RESEARCH

Open Access

Stereoscopic 3D video quality assessment based on depth maps and video motion

Juan Pedro López[*], Juan Antonio Rodrigo, David Jiménez and José Manuel Menéndez

Abstract

In this paper, we propose techniques to assess the objective quality for stereoscopic 3D video content, related to motion and depth map features. An analysis has been carried out in order to understand what causes the generation of visual discomfort in the viewer's eye when visualizing a 3D video. Motion is an important feature affecting 3D experience but is also often the cause of visual discomfort. Guidelines are obtained after applying the algorithm to quantify the impact over viewer's experience when common cases happen, such as high motion sequences, scene changes with abrupt parallax changes, or complete absence of stereoscopy.

1 Introduction

The success of new 3D services is a reality due to the increase of quality of experience. Although there are some factors and initial measurement devices in this field, there are still no common way and procedure to compare 3D video contents and integrated solutions and obtain an evaluation of quality.

Considering the Qualinet White Paper, quality of experience (QoE) is defined as the degree of delight or annoyance of the user of an application or service [1]. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state. Quality is influenced by different factors such as content, network, device, application, user expectations and context, and more when 3D video is considered.

While evaluating visual quality assessment over standard 2D video, there are several factors that are typically taken into account, such as sharpness, blocking effect, or blurring. The difference in perception between regular 2D video and 3D stereoscopic video extends the list of factors that will affect its visualization, thus, emerging the need to study them in order to offer a more complex 3D quality assessment.

Stereoscopic 3D video perception is based on the fact that two different video signals are captured in order to feed each of the viewer's eyes. There is a signal aimed to be received by the left eye and another one aimed to be received by the right eye. This system tries to recreate the experience of watching a real world scene, where two different images are captured by each eye and the difference between them depends on the position of the elements in the world related to the viewer's position. This means that the system is feeding the observer with a disparity depth cue.

But the experience of watching 3D TV is significantly different from a natural view, as the point of view is prefixed by the fixed point of view of the camera lenses that have captured the scene and so is the focus. Furthermore, in natural viewing, the eyes focus (accommodate) and converge to the same distance, but when looking at a 3D object displayed on a screen, a viewer's eyes must focus on the screen while, at the same time, they converge on a point in space that may be located beyond the screen, on the screen, or in front of the screen. This is known as the *vergence-accommodation conflict*. This conflict limits the amount of parallax that a viewer can tolerate avoiding discomfort, also known as the *zone of comfort* (it will be studied in more detail in the next section).

This paper aims to study the effects of stereoscopic disparity in quality assessment through the analysis of depth maps of a sequence and its temporal evolution. We try to quantify objectively the effects of parallax, depth and motion, exporting the common situations in which discomfort is substantial, from opinions of observers derived from empirical and subjective tests.

* Correspondence: jlv@gatv.ssr.upm.es
Escuela Técnica Superior de Ingenieros de Telecomunicación de la
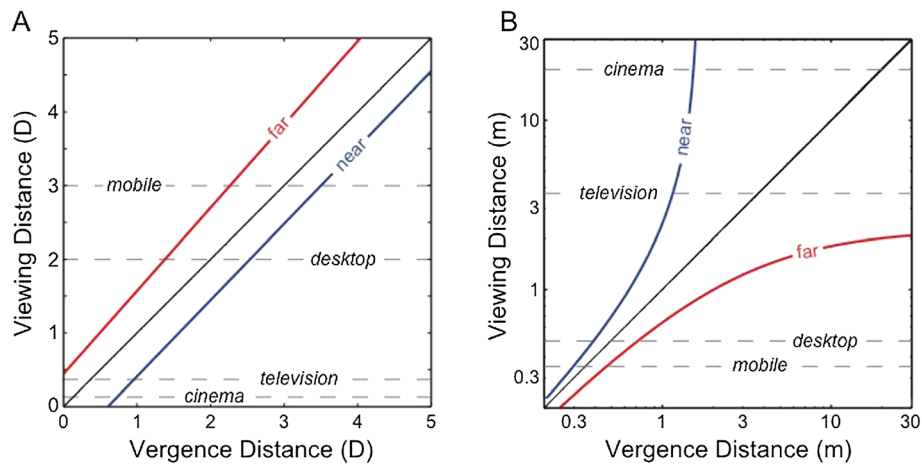Universidad Politécnica de Madrid, Madrid 28040, Spain

Figure 1 Shibata's zone of comfort with distance in (A) diopters and (B) meters.

In the following text, we will first make an overview of the related studies in Section 2 and the description of subjective assessment developed which is compiled in Section 3, define the aims and questions to be answered in Section 4.The test results are given in Section 5 and Section 6 concludes the paper.

2 Related work

Image quality assessment is a difficult process which plays a major role in various processing applications [2]. A lot of work has been developed in this field, defining metrics and algorithms to predict the quality of a video sequence, avoiding new subjective studies, which are necessary to demonstrate the validity of these objective metrics. An overview of the extensive and most interesting work in quality assessment is collected in [3-5]. But objective video quality is always related to subjective quality assessment, because the final conclusion to obtain is the impact on the observer. Most of these subjective studies published adhere to the procedures contained in the Recommendation ITU-R BT.500 [6] and in ITU-T P.910 [7], in order to define settings to

facilitate the comparison between studies developed by different researchers.

In 3D media, not only new factors are concerned in order to assess quality and artifacts that commonly appear on 2D video are inherited, but also new topics related to the optic effect of stereoscopy, such as visual discomfort or perceptual inconsistencies between depth cues, as stated in [8]. Much work has also been developed in this field relating to depth and motion, such as in [9], where filtering is used to reduce visual discomfort on screens, and also studies in [10,11]. Additionally, interesting quality model articles about QoE in 3DTV are proposed in [12].

In [13], an overview describing the main topics relevant to comfort in viewing stereoscopic television is developed and analyzed after subjective tests related to accommodation-vergence conflict, parallax distribution, binocular mismatches, and depth, and cognitive inconsistencies. In [14], it is reported that depth and motion are closely related in terms of calculating visual discomfort. In addition, authors in [15] offer a visual comfort model for detecting salient object's motion features in

Figure 2 Examples of left frames in sequence Modernism.

Figure 3 Effect with anaglyph examples of parallax variation in sequence Palco HD.

depth of field; also models for viewing comfort prediction were developed by Richard et al. [16]. Moreover, studies were developed by Barkowsky et al. about visual discomfort in stereoscopic 3D video sequences in [17,18].

An interesting subjective evaluation of visual discomfort is developed in [19], where parallax limits and regions of comfort, depending on the screen size, disparity, and viewing time, are obtained.

Other artifacts such as stereo window violation (SWV) and temporal continuity of the disparity (TCD) have been studied in [20]. Guidelines to create comfortable and faster stereoscopic films are included in this work, based on depth maps and 3D cinematography principles, and also in works developed by Chen et al. [21].

The zone of comfort (ZoC) was first introduced by Percival [22], when, based on experiments with spectacles, he suggested the limits to vergence-accommodation postures that could be achieved without causing discomfort. Recent studies, such as Shibata et al. [23], conclude that the ZoC may differ from Percival's ZoC when the experiments are based on stereoscopic vision rather than based on vision through spectacles. In stereo vision, the vergence-accommodation conflict constantly changes; while in a lens or spectacles system, it is maintained fixed [24]. Figure 1 shows Shibata's ZoC for different accommodation distances in stereo vision, both in diopters (optic measurement and meters (length)). Vergence-accommodation distance in meters will help understand a real deterministic situation where the viewer is positioned at a precise distance from the screen where the

stereo image is produced. According to this diagram, images with positive parallax have little or no capability to induce discomfort, while negative parallax is most likely to cause discomfort if not controlled.

