Visual discomfort of stereoscopic 3D videos: Influence of 3D motion

Jing Li*, Marcus Barkowsky, Patrick Le Callet

LUNAM Université, Université de Nantes, IRCCyN UMR CNRS 6597, Polytech Nantes, rue Christian Pauc BP 50609, 44306 Nantes Cedex 3, France

ARTICLE INFO

Article history:
Received 6 April 2013
Received in revised form 2 January 2014
Accepted 12 January 2014
Available online 22 January 2014

Keywords:
Visual discomfort
3D motion
Stereoscopic video
Paired comparison
Objective model

ABSTRACT

Visual discomfort is one of the most frequent complaints of the viewers while watching 3D images and videos. Large disparity and large amount of motion are two main causes of visual discomfort. To quantify this influence, three objectives are set in this paper. The first one is the comparative analysis on the influence of different types of motion, i.e., static stereoscopic image, planar motion and in-depth motion, on visual discomfort. The second one is the investigation on the influence factors for each motion type, for example, the disparity offset, the disparity amplitude and velocity. The third one is to propose an objective model for visual discomfort. Thirty-six synthetic stereoscopic video stimuli with different types of motion are used in this study. In the subjective test, an efficient paired comparison method called Adaptive Square Design (ASD) was used to reduce the number of comparisons for each observer and keep the results reliable. The experimental results showed that motion does not always induce more visual discomfort than static conditions. The in-depth motion generally induces more visual discomfort than the planar motion. The relative disparity between the foreground and the background, and the motion velocity are identified as main factors for visual discomfort. According to the subjective results, an objective model for comparing visual discomfort induced by different types of motion is proposed which shows high correlation with the subjective perception.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

As one of the three dimensions of Quality of Experience (QoE) in 3DTV [1], visual discomfort is getting more and more attention recently as viewers often complain about it while or after watching stereoscopic 3D images and videos. With the boom of stereoscopic 3D movies, visual discomfort becomes important and urgent as it is related to the safety and health issues of viewers.

The concepts of visual discomfort and visual fatigue in 3DTV are often confused and interchangeably used. In [2,3], the authors provided distinct definitions for them. In this paper, we adopt the definition from [4], the details are shown as follows:

- Visual discomfort is a physical and/or a psychological state assessed by the participants themselves, as a presently perceived degree of annoyance. As such, it may be related to experienced symptoms, perceived difficulties when performing a visual task, or any negative sensation associated with this task. Thus, visual discomfort can be measured by asking the viewer to report its level. Visual discomfort appears and disappears with any of these negative associations, and is supposed to have a steep rise time and a steep fall time contrary to visual fatigue. In other words, visual discomfort disappears rapidly when the visual task is interrupted, either by asking the observer to close his eyes or by terminating the visual stimulus. In this paper, we only focus on the visual discomfort.
- Visual fatigue is defined in this context as a syndrome. Its presence is assessed by the observation of zero, one or several symptoms (reported by the observer such as nausea) and zero, one or several clinical signs (measured objectively such as eye blinking rate). Usually it is supposed to have a longer rise time and a longer fall time than visual discomfort, it is not instantaneously diagnosed in conjunction with a certain 3D stimulus and remains a certain time after the 3D viewing has finished.

In this paper, we focus on the visual discomfort issue. The vergence–accommodation conflict is one of the widely accepted causes of visual discomfort [5,6]. When viewing an object by means of a 3D screen, the eyes converge to the virtual object position but the accommodation has to be performed at the screen depth level, which is unnatural and will not happen in our daily life. The larger this discrepancy on vergence and accommodation, the more possibility that observers perceive visual discomfort.

The comfortable viewing zone, which determines the threshold of the distance between the virtual object and the screen that
would not induce visual discomfort, was thus investigated and defined in different ways. For example, ±0.2D of depth of field (DOF) [7–9], ±1 arc degree of visual angle [10,11] or ±3% of the screen width for 3D television [3].

The comfortable viewing zone is well investigated for stereoscopic still images. However, recent studies showed that fast motion can induce visual discomfort even if the object is within the comfortable viewing zone [9,11,12]. Thus, motion in stereoscopic videos, including planar motion and in-depth motion, is considered as an important factor of visual discomfort in 3D video sequences.

