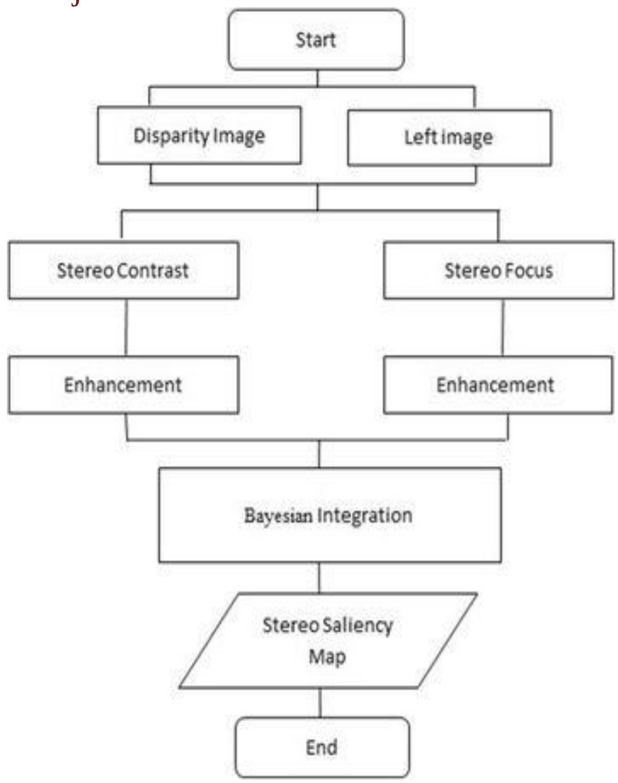


Figure 1: Evaluation in 3D visual data



CONCLUSION

In this paper we presented our method for evaluation of 2D-to-3D conversion quality. We described algorithms for detecting edge-sharpness mismatch artifacts and the cardboard effect. To demonstrate the utility of our proposed method, we evaluated several recent Blu-ray 3D releases and presented some of the detected artifacts. Possible directions of further work include developing new metrics, decreasing the number of false positives in current metrics, and subjective testing to determine the noticeability of artifacts to viewers.Another major topic is subjective testing and development of quantitative metrics that correlate well with human perception of 2D-to-3D conversion artifacts. The results of human-perception tests may depend on the type and quality of the display device; also, the results are susceptible to viewers' visual acuity and their ability to perceive 3D. This topic should therefore receive separate attention.

References

[1] ITU-R Recommendation BT.2021, "Subjective methods for the assessment of stereoscopic 3DTV systems," tech. rep. (Aug 2012).

[2] Boev, A., Hollosi, D., Gotchev, A., and Egiazarian, K., "Classification and simulation of stereoscopic artifacts in mobile 3dtv content," *Proc. SPIE* 7237, 72371F–72371F–12 (2009).

[3] Voronov, A., Vatolin, D., Sumin, D., Napadovsky, V., and Borisov, A., "Methodology for stereoscopic motion-picture quality assessment," Proc. SPIE 8648, 864810–864810–14 (2013).

[4] Yamagishi, K., Kawano, T., and Kawashima, K., "QoE Assessment Methodologies for 3D Video Services," NTT Technical Review (May 2013).

[5] Shao, H., Cao, X., and Er, G., "Objective quality assessment of depth image based rendering in 3DTV system," in [3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2009], 1–4 (2009).

[6] Fehn, C., "A 3D-TV approach using depth-image-based rendering (DIBR)," in [Proc. of VIIP], 3 (2003). [7] Solh,

M. and AlRegib, G., “A no-reference quality measure for DIBR-based 3D videos,” in [Multimedia and Expo (ICME), 2011 IEEE International Conference on], 1–6, IEEE (2011).

[7] Bouma, H., Dijk, J., and van Eekeren, A. W. M., “Precise local blur estimation based on the first-order derivative,” Proc. SPIE 8399, 839904–839904–8 (2012).

[8] Canny, J., “A computational approach to edge detection,” Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI-8(6), 679–698 (1986).

[9] Simonyan, K., Grishin, S., Vatolin, D., and Popov, D., “Fast video super-resolution via classification,”in [Proceedings of IEEE International Conference on Image Processing], 349–352 (2008)