In order to adapt a stereoscopic ZoC to 3D video, it is necessary to take into account motion and time of exposure to a stereo scene. The ZoC will be further reduced when these elements appear. The time to convergence and accommodation in this case is relevant; thus, there is a need to adapt the concept of ZoC. In [17] the variation of time of exposure is studied in order to determine its effects of visual discomfort, thus, relating it to the ZoC.

The determination of an image parallax range is obtained through an associated depth map. There are several ways to obtain a depth map from a stereo image, depending on the computation complexity and accuracy restrictions. As a rule of thumb, it can be stated that complexity is proportional to accuracy, thus, low complexity algorithms such as sum of absolute differences (SAD) can perform well under certain circumstances, as stated in [25,26]. Scharstein and Szeliski [27] described the procedures to define the Middlebury Database [28], which offers the most complete stereoscopic algorithms benchmark to date, and software to evaluate new algorithms to predict image parallax. SAD-based algorithms are among the least complex and more often used. Census-based algorithms [29-32], first introduced in [33], are common in real-time hardware-based systems [34-38] and may work better in the homogeneous zones
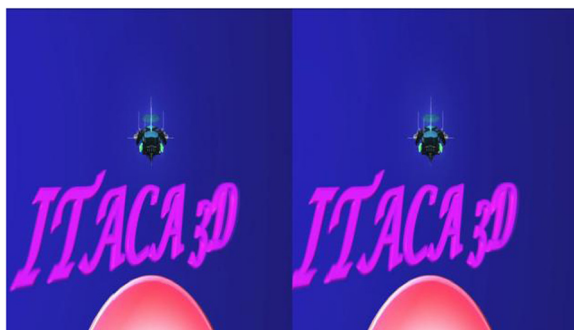

Figure 4 Sequence 'Itaca 3D' with window violation (WV) on the left side of image.

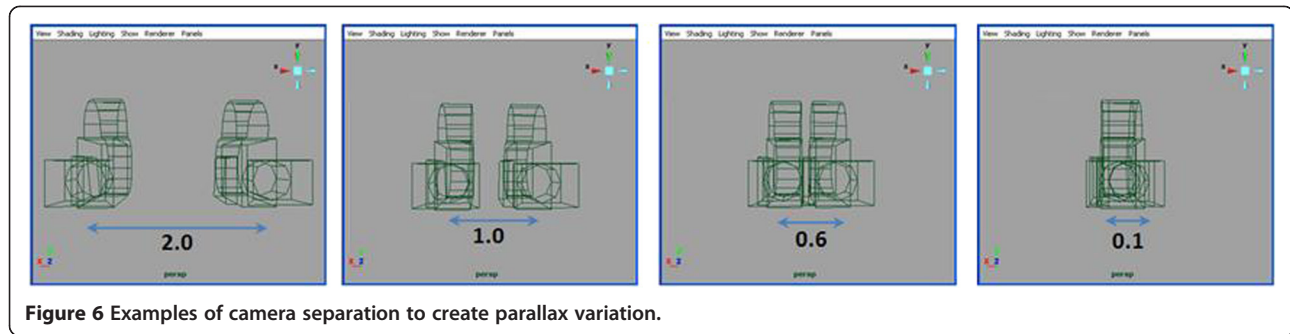**Figure 5 Sequence Itaca 3D with WV on the right side of the image.**

Figure 6 Examples of camera separation to create parallax variation.

of the image. Its complexity increases when used in software-based systems because of its bit-based nature. Some systems mix both algorithms in order to obtain the best results from each one of them, such as [39], ranked in second place in the Scharstein and Szeliski list, or [40] based on dense stereo matching.

Finally, works by Nojiri et al. [41] are centered in visual comfort and sense of presence in stereoscopic video, relating the distribution of parallax and motion in a sequence. Results are based on subjective assessment, using a single-stimulus continuous quality evaluation method.

3 Determination of visual discomfort sources

Tests have been run over a set of 3D video sequences to understand and analyze different features which generate visual discomfort or quality reduction. A group of ten observers were asked to rank the sequences, taking into account their 3D quality. Results were compared to the objective data obtained through our developed tools to decide which features would be a possible cause of visual discomfort and how to modify them to obtain good 3D experiences. All tests were carried out on a 46-in. 3D LCD monitor model JVC GD-463D10 (JVC, Wayne, NJ, USA) [42] with passive glasses.

A group of 16 subjects took part in the tests, taking into account that Winkler considers enough 15 subjects for subjective assessment in [43], evaluating their stereoscopic quality in Mean opinion score (MOS). The test consisted in observing each sequence and giving them a MOS from 5 to 1, considering the meaning of each note: 5 (perfect), 4 (fair), 3 (annoying), 2 (very annoying), and 1 (impossible to communicate). The experiment followed the ACR (absolute category rating) explained in recommendation P.910 [7].

Sequences used for the assessment included a 3-min sequence called 'Modernism' (Figure 2), created by Mediapro company, in which different scenes appeared with different levels of motion and depth. Sequences 'Rain Fruits' and 'Fountain' from EBU were also used [44].

Synthetic test sequences 'Palco HD' (Figure 3, which shows its anaglyph version) and 'Itaca' (Figures 4 and 5, which show the spinning logo and moving ship with window violations in left and right side) were also used, which included parallax and object distance variation. These sequences were created specifically for these tests using the software tool *Autodesk Maya* to develop an environment with virtual pairs of cameras, to enable an easy method to alter distance between the pair, generating high negative parallax images (Figure 6).
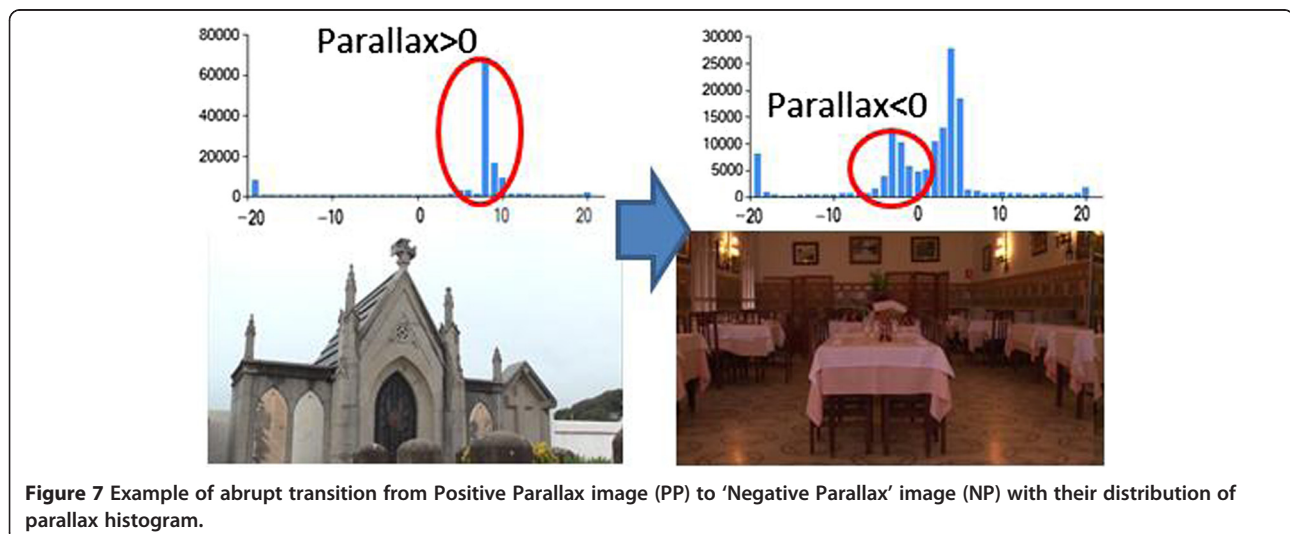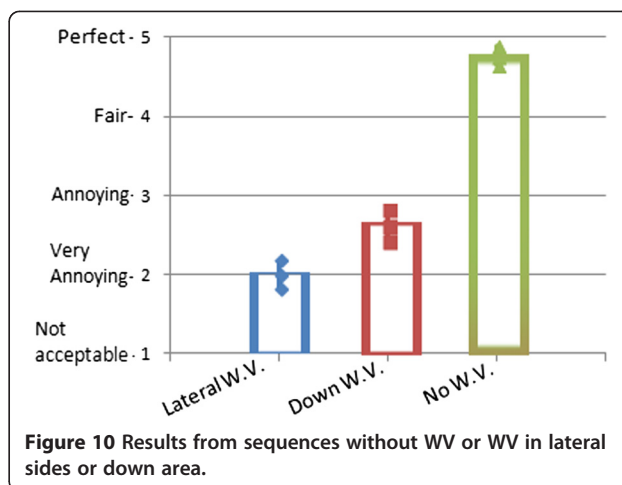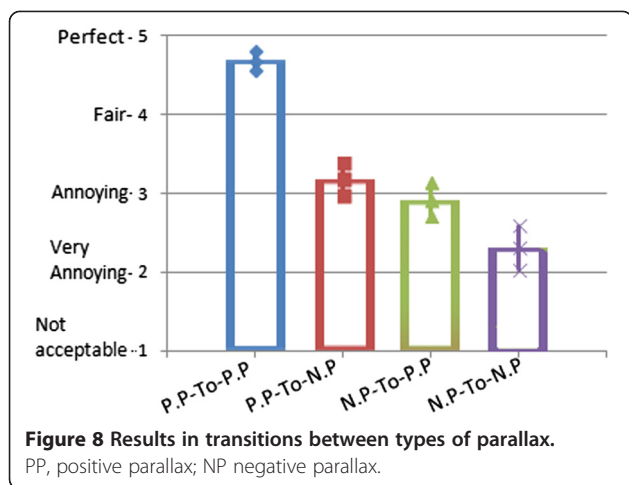