In-depth motion is one of the significant factors that may cause visual discomfort. Studies already showed that visual discomfort increases with the in-depth motion velocity [3,9,11,13,14]. However, the influence from disparity amplitude (disparity range) and the disparity type (crossed or uncrossed) of in-depth motion on visual discomfort are still under study. In [11], the results showed that disparity amplitude of the moving object is not a main factor. However, in their recent study [3] it is shown that visual discomfort increases with the disparity amplitude. Furthermore, the results also showed that the in-depth motion with crossed disparity would induce significantly more visual discomfort than the uncrossed and mixed conditions. In [13], as they only analyzed the in-depth motion in the disparity range of ±1° with different velocities, there is no conclusion about the influence of crossed or uncrossed disparity amplitude on visual discomfort.

The influence of the planar motion on visual discomfort was investigated as well [3,12,13,15]. These studies showed high consistency on the conclusion that visual discomfort increases with the motion velocity. The influences of the disparity on visual discomfort led to different conclusions in these studies. In [13], the results indicated that the disparity type, i.e., crossed and uncrossed disparity, did not affect the visual discomfort thresholds. However, in [3,12,15], the results showed that the crossed disparity will generate more visual discomfort than the uncrossed disparity. The possible explanation for this conflict might be the position of the background. In [13], the background was positioned at the screen plane. In [12,15], the background was placed at a fixed position of 46.28 cm behind the screen (with disparity of −1.4°) and in [3], the position of the background was not depicted but in their previous study [11], the background was positioned at 134 cm away from the camera (with disparity of −2.6°). The impact of the position of background on visual discomfort may therefore require further study.

Most of the studies mentioned above investigated the influence of the in-depth motion and planar motion on visual discomfort separately. For quantifying the influence of static situations, planar and in-depth motion, it would be important to directly compare their impact on visual discomfort. Thus, this study focused on the influence of motion on visual discomfort of 3DTV, including the comparative analysis on the influence of different motion types on visual discomfort, the influence of disparity and velocity within a certain motion type and the proposal of an objective visual discomfort model based on the results. As observers may experience difficulties when using traditional 2D quality methods, e.g., Absolute Category Rating (ACR), to assess the visual discomfort induced by stereoscopic images [16], the pair comparison method was used in the subjective test as it would be easier for observers to select their preference in each pair, thus, the results are more reliable.

To make the test feasible, the Adaptive Square Design (ASD) method was employed which reduces the number of comparisons significantly and generates reliable results.

The rest of this paper is organized as follows. Section 2 introduces the subjective experimental setup. Then follows the experimental results in Section 3. The proposed objective model is introduced in Section 4. Section 5 concludes the paper.

### 2. Experiment 1: influence from motion

#### 2.1. Definitions

When displaying 3D images on flat panel stereoscopic displays, the binocular angular disparity can be expressed by degree of visual angle [17], as shown in Fig. 1, the binocular angular disparities \( \phi_A \) and \( \phi_B \) for points A and B can be calculated by Eqs. (1) and (2).

For a point which is on the screen plane, the binocular angular disparity is 0° which means there is no disparity between the two retinal images. A positive value represents crossed disparity, such as the point A; a negative value represents uncrossed disparity, such as the point B.

\[
\phi_A = \beta - \alpha \quad (1)
\]

\[
\phi_B = \gamma - \alpha \quad (2)
\]

Motion in 3DTV may be decomposed into planar motion (or lateral motion) and in-depth motion. Planar motion means that the object moves only in a plane parallel to the screen, i.e., the disparity does not change temporally. In-depth motion, which is also called motion in depth or z-motion, is defined as that the object moves towards or away from the observers. For planar motion, both eyes make the same conjunctive eye movements, called version [14]. For in-depth motion, the eyes make opposite, disjunctive eye movement, called vergence [18]. The eye movements for the planar motion and in-depth motion are shown in Fig. 2.

The speed of the planar motion and in-depth motion can be expressed by the change of distance per second or the change of the visual angle (version or vergence) per second. The change of the visual angle is chosen in the study with the unit of “deg/s”.