Figure 7 Example of abrupt transition from Positive Parallax image (PP) to 'Negative Parallax' image (NP) with their distribution of parallax histogram.

Figure 8 Results in transitions between types of parallax.
PP, positive parallax; NP negative parallax.

Figure 10 Results from sequences without WV or WV in lateral sides or down area.
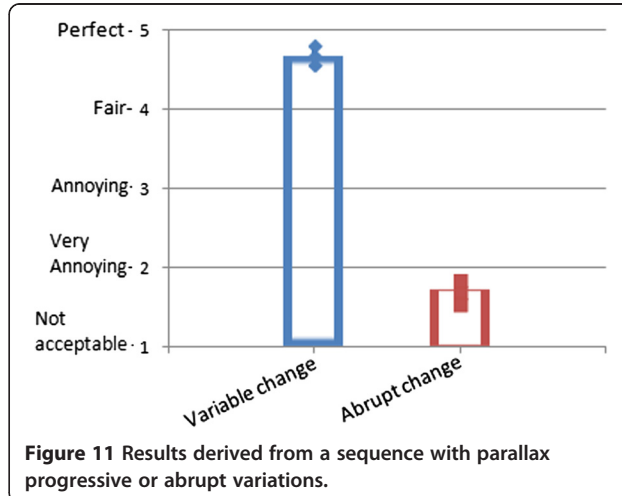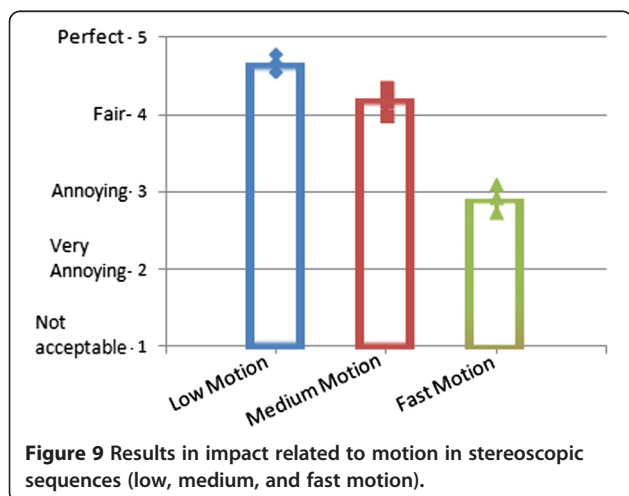
All these sequences have high-definition resolution (1,920 × 1,080), with no compression, and available in side-by-side formats as half-resolution. Sequences generated by EBU are also available in side-by-side with complete HD resolution for both left and right views, which is interesting to test the necessity of higher resolution.

The sequences used were distributed among six different groups, each one related to one experiment to analyze. The user gave a score to every sequence (or pair of sequences) and had free space to write a general comment on each one, in order to draw conclusions about their opinions:

- Pair of sequences with transitions from different types of parallax (example in Figure 7), negative and positive, to detect the impact over abrupt stereoscopic changes. A 'positive parallax' sequence (PP) is considered when it has no remarkable negative parallax and pixels, with the positive parallax representing more than 25% of an image. On the

other side, negative parallax sequences (NP) are sequences whose images posses more than 15% of pixels in negative parallax (assuming an environment of positive parallax). See Figure 8 for results.

- Same static images and video sequences with different values of parallax, from zero and positive parallax, to negative out of ZoC values
- Negative parallax sequence with different levels of motion: low, medium, and high. Test statistics are collected in Figure 9.
- Sequences with WV produced in different sides of the image, in lateral or top/bottom regions of the image. Examples are shown in Figures 4 and 5. See subjective results in Figure 10.
- Negative parallax variations to compare zones inside and outside of Shibata's ZoC
- Long sequence with soft variation of parallax, at the end the sequence, starts from the beginning producing an abrupt parallax change. Results from this experiment appear in Figure 11.
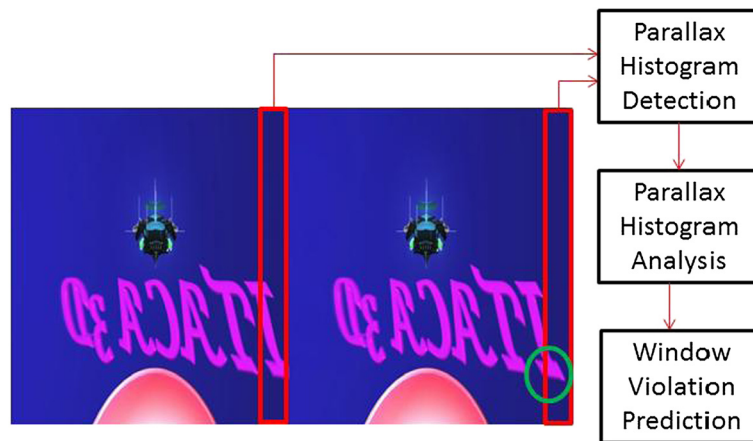

Figure 9 Results in impact related to motion in stereoscopic sequences (low, medium, and fast motion).

Figure 11 Results derived from a sequence with parallax progressive or abrupt variations.

Figure 12 Working scheme for detecting window violation with example of sequence Itaca 3D.

The session had duration of approximately 20 min and the users did the tests individually. Prior to the tests, some sequences were used to prove the observers' normal stereopsis, as recommended in [45].

Compared to Nojiri et al. [41] tests, this subjective assessment is more exhaustive and centered in analyzing concrete cases of visual discomfort because these tests isolate each one of the cases, such as detecting window violations or abrupt changes of parallax, based on parallax distribution and motion to generate a global algorithm to detect cases of visual discomfort.

3.1 Conclusions related to the subjective assessment

Analyzing the graphs presented and the observers' opinions obtained from the study, it is possible to conclude the following:

1. Derived from variation of parallax in static images:

    (a) Negative parallax with objects situated in the edges of the picture frame generates disparities in visualization of objects closer to the observer, by the effect of a window violation.

    (b) Sequences with objects out of the ZoC generate visual discomfort and fatigue in the observer's eye.

    (c) The variance in the parallax histogram, especially in negative parallax areas but also in positive areas, has a high influence in 3D experience and, as a consequence, in quality in general.

2. Derived from variation of parallax in motion sequences:

    (a) When a scene change occurs, the abrupt variation from high-percentage positive parallax pixels to sudden emergence of negative parallax objects is very annoying for human eye, if it has not enough time for accommodation and focus on the image. As seen on abrupt changes graph Figure 11.

    (b) In fast motion video sequences, even positive or negative parallax could be a cause of visual discomfort (see Figure 8), but it has been
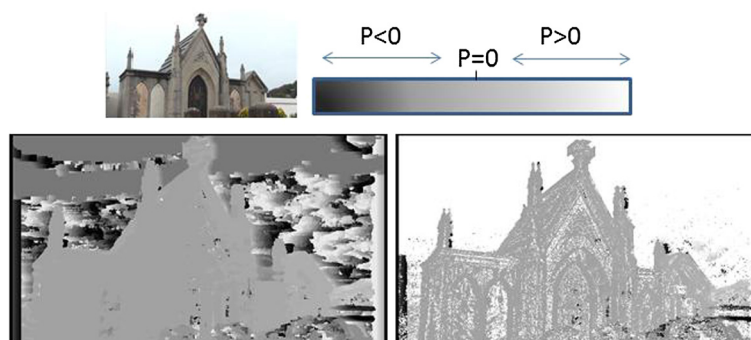

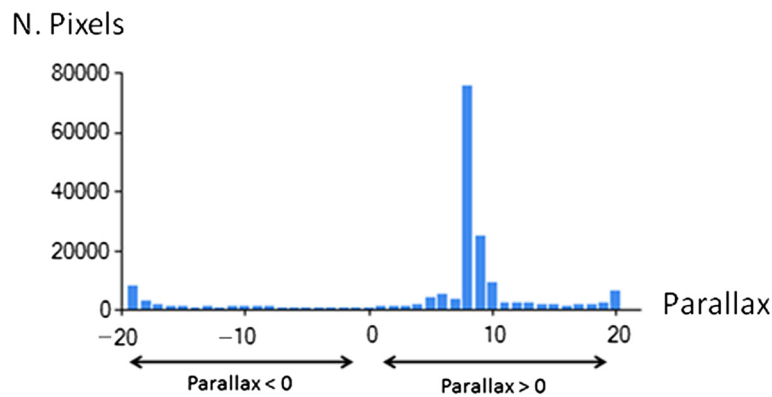**Figure 13 Depth map.** *P* refers to parallax.

Figure 14 Example of parallax histogram corresponding to positive parallax image.

concluded after several experiments that negative parallax is especially annoying for human eye.

(c) In cases of low motion sequences, the time of convergence and accommodation is higher, but the high variations are still considered.

(d) Images with motion in objects with different values of negative parallax are difficult to focus simultaneously (Figure 8), as demonstrated empirically in sequence 'Itaca' sequence which appears in Figure 4, with a logo turning in circles, while a starship is getting close to observer. Observers' commentaries about this sequence revealed their difficulty to focus both objects, because of the difference in negative parallax.

4 Work implementation

In this section, the work developed is described in two subsections. First (Section 4.1), the tools implemented in order to obtain quality through depth map histograms, calculating degradations related to each individual frame, are described in detail. Based on cases of study

summarized in last section, techniques are developed to quantify these effects, annoying for human eye.

In the second part (Section 4.2), the work for static images is extended to sequences, analyzing video motion, and the effect of depth when there are variations in parallax, derived from depth maps. Also, some cases of the study are analyzed when combining depth and motion.

4.1 Quality in static images using depth map histograms

To resolve the validity of a stereoscopic image, it is required to determine whether it delivers visual discomfort or annoyance. The developed algorithm obtains parallax information through the computation of a depth map.

First of all, the depth map histogram is compared to the suited ZoC in order to check if it does not fall out of its boundaries. Vergence-accommodation conflict needs to be confined between ZoC limits to prevent visual discomfort.

In order to evaluate the disparity results, it is necessary to understand the relation between disparity in pixels and virtual perception of depth. Vergence distance is determined by several factors such as image resolution
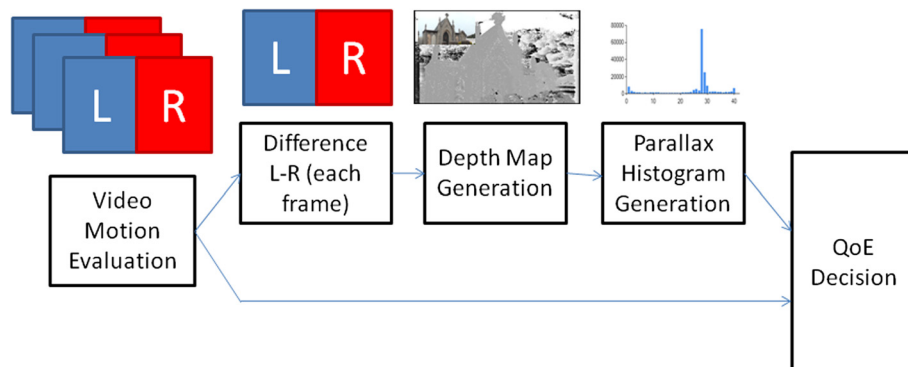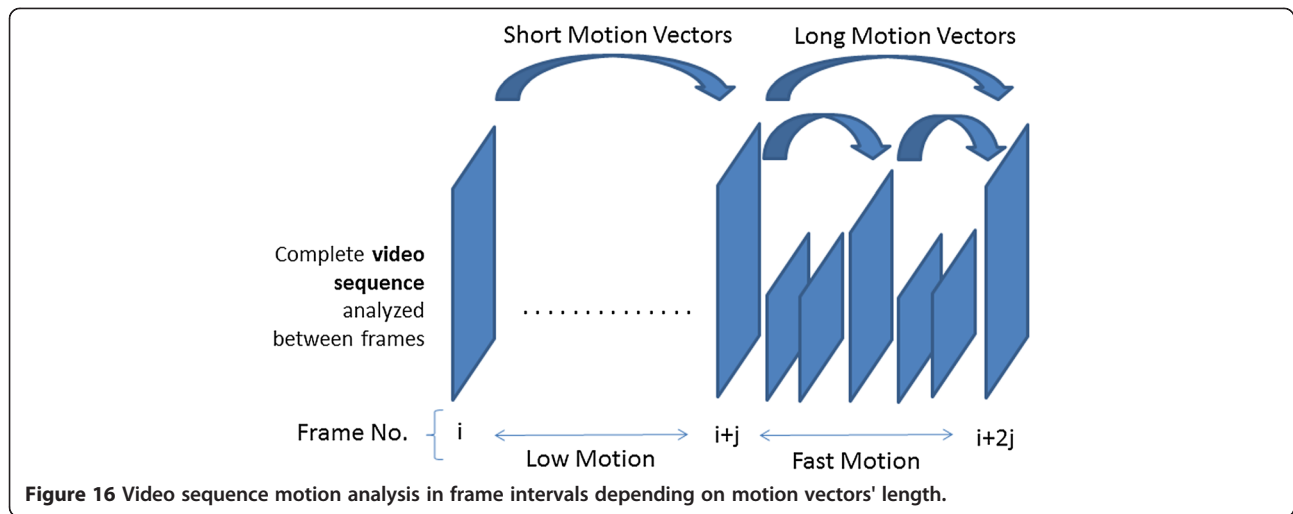

Figure 15 Scheme of QoE decision algorithm.