In this study we use disparity amplitude and disparity offset to define the motion in stereoscopic videos. The disparity amplitude \( d_A \) between points A and B can be expressed by Eq. (3) which represents the difference of angular disparity between the two points. The disparity offset \( d_o \) between the two points A and B can be expressed by Eq. (4) which represents the center of the angular disparity between the two points.

The static and planar motion stimuli can be expressed by the disparity offset and the velocity, where the disparity amplitude equals zero. The in-depth motion stimuli can be defined by the disparity amplitude, the disparity offset and the velocity.

\[
d_A = |\phi_A - \phi_B| \quad (3)
\]

\[
d_o = \frac{1}{2} (\phi_A + \phi_B) \quad (4)
\]

#### 2.2. Experimental design

To avoid the influence of other factors on visual discomfort, 36 synthetic video stimuli were used in this study for precise control...
on the possible influence factors, including motion type, velocity, disparity offset and disparity amplitude.

For the planar motion stimuli, we keep them consistent with our previous experiment [12,15]. Five angular disparity offset levels (0°, ±0.65°, and ±1.3°) and three velocity levels (4, 10, and 16 deg/s, which represent slow, medium, and fast, respectively) were selected assuming that the interpupillary distance was 65 mm and the viewing distance was fixed to 90 cm. A background was placed at a fixed position (−1.4°) with a virtual distance of 136.28 cm away from the observers. This design is consistent with a typical natural video content where the background is almost fixed and placed behind the screen. Fig. 3 shows the disparities used in the planar motion stimuli and their relationship with the comfortable viewing zone.

For the static condition, five disparity offset levels were selected which were the same as the planar motion design. The foreground object keeps static in the center of the screen at a certain disparity plane.

For the in-depth motion condition, four disparity amplitude levels (0.65°, 1.3°, 2° and 2.6°), three disparity offset levels (−0.65°, 0°, 0.65°) and three velocity levels (1, 2, and 3 deg/s, binocular angular degree) were selected. The reason for choosing binocular angular disparity speed was that the object’s velocity appears visually constant which is not the case for a constant value in the unit of cm/s. The direction of the movement is inverted at the far or at the near end of the movement so the object in the experiment moved forth and back in an endless loop. The three velocity levels 1, 2 and 3 deg/s represent slow, medium and fast, respectively. Fig. 4 shows the design of the disparity amplitude and disparity offset for the in-depth motion.

The 15 planar motion stimuli, 5 static stimuli and 16 in-depth motion stimuli used in the subjective experiment are listed in Table 1 with their stimulus serial number, disparity offset $d_o$, disparity amplitude $d_a$, planar motion velocity $v_p$ and in-depth motion velocity $v_d$.

2.3. Apparatus

The stereoscopic sequences were displayed on a Dell Alienware AW2310 23-inch 3-D LCD screen (1920 × 1080 full HD resolution, 120 Hz), which featured 0.265-mm dot pitch. The display was adjusted for a peak luminance of 50 cd/m² when viewed with the active shutter glasses. The graphics card of the PC was an NVIDIA Quadro FX 3800. Stimuli were viewed binocularly through the NVIDIA active shutter glasses (NVIDIA 3D vision kit) at a distance of about 90 cm, which was approximately three times the picture height. The peripheral environment luminance was adjusted to about 44 cd/m². When seen through the eye-glasses, this value corresponded to about 7.5 cm/m² and thus to 15% of the screen’s peak brightness as specified by ITU-R BT.500 [19].

2.4. Stimuli

The stereoscopic sequences consisted of a left-view and a right-view image which were generated by the MATLAB psychtoolbox [20,21]. Each image contained a foreground object and a static background. A black Maltese cross which was frequently used in such kind of psychometric experiment [22,23] was used as the foreground object with a resolution of 440 × 440 and the visual angle of 7.6°. As it contained both high and low spatial frequency

Fig. 2. The left figure shows the eye movement of the planar motion object. Planar motion velocity is the amount of the change of the version per second. The right figure shows the eye movement of the in-depth motion object. In-depth motion velocity is the amount of the change of the vergence per second. $A_0$ represents the perceived virtual object at frame $N$, $A_{N+1}^L$ and $A_{N+1}^R$ represent the left and right view images on the screen at frame $N$.