Figure 16 Video sequence motion analysis in frame intervals depending on motion vectors' length.

and the screen size distance between the viewer and the screen. Viewers stood at a distance from screen of approximately 2.4 m. Following ITU P.910 recommendations [7], it is the comfortable distance according to high-quality sequences. With those parameters, Shibata's ZoC has a parallax range limited to [−125, 107], measured in pixels. Anything relevance that the algorithm finds out of those bounds should be directly considered as a cause of visual discomfort. Note that the algorithm should be adjusted if any of the former parameters should change.

The other feature measured, which is the usual suspect of causing visual annoyance, is the *window violation*. Window violation occurs when an object with negative parallax does not fit the screen and, therefore, is cut by the screen edges. Having negative parallax, it is supposed to be out of the screen, which means that the view of screen edges should not be hidden. This generates an incoherent depth cue situation.

In order to measure this feature, the algorithm will examine the depth map's limits looking for negative parallaxes, which will be computed as a factor of visual annoyance. Due to the difficulties of obtaining disparities near the edges, a group 10 pixel rows or columns closest to the edges are analyzed to decide whether or not a window violation is detected. Figure 12 represents how the algorithm is computed.

4.2 Depth map calculation

To compute depth maps from stereoscopic images, the system performs a SAD-based algorithm. As it has been stated earlier, other algorithms are more powerful, but require much higher computational demand. We need to obtain the general depth characteristics of a scene and its evolution, even though pixel depth accuracy of the whole image is not necessary. SAD-based algorithms work well enough to fulfill our goal and are less computational demanding.

The weakest detections with SAD algorithms occur in homogeneous zones, where the capability to discern between possible pair candidates is low. In order to alleviate these probable errors, the system performs a difference between both views in order to calculate depths only over those pixels that will differ from one image to another, reducing homogeneous zones and, therefore, noise in resulting depth maps. Discarded pixels would
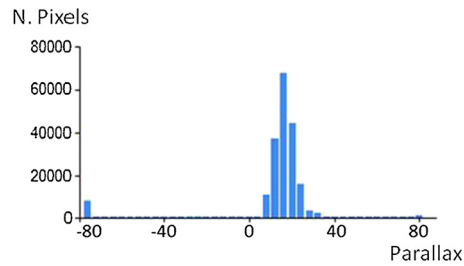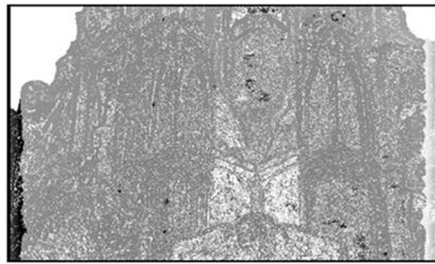

Figure 17 Example of motion vector detection indicating modulus and angle.
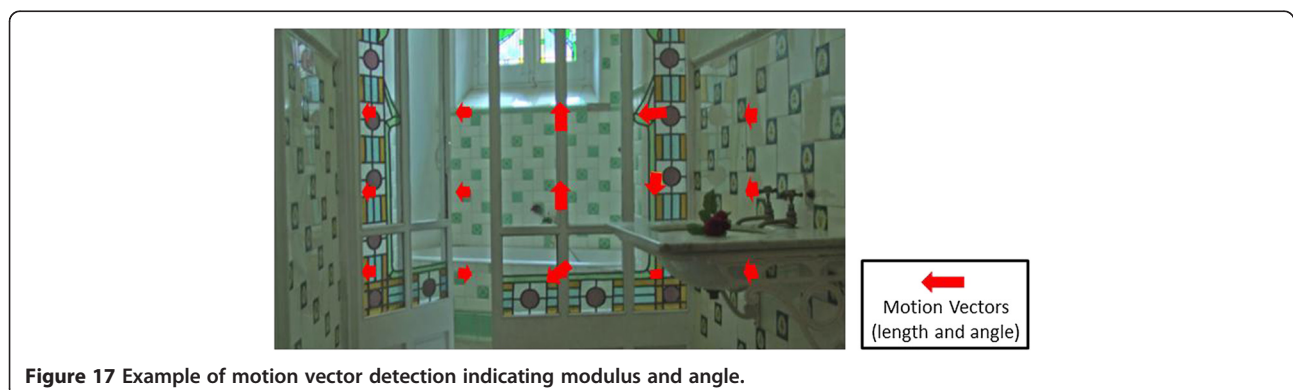
Figure 18 Church depth map and histogram.

not be taken into account for statistical calculations. The histogram will contain less depth information, but will be more accurate. Figure 13 shows the original depth map (left) and the filtered depth map (right). In the original depth map, there are several errors in the background zone, where the sky is homogeneous. In the left image, this zone is not calculated and therefore not taken into account to classify image's general depth.

Figure 14 shows the histogram calculated for the previous depth map. All the elements in the scene have positive parallax. There is a very small amount of negative parallax pixels which represent noise (bad information) that result from the depth map algorithm calculation.

4. 3 Depth-and-motion QoE decision algorithm

In 3D stereoscopic video, motion is a basic element to take into account, when assessing the quality, as a primal reason for visual discomfort, related to high depth levels which combine areas with negative and positive parallax.

The steps to for offering a QoE decision are the following: On the first instance, the complete sequence is processed to obtain the motion vectors in order to find the scene changes. Once the instance where the scene change is detected, each scene is isolated and motion vectors are calculated between consecutive frames to obtain the level of motion in that specific scene.

Depending on the level of motion, the scene is classified as slow, medium, or fast. This is necessary to decide the necessity of calculating new depth maps for various

frames if it is in fast motion or assuming the same disparity for a collection of similar consecutive frames in slow motion, saving computing time.

After deciding the key frames (one or more), depth maps are obtained for each of these frames by using the difference image between left and right view, which is used as a mask to simplify the process. Depth map is calculated as explained in the previous section about static images. The comparison between parallax histogram, derived from each key frame, allows us to make a statistic about the variations in objects depth and, consequently, quantify the probability of visual discomfort appearance when observing the sequence.

The final QoE decision algorithm offers guidelines and prediction about visual discomfort in human eye, analyzing individually the sources of discomfort, calculating the probability for each one to happen (Figure 15). Sources of discomfort considered are the following: the probability of detecting WV, the probability of finding abrupt transition between frames, and the probability of finding excessive negative parallax on frames, affecting to visual discomfort. QoE can be calculated as follows:

$$
\begin{aligned}
QoE = {} & \alpha_1 p\ (\mathrm{WV}) + \alpha_2 p\ (\mathrm{Abrupt\ transition}) \\
& + \alpha_3 p(\mathrm{High\ NP}).
\end{aligned}
\tag{1}
$$

In the formula (1) 'p' means the probability of an event to happen and α_i are constants dependent on motion and distribution of parallax.
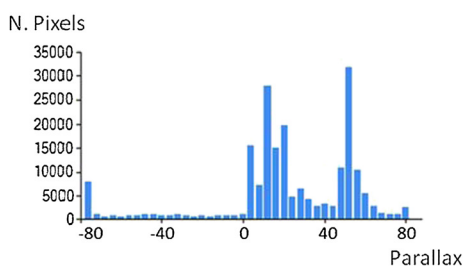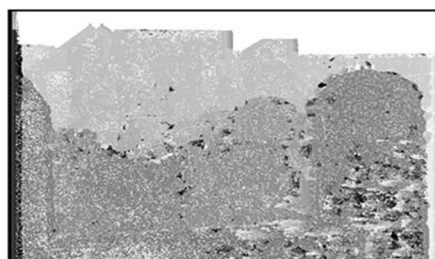

Figure 19 Cemetery depth map and histogram.

Figure 20 Library depth map and histogram.

4.3.1 Motion vectors and motion estimation

The work derived from the static image process is related to motion. It is necessary to evaluate the motion level in a video sequence to conclude how much this motion affects to the perception of the third dimension in stereoscopic video. For this purpose, motion vector calculation is obtained.

The sequence is analyzed in gaps to predict video motion, detecting the key frames. In consecutive static frames or areas with low motion, the depth map is assumed to be the same for that sequence of frames. On the other side, when medium or fast motion happens, more depth map information is necessary to compare the results.

As seen in Figure 16, the variance and average length of motion vectors is calculated between the frames separated in a fixed distance (in the image indicated as 'j'). If the length variation is low among motion vectors, the image is considered to be static or only low motion is generated. As a consequence, the treatment in this area of the sequence is the same as explained in Section 4.1. Motion is calculated as the average valid motion vectors (without the discarded incoherent ones), always related to the variance of motion lengths:

$$\text{Motion}_{\text{frame}i} = \left| \text{Average}\left(|MV|_{\text{frame}i} \right) \right| \qquad (2)$$

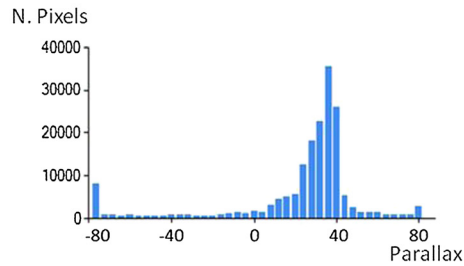For motion vectors calculation, only the left stereoscopic image is selected, and a grid is created to detect the blocks motion. In the case of the example of Figure 17, a grid with three lines and five columns allows us to obtain 15 different motion vectors. The blocks between 9×9 and 15×15 pixels are searched in the next frame left image; homogeneous blocks are discarded to avoid false detections. The motion must be coherent in distance, so vectors with length values over two times the variance are also discarded. The purpose of formula (2) is classifying the sequence intervals into four difference categories depending on the motion: static and low and medium or fast.

The last case is when a scene change occurs, the depth maps from both previous and next frame are processed. This is a concrete case of motion vector abrupt variation, in which the variance of the vectors' length is higher than when fast motion happens.

As manifested from observers, the abrupt changes of negative parallax in a positive/negative parallax environment provoke a high visual discomfort in the observer's eye:

$$\text{Discomfort} \equiv \left| \text{Average}\left(|P < 0|_{\text{frame}i} \right) - \text{Average}\left(|P < 0|_{\text{frame}(i+n)} \right) \right| > 10\%$$
$$(3)$$

The formula (3) defines the discomfort as an abrupt change in negative parallax, where $P < 0$ corresponds to the negative parallax in actual frame 'i' and 'n' is the difference between two separate frames.

Discomfort is only produced in an environment of significant variance of positive parallax and motion, even
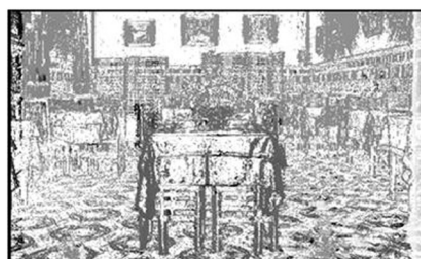

Figure 21 Table depth map and histogram.

Table 1 Histogram variances

Image	Positive variance (pixels)	Negative variance (pixels)
Church	10	-
Cemetery	19.6	-
Library	12.8	23.2
Table	15.6	16.4

with low and medium motion and especially in fast motion sequences.

5 Test results

With results obtained from the subjective assessment, objective metrics and studies are developed in static stereoscopic images and in motion video sequences.

5.1 Quality in static images using depth map histograms

Tests have been divided in two groups. First of all, the tests were run over still images to classify stereo features without dealing with motion effects. These tests were focused on ZoC measurements, window violations, and depth distribution.

To evaluate the effects of parallax out of ZoC, we have rendered virtual images such as the one shown in Figure 3. When disparity was forced to be near Shibata's ZoC, the perception of the observers was negative, even when disparity fell behind 70% of ZoC range, i.e., a high percentage of ZoC to be considered. In order to secure a good comprehension of the scene, the threshold was fixed at two thirds of the total Shibata's ZoC. Further, the vergence-acommodation conflict was found to be nearly unsolvable or, at least, it took a lot of time to be solved. This time-related effect will be dealt within subsequent sections. Out of ZoC, violations were easily detected by the algorithm analyzing the resultant parallax histogram. Other still images were used to quantify

window violation cue conflict. During the sequence, the text is turning and, from time to time, some of that text crosses the screen's limits. As the text has a negative parallax, it should never touch the borders. Out of tests results, it was determined that window violation cue conflict became difficult to overcome when at least 20% of the screen edge was filled with negative parallax pixels. Again, we were able to detect window violations measuring negative parallax pixels over the edges from computed depth maps results.

The last still image test was related to quality of experience rather than annoyance or discomfort. In this case, a set of images was ranked for their 3D effectiveness. The results were compared to their depth map histogram distribution. Figures 18, 19, 20, 21 show depth maps and histograms of the images submitted to test. Table 1 holds variance statistics for all the images tested. Note that histogram value for −80 pixels always shows a peak. This peak is considered as noise related to depth map calculation techniques and will not be taken into account when statistics are calculated.

When asked about the 3D perception between the church and the cemetery image, the observers would be inclined to prefer the second image because there is a wider range of depth. This is statistically measured as a bigger positive parallax variance. Library and, above all, table, were found to be the favorite images due to its variety of depths, from positive to negative parallax.

The conclusions derived from the subjective tests revealed that distribution of parallax, especially negative parallax, is one of the main sources of visual discomfort but, on the other side, is the more attractive source of 3D enjoyment, so an agreement between these two concepts must be found. Our developed tools confirmed these features in each of the images through statistical depth analysis, which led us to believe that the system is
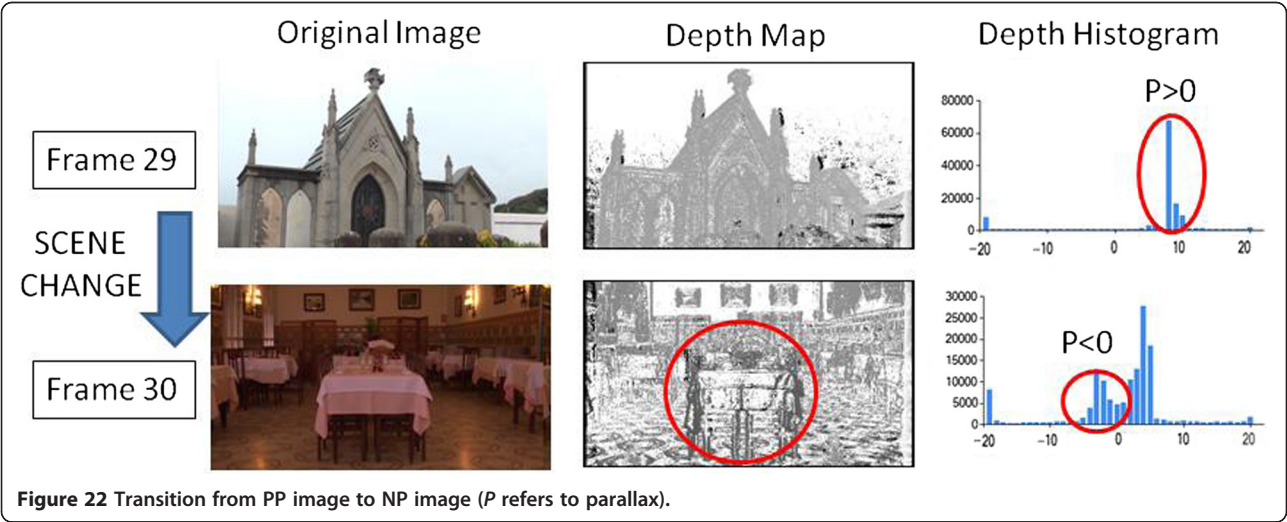

Figure 22 Transition from PP image to NP image (*P* refers to parallax).