Fig. 3. The relationship of the foreground and the background position and the comfortable viewing zone in planar motion stimuli.
components, it was supposed to limit the influence of one particular spatial frequency in the experiment [24].

The background was generated by adding salt and pepper noise to a black image of full HD resolution, and then filtered by a circular averaging filter. The reason for using this kind of image as the common background of all stimuli was that it could preclude all of the monocular cues on stereopsis.

For the planar motion stimuli, the trajectory of the moving object is a circle with center point at the center of the screen, and radius of 300 pixels, approximately $10^\circ$ of visual angle. The motion direction of the object was anti-clockwise. An example of the stimuli is shown in Fig. 5, in which the foreground object is placed in front of the screen with an angular disparity of $1.3^\circ$. For the static stimuli, the Maltese cross was positioned at the center of the screen. For the in-depth motion stimuli, the Maltese cross was positioned in the center of the screen and moved back and forth to the observers. An example is shown in Fig. 6, in which the foreground object is moving in the depth plane with disparity amplitude of $2.6^\circ$ and offset of $0^\circ$. All these stimuli can be downloaded from the website (will be provided after the first round of review).

2.5. Viewer

Forty two viewers participated in this subjective experiment. Twenty one are male, 21 are female. They are all non-experts in subjective experiment, image processing or 3D related field. Their age ranged from 19 to 48 years with an average age of 26.8. All have either normal or corrected-to-normal visual acuity. The visual acuity test was conducted with a Snellen Chart for both far and near vision. The Randot Stereo Test was applied for stereo vision acuity check, and Ishihara plates were used for color vision test. All of the viewers passed the pre-experiment vision check.

2.6. Assessment method

The paired comparison method was used in this test. To reduce the number of comparisons, the ASD method was used which was proposed in [25] and validated by [26]. The detail of the ASD method is:

![Fig. 4. The disparity amplitude and offset design for in-depth motion stimuli. The arrows represent the depth interval in which the object moves.](image)

![Table 1](image)