Table 2 Positive and negative parallax from histogram analysis from sequence Modernism

Frame	Positive parallax (% pixels)	Negative parallax (% pixels)
29	44.26	2.60–
30	31.58	18.15

well suited to detect possible indications of 3D quality of experience through objective analysis of still images.

5.2 Depth and motion in video sequences affecting to perceived quality

First of all, the scene changes were analyzed with different variations of negative parallax in an environment of positive high-variance. As seen in Figure 22, the depth map and histogram are calculated and their related statistics are evaluated to detect the scene changes. The statistics related to both depth histograms is included in Table 2.

The variation of negative parallax from frame 29 to the next one is more than a 15%, taking into account that although the negative parallax in frame 29 tends to zero, the positive parallax is very significant, with more than a 25%, which means that there is a high probability of detecting visual discomfort, as observers manifested in subjective tests, who needed time to focus the objects in negative areas. Similar results have been obtained with scene changes with variations of negative parallax higher than 10%. Tests related to motion with high parallax variation offer similar results.

Fast motion is detected in some sequences when the motion vectors reveal a movement higher than 2 pixels per frame. In Figure 23, both negative and positive parallax percentage are presented in parallel to motion description. It is remarkable that the abrupt increase of negative parallax is not enough for visual discomfort to be detected. It is necessary to perform an environment with parallax variance and motion. The probability of discomfort appearance is higher when motion is fast than when motion is slower.
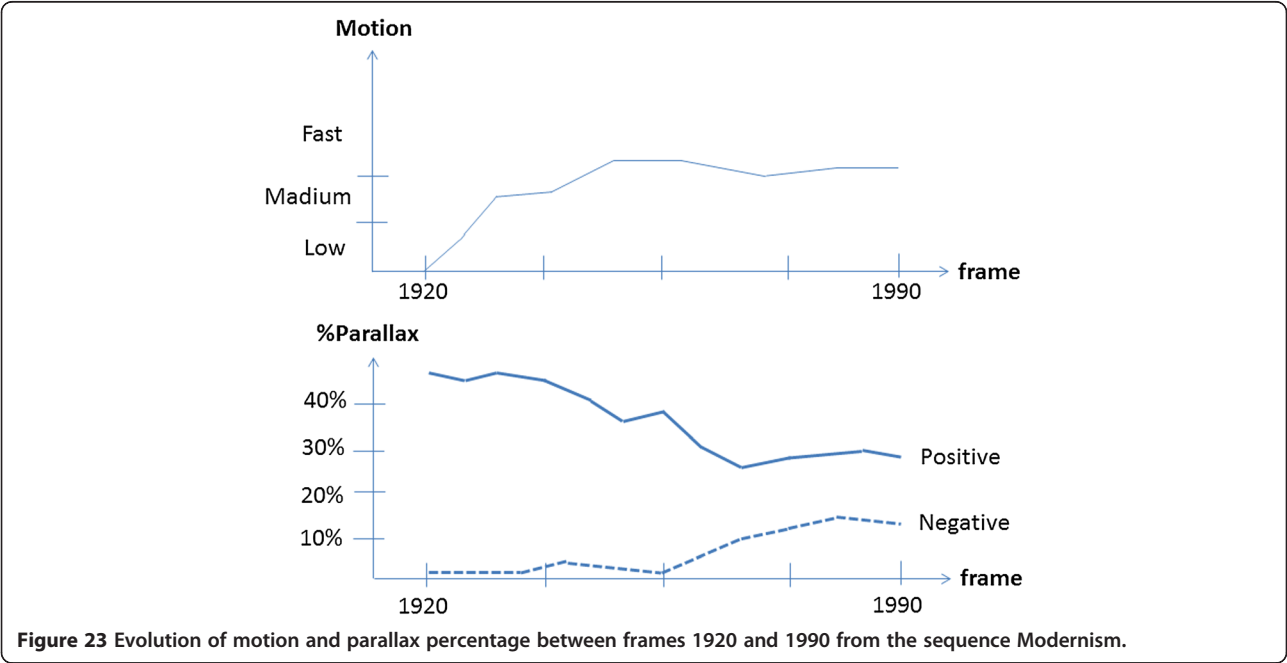
6 Conclusions

Depth and motion are the main factors in perceived quality of experience. Information provided by depth maps and estimated motion vectors is useful to avoid effects that can cause visual discomfort and fatigue in observers when contemplating 3D stereoscopic contents.

Subjective assessment allowed us to isolate the main features to be detected, in order to perform an algorithm which could translate the user's opinion to an automatic objective system.

Information obtained from depth maps and its associated histograms and parallax distribution are the main features that were decided to use in developing our quality measurement algorithm.

The presence of objects with negative parallax on a static image, and especially when motion is detected in the video sequence which contains that image, requires quantifying the probability of observer's annoyance. In graphics comparing parallax and motion, evolution is remarkable to the relationship between both parameters in the final experience of users. Previous ZoC studies have been found to be greatly affected by motion and time of visualization, diminishing its range significantly.

Figure 23 Evolution of motion and parallax percentage between frames 1920 and 1990 from the sequence Modernism.

Parallaxes getting near the ZoC edges (especially negative parallax) have been proved to be undesirable when fast motion or high parallax variance appears.

The tests that have been developed through this work show good results when applying the techniques to video sequences that contain effects which could be considered annoying for the human eye. Algorithms were validated with a collection of sequences with positive results.

Results obtained offer guidelines for stereoscopic video creation, extracting the probabilities of visual discomfort and fatigue and reaching a consensus between 3D perception and annoyance over observer's eye. Nevertheless, in this context, the content provider and the user have the final decision to accept or not a particular content.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