<table>
<thead>
<tr>
<th>Number</th>
<th>$d_o$ ($^\circ$)</th>
<th>$d_a$ ($^\circ$)</th>
<th>$v_p$ (deg/s)</th>
<th>$v_d$ (deg/s)</th>
<th>BT score</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.3</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0.42</td>
</tr>
<tr>
<td>2</td>
<td>-1.3</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>1.70</td>
<td>0.33</td>
</tr>
<tr>
<td>3</td>
<td>-1.3</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>3.56</td>
<td>0.27</td>
</tr>
<tr>
<td>4</td>
<td>-0.65</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1.03</td>
<td>0.37</td>
</tr>
<tr>
<td>5</td>
<td>-0.65</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>2.09</td>
<td>0.31</td>
</tr>
<tr>
<td>6</td>
<td>-0.65</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>3.40</td>
<td>0.27</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1.80</td>
<td>0.35</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>2.44</td>
<td>0.31</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>3.61</td>
<td>0.26</td>
</tr>
<tr>
<td>10</td>
<td>0.65</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>2.36</td>
<td>0.31</td>
</tr>
<tr>
<td>11</td>
<td>0.65</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>2.84</td>
<td>0.29</td>
</tr>
<tr>
<td>12</td>
<td>0.65</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>3.85</td>
<td>0.27</td>
</tr>
<tr>
<td>13</td>
<td>1.3</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>3.10</td>
<td>0.27</td>
</tr>
<tr>
<td>14</td>
<td>1.3</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>3.58</td>
<td>0.27</td>
</tr>
<tr>
<td>15</td>
<td>1.3</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>4.22</td>
<td>0.26</td>
</tr>
<tr>
<td>16</td>
<td>-1.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.90</td>
<td>0.38</td>
</tr>
<tr>
<td>17</td>
<td>-0.65</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.39</td>
<td>0.30</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.89</td>
<td>0.26</td>
</tr>
<tr>
<td>19</td>
<td>0.65</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.86</td>
<td>0.24</td>
</tr>
<tr>
<td>20</td>
<td>1.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7.39</td>
<td>0.25</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>1.3</td>
<td>0</td>
<td>1</td>
<td>4.70</td>
<td>0.25</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>1.3</td>
<td>0</td>
<td>2</td>
<td>5.35</td>
<td>0.24</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>1.3</td>
<td>0</td>
<td>3</td>
<td>5.76</td>
<td>0.24</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>4.82</td>
<td>0.25</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>5.45</td>
<td>0.24</td>
</tr>
<tr>
<td>26</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>6.14</td>
<td>0.24</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>2.6</td>
<td>0</td>
<td>1</td>
<td>4.99</td>
<td>0.25</td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>2.6</td>
<td>0</td>
<td>2</td>
<td>5.95</td>
<td>0.24</td>
</tr>
<tr>
<td>29</td>
<td>0</td>
<td>2.6</td>
<td>0</td>
<td>3</td>
<td>6.24</td>
<td>0.24</td>
</tr>
<tr>
<td>30</td>
<td>-0.65</td>
<td>1.3</td>
<td>0</td>
<td>1</td>
<td>4.00</td>
<td>0.26</td>
</tr>
<tr>
<td>31</td>
<td>-0.65</td>
<td>1.3</td>
<td>0</td>
<td>3</td>
<td>5.50</td>
<td>0.24</td>
</tr>
<tr>
<td>32</td>
<td>-0.65</td>
<td>0.65</td>
<td>0</td>
<td>1</td>
<td>3.97</td>
<td>0.27</td>
</tr>
<tr>
<td>33</td>
<td>-0.65</td>
<td>0.65</td>
<td>0</td>
<td>3</td>
<td>5.00</td>
<td>0.25</td>
</tr>
<tr>
<td>34</td>
<td>0.65</td>
<td>1.3</td>
<td>0</td>
<td>1</td>
<td>5.34</td>
<td>0.24</td>
</tr>
<tr>
<td>35</td>
<td>0.65</td>
<td>1.3</td>
<td>0</td>
<td>3</td>
<td>6.04</td>
<td>0.24</td>
</tr>
<tr>
<td>36</td>
<td>0.65</td>
<td>0.65</td>
<td>0</td>
<td>3</td>
<td>5.99</td>
<td>0.24</td>
</tr>
</tbody>
</table>

![Fig. 5. An example of stimulus with planar motion in the experiment. The foreground object is moving at the depth plane with a disparity of $1.3^\circ$. The background is placed at a fixed depth plane of $-1.4^\circ$. The motion direction of the Maltese cross is anti-clockwise.](image)
1. For the first observer, the indices of the 36 stimuli were randomly placed into a square matrix $R$ of size $6 \times 6$. Only the stimulus pair whose indices are in the same column or row of the matrix $R$ are compared.

2. According to the obtained paired comparison data from all observers who have conducted the test, the Bradley–Terry scores (how to obtain Bradley–Terry scores can be found in Section 3.1) and the rank orders of all stimuli can be obtained. Supposing that the ascending ordering index vector for all stimuli is $d = (d_1, d_2, \ldots, d_{36})$, the square matrix $R_{\text{ASD}}$ is arranged in such a way that the elements of the vector $d$ are placed along a spiral as shown in Fig. 7.

3. For the $k_{th}$ observer ($k \geq 2$), only the stimulus pairs $\{S_i S_j\}$ are compared if and only if $(i, j) \in C$, where $C$ is defined as:

$$C = \{(x, y) \mid p = p' \lor q = q', \quad x = r_{pq}, \quad y = r_{p'q'} \text{ in } R_{\text{ASD}}\}$$

In other words, only the stimulus pairs whose indices are in the same column or row of the matrix $R_{\text{ASD}}$ are compared.

4. Repeat steps 2 and 3, until all 42 observers finished the test ($k = 42$). Please note that $d$ and $R_{\text{ASD}}$ are updated for each observer.

According to this method, 36 stimuli lead to in total 180 pairs for each observer. In the subjective test, the viewers watched a pair of stimuli at one trial, and then they were asked to select the one which they feel more uncomfortable. The presentation order of each pair was random and balanced by all observers, i.e., all stimuli had the same frequency on first order and second order of presentation.