References

1. P Le Callet, S Moller, A Perkis, Qualinet white paper on definitions of quality of experience (2012). European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), version 1.2, (Qualinet, Lausanne, 2013)
2. Z Wang, AC Bovik, L Lu, Why is image quality assessment so difficult? IEEE Int. Conf. Acoust., Speech, Signal Process. **4**, 3313–3316 (2002)
3. S Winkler, *Digital video quality: vision models and metrics* (Wiley, Hoboken, New Jersey, USA, 2005)
4. HR Wu, KR Rao, *Digital Video Image Quality and Perceptual Coding (Signal Processing and Communications)* (CRC, Taylor and Francis Group, London, 2005)
5. Z Wang, A Bovik, *Modern Image Quality Assessment (Synthesis Lectures on Image, Video, & Multimedia Processing)* (Morgan & Claypool, San Rafael, CA, USA, 2006)
6. ITU-R. Recommendation ITU-R BT.500-11, standardization sector of ITU, 2002. *Methodology for the subjective assessment of the quality of television pictures*, 2002
7. ITU-T P.910. *Recommendation Telecom. Standardization Sector OF ITU Subjective video quality assessment methods for multimedia applications*, 2008
8. JA Rodrigo, D Jiménez, JM Menéndez, *Real-Time 3-D HDTV Depth Cue Conflict Optimization, in 2011 IEEE International Conference on Consumer Electronics ICCE, Berlin*, 2011
9. YJ Jung, H Sohn, S-I Lee, F Speranza, YM Ro, Visual importance- and discomfort region-selective low-pass filtering for reducing visual discomfort in stereoscopic displays. IEEE Trans Circ Syst Video Tech **23**(8), 1408–1421
10. P Hanhart, F De Simone, T Ebrahimi, *Quality assessment of asymmetric stereo pair formed from decoded and synthesized views, in 2012 Fourth International Workshop on Quality of Multimedia Experience (QoMEX)* (Yarra Valley, 2012), pp. 236–241
11. E Bosc, R Pepion, P Le Callet, M Koppel, P Ndjiki-Nya, M Pressigout, L Morin, Towards a new quality metric for 3-D synthesized view assessment. IEEE J. Sel. Top. Signal Process. **5**(7), 1332–1343 (2011)
12. W Chen, J Fournier, M Barkowsky, P Le Callet, *Quality of experience model for 3DTV, in Proceedings of the SPIE 8288, Stereoscopic Displays and Applications XXIII* (Burlingame, 2012)
13. T Wa James, F Speranza, S Yano, K Shimono, H Ono, Stereoscopic 3D-TV: visual comfort, broadcasting. IEEE Trans. **57**(2), 335–346 (2011)
14. F Speranza, WJ Tam, R Renaud, N Hur, Effect of disparity and motion on visual comfort of stereoscopic images. Proc. of SPIE **6055**, 94–103 (2006)
15. YJ Jung, S Lee, H Sohn, HW Park, YM Ro, Visual comfort assessment metric based on salient object motion information in stereoscopic video. J. Electron Imaging **21**(1) (2012)
16. C Richardt, L Swirski, IP Davies, NA Dodgson, Predicting stereoscopic viewing comfort using a coherence-based computational model, in *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging* (ACM, New York, 2011), pp. 97–104
17. J Li, M Barkowsky, P Le Callet, *The influence of relative disparity and planar motion velocity on visual discomfort of stereoscopic videos, in Proceedings of the International Workshop on Quality of Multimedia Experience QoMEX* (Mechelen, Belgique, 2011), pp. 1–6
18. J Li, M Barkowsky, P Le Callet, *Visual discomfort is not always proportional to eye blinking rate: exploring some effects of planar and in-depth motion on 3DTV QoE, in Proceedings of VPQM 2013* (États-Unis, Scottsdale, 2013), pp. 1–6
19. C Sang-Hyun, K Hang-Bong, *Subjective evaluation of visual discomfort caused from stereoscopic 3D video using perceptual importance map, in IEEE Region 10 Conference, TENCON 2012* (Cebu, 2012), pp. 1–6
20. T Kun-Lung, H Wei-Jia, L An-Chun, H Wei-Hao, Y Yin-Chun, C Wen-Chao, *Automatically optimizing stereo camera system based on 3D cinematography principles, in 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)* (Zurich, 2012), pp. 1–4
21. W Chen, J Fournier, M Barkowsky, P Le Callet, *New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone. Paper presented at the stereoscopic displays and applications XXII* (San Francisco, California, 2011), pp. 786310–786313
22. AS Percival, The relation of convergence to accommodation and its practical bearing. Ophtalmol. Rev. **11**, 313–328 (1892)
23. T Shibata, J Kim, DM Hoffman, MS Banks, The zone of comfort: predicting visual discomfort with stereo displays. J. Vision **11**, 1–28 (2011)
24. MS Banks, JCA Read, RS Allison, SJ Watt, Stereoscopy and the human visual system. SMPTE Motion Imaging J. **121**, 24–43 (2012)
25. P Leclercq, J Morris, *Assessing stereo algorithm accuracy, in Proceedings of Image and Vision Computing, IVCNZ '02* (, Auckland, 2002)
26. P Leclercq, J Morris, *Robustness to Noise of Stereo Matching, in Proceedings of 12th International Conference on Image Analysis and Processing* (Mantova, Italy, 2003)
27. D Scharstein, R Szeliski, A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. Int. J. Comput. Vis. **47** (2002)
28. *The Middlebury Database - Stereo Evaluation (version 2)*. http://vision.middlebury.edu/stereo/eval/
29. Y Miyajima, T Maruyama, *A real-time stereo vision system with FPGA, in Proceedings of the 30th Conference of IEEE Industrial Electronics Society* (Lisbon, 2003)
30. J Yi, J Kim, L Li, J Morris, G Lee, P Leclercq, Real-time three dimensional vision. Lect. Notes Comput. Sci **3189**, 309–320 (2004)
31. S Lee, J Yi, J Kim, Real-time stereo vision on a reconfigurable system. Lect. Notes Comput. Sci **3553**, 299–307 (2005)
32. K Ambrosch, M Humenberger, W Kubinger, A Steininger, *SAD-based stereo matching using FPGAs, embedded computer vision, part II* (Springer, Dordrecht, 2009)
33. R Zabih, J Woodfill, *Non-parametric local transforms for computing visual correspondence, in Proceedings of 3rd European Conference on Computer Vision* (Stockholm, 1994), pp. 150–158
34. A Wilson, Census transform brings stereo to embedded systems, OptoIQ. Walling, Fran, Vision Systems Design **11**(9), 14 (2006)
35. B Cynagek, *Adaptive window growing technique for efficient image matching* (Technical Report (AGH-University of Science and Technology, Krakow, 2004)
36. J Woodfill, B Von Herzen, *Real-time stereo vision on the PARTS reconfigurable computer, in Proceedings of the 5th IEEE Symposium on FPGAs for Custom Computing Machines* (Napa Valley, 1997)
37. S Jin, J Cho, X Pham, K Lee, S-K Park, M Kim, J Jeon, FPGA Design and Implementation of a Real-Time Stereo Vision System. IEEE Trans. Circ. Syst. Video Tech. **20**(1), 15–26 (2010)

38. NY-C Chang, T-H Tsai, B-H Hsu, Y-C Chen, T-S Chang, Algorithm and architecture of disparity estimation with mini-census adaptative support weight. IEEE Trans. Circ. Syst. Video Tech. **20**(6), 792–805 (2010)

39. X Mei, X Sun, M Zhou, S Jiao, H Wang, X Zhang, *On building an accurate stereo matching system on graphics hardware, Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, vol., no., pp.467,474, 6–13 Nov. 2011, Barcelona, Spain*

40. S Chambon, A Crouzil, *Combination of correlation measures for dense stereo matching, in International Joint Conference on Computer Vision Theory and Applications, VISAPP* (France, 2011)

41. Y Nojiri, H Yamanoue, S Ide, S Yano, F Okana, *Parallax distribution and visual comfort on stereoscopic HDTV, in Proceedings of IBC* (, Lisbon, 2006), pp. 373–380

42. *Professional 46-inch 3D display monitor GD-463D10U*. http://pro.jvc.com/prof/attributes/features.jsp?model_id=MDL101867 Accessed April 2009

43. E Winkler, *On the properties of subjective ratings in video quality experiments, in Proceedings of the International Workshop on Quality of Multimedia Experience (QoMEX)* (San Diego, 2009)

44. *EBU, EBU Test Sequences*. http://tech.ebu.ch/testsequences. Accessed 20 May 2006

45. *ITU-R, Recommendation BT.1428. Subjective assessment of stereoscopic television pictures (question ITU-R 234/11)*, 2001