2.7. Procedures

The subjective experiment contained a training session and a test session. Five pairs of stimuli were shown in the training session. The observers were asked not to stare at the moving object but to watch the whole stereoscopic sequence. Then, they should select the one which is more uncomfortable, concerning e.g., mental uneasiness. The users viewed two keys to switch between the pair of stimuli on one screen. There was a minimum duration for the display of each stimulus before making a decision by pressing a specified button. The minimum duration is defined as the longer one between 5 s and the duration of a complete cycle of movement (the moving object went back to its start point). During the training session, all questions of the viewers were answered. We ensured that after the training session, all of the viewers knew about the process and task of this experiment clearly.

In the main test session, 180 pairs were compared. As the duration of the whole test was different due to individual differences of each viewer, and to avoid visual fatigue caused by long time watching affecting the experimental results, the test was split into two sub-sessions. Each session contained 90 stimuli. There was a 10-min break between the two sub-sessions.

3. Experimental results

3.1. Bradley–Terry model: convert pair comparison data to scale values

Generally, the outcome of the paired comparison test is a pair comparison matrix $A$, where $A = (a_{ij})_{m \times n}$, $a_{ij}$ is the total count of preference of stimulus $S_i$ over $S_j$ for all observers. $a_{ii} = 0$ for $i = 1, 2, \ldots, m$. The total number of comparisons for stimulus pair $(S_i S_j)$ is $n_{ij} = a_{ij} + a_{ji}$.

The Bradley–Terry (BT) model [27,28] is a well-known model to convert pair comparison data to psychophysical scale values for all stimuli which is a function defined as follows:

$$V_i - V_j = \log \frac{p_{ij}}{1 - p_{ij}}$$

(5)

$P_{ij}$ represents the probability that stimulus $S_i$ is preferred to $S_j$. $P_{ij}$ can be estimated by $p_{ij}$, where $p_{ij} = a_{ij}/n_{ij}$ when $n_{ij}$ is large enough. The outputs are the differences of the BT scale values between stimuli $S_i$ and $S_j$. By utilizing least squares estimation or maximum likelihood estimation, the scale value $V_i$ for each stimulus, $i = 1, \ldots, m$ can be estimated. Please note that the scale value $V_i$ is a relative value which can be added with an offset, but cannot be re-scaled by a factor. In this study, BT scores represent the degree of visual discomfort. The higher the BT scores, the more pronounced the visual discomfort.

The BT scores and confidence intervals of all stimuli are shown in Table 1. For better comparison, the lowest BT score is set to 0.

3.2. Planar motion and static conditions

The BT scores for the planar motion stimuli and static stimuli are shown in Fig. 8 where the static condition can be considered as a special case of the planar motion.

The experimental results on the planar motion stimuli validated the conclusions from [12,15], which were:

- Visual discomfort increases with the planar motion velocity.
- The vergence–accommodation (VA) conflicts might not significantly affect the visual discomfort. As shown in Fig. 8a, the visual discomfort neither reach the minimum at the screen plane nor increase with the absolute value of disparity offset. The possible explanation might be the existence of the background. During the test, the viewers switched their attention between the background and the foreground, the larger this switch distance, the more changes on the mismatch of VA conflict, which leads to more visual discomfort.
The relative disparity $r_o$ between the foreground and the background ($r_o = d_o + 1.4$ in this study) determines the visual discomfort, i.e., visual discomfort increases with the relative disparity.

For the static stimuli as shown in Fig. 8a, the visual discomfort increases with the relative disparity as well. Under the condition of small relative disparity, i.e., $r_o = 0.1$ (d.o. = −1.3°), the visual discomfort induced by static stimuli is less than the planar motion stimuli with medium or fast velocities. However, the gradient of the curve for the static case is steeper than the planar motion conditions. Thus, the visual discomfort increases faster with the disparity offset for the static stimuli than the planar motion stimuli. In the condition of disparity offset equals zero degree ($r_o = 1.4°$), the static stimuli can generate similar visual discomfort as the fast planar motion stimuli. When the relative disparity is larger than $1.4°$, the visualization of static objects seems to induce much more visual discomfort than planar motion stimuli. This result also can be interpreted as that when the relative disparity is increasing and the disparity offset becomes crossed, the planar motion seems to help to reduce visual discomfort when compared to static condition. In our experiment, the observers explained that when watching the planar motion stimuli, it was easier to fuse the Maltese cross compared to the static conditions, in particular when the disparity is crossed.

As shown in Fig. 8b, for the planar motion condition, there might be a minimum in the curve of visual discomfort that would be located at some velocity in between static and the slowest velocity that was included in this study.

### 3.3 In-depth motion and static conditions

The Bradley–Terry scores for in-depth motion stimuli are shown in Fig. 9. According to the results, we may extract the following conclusions:

- As shown in Fig. 9a, generally, disparity amplitude may not affect the visual discomfort significantly. For example, in the condition of $d_o = −0.65°$ and slow velocity, the visual discomfort induced by the stimulus with disparity amplitude of $0.65°$ is not significantly different from the stimulus with disparity amplitude of $1.3°$. However, for the fast motion condition, the impacts of disparity amplitude on visual discomfort were observed.

- Visual discomfort increases with the disparity offset as shown in Fig. 9a. It seems that the vergence–accommodation conflict did not have significant impact on the results. This is in line with the results of the planar motion conditions. The relative disparity might be a main factor.

- As shown in Fig. 9b, visual discomfort increased with the in-depth motion velocity.

The static condition can be considered as a special case of the in-depth motion as well. The BT scores of the static stimuli were compared with the in-depth motion stimuli which are shown in Fig. 10.

As shown in Fig. 10a, the gradient of the curve for in-depth motion is much flatter than the static conditions. When the disparity offset is less than $0.65°$, the in-depth motion will generate more visual discomfort than the static stimuli. For example, when compared with the static stimuli with the disparity of $0°$, all the in-depth motion stimuli in our study generated more visual discomfort. However, when the relative disparity is larger than $2.05°$ ($d_o = 0.65°$), we may extrapolate that the visual discomfort induced by the static stimuli would be much higher than the in-depth motion stimuli.

Considering the velocity, as shown in Fig. 10b, when the relative disparity is less than $2.05°$ ($d_o = 0.65°$), the visual discomfort increases with the velocity. However, if the relative disparity is larger than $2.05°$, the static stimuli might generate more visual discomfort than the in-depth motion stimuli with slow velocity.

### 3.4 Time course analysis

The sensation of visual discomfort may be increasing with the exposure duration of the stimulus. In [29], this observation was verified by experimental results in stereo displays. In our study, the presentation order of the stimuli was randomized and balanced to avoid influence of this effect. Nonetheless, considering that the test duration is longer than a typical ACR or SS test in visual discomfort of 3DTV (e.g., [3,13]), it would be interesting to analyze the effect of time course on the paired comparison results. Aiming at this objective, each sub-session of our test is divided into two sub-subsessions, thus sub-subsession 1, 2, 3, 4 represent the first half, the second half of the sub-session 1, and the first half and the second half of the sub-session 2, respectively. Then, the Barnard’s exact test [30] was applied on each paired comparison data $p_i$ of each two sub-subsessions to check whether they are
significantly different ($p$-value < 0.05). The total number of significantly different pairs between the results of each two sub-subsessions are shown in Table 2.

Under the assumption of time-course effect, if there is time-course effect in our test, the number of significantly different pairs between sub-subsession 1 and 4 should be the largest among all results as sub-subsession 1 is conducted at the beginning of the test and sub-subsession 4 is conducted at the end of the test. To test it, we utilized the Barnard’s exact test again to check whether the numbers of the significantly different pairs between two sub-subsessions are significantly different, e.g., whether 11/413 (for sub-subsession 1 vs 2) is significantly smaller than 7/412 (for sub-subsession 1 vs 4). We tested all combinations, the Barnard’s exact results showed that all of them are not significantly different ($p$-value > 0.05).

According to the obtained test results, there may be no significant time course effect on the paired comparison results. In other words, observers’ discriminability on visual discomfort of the test stimuli did not depend on the temporal position of the stimulus presentation in the experiment. This may be considered as another indication that paired comparison is a reliable subjective assessment method for visual discomfort in 3DTV.

### 3.5. Discussions

In literature it is often mentioned that the motion in stereoscopic videos would induce more visual discomfort than static conditions. However, in this study, a counter indication was found. All three motion types showed that the relative disparity between the foreground and the background is a main factor in visual discomfort, i.e., visual discomfort increases with the relative disparity. The gradient of visual discomfort with relative disparity is the highest for the static stimuli, followed by in-depth and then the planar motion stimuli. This implies that static stimuli induce more visual discomfort when the relative disparity exceeds a...
certain value. This value is approximately 1.4° for planar motion and 2.05° for in-depth motion.

The gradient analysis also reveals that there is no “crossing point” between the planar motion and the in-depth motion in the positive three-quarters of the disparity space, i.e., \( d_p \) from \(-0.65°\) to \(1.3°\). The in-depth motion stimuli are always more uncomfortable than the planar motion stimuli in this study. However, for the condition that the disparity offset is less than \(0.65°\), we may extrapolate that the slow in-depth motion stimuli might generate less visual discomfort than the fast planar motion stimuli. However, further studies are required.

4. Linear regression analysis: towards an objective visual discomfort model

To investigate the influence factors of each motion type, multiple linear regression analysis is used in this study which attempts to model the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data. For the static situation, there is only one possible factor which is the relative angular disparity. For motion stimuli, the relative disparity offset, disparity amplitude, planar motion velocity, in-depth motion velocity, and their interactions are possible factors. Inspired by the influence of frequency characteristics of in-depth motion on visually induced motion sickness [31] where it has been analyzed whether the oscillation frequency of the in-depth motion also has effect on visual discomfort in 3DTV, another factor called “frequency” was added for analysis in our study. It can be obtained by \( \frac{v_p}{d_p} \).

The stepwise regression function in Matlab was used to select the most significant factors or remove the least significant factors[32]. The output of the stepwise regression includes the estimates of the coefficients for all potential factors, with confidence intervals, the statistics for each factor and for the entire model. To avoid model overfitting, the Leave-one-out Cross Validation (LOOCV) method was used to all possible models to find the model with the minimum averaged RMSE. The final selected models are shown in Table 3.

All the factors shown in Table 3 are statistical significant factors with \( p-value \) of the Student’s-\( t \)-test <0.05. The \( coefficient \) in the table is the coefficient of the corresponding factor in the linear model. The \( model \ analysis \) shows the linear model for each motion type. The \( R^2 \), RMSE, the \( F \)-statistic and its \( p-value \) are provided as the evaluation results of this model. It is shown that for the planar motion stimuli, the relative disparity, planar motion velocity and their interaction term are important factors for visual discomfort. For the static stimuli, the relative disparity offset in this study shows its predominant effect. For the in-depth motion stimuli, the disparity amplitude is not a main factor which is consistent with the conclusions of Section 3.3 and [11]. However, the interaction term of the disparity amplitude and the velocity, and the combination of the three factors (velocity, disparity amplitude, relative disparity offset) play important roles in determining visual discomfort. In addition, unlike the visually induced motion sickness condition, the frequency of the oscillation of the in-depth motion is not a significant influence factor on visual discomfort in 3DTV.

According to the regression analysis results, an objective model for comparing visual discomfort of still stereoscopic images, planar motion stimuli and in-depth motion stimuli is developed. The unit for disparity and velocity are all in visual angular degree. Here we summarize it as:

\[
V = \begin{cases}
2.53r_o + 0.54 & \text{static condition} \\
1.45r_o + 0.28v_p - 0.07r_o v_p - 1.11 & \text{planar motion} \\
0.31r_o + 1.23v_d + (0.45 - 0.21r_o) d_p v_d + 2.51 & \text{in-depth motion}
\end{cases}
\]  

(6)

The scatter plot of the objective and subjective results is shown in Fig. 11. The Pearson Linear Correlation Coefficient (PLCC), Spearman’s Rank Correlation Coefficient (SROCC) and Root Mean Square Error (RMSE) are used to evaluate the correlation between the objective scores and the subjective results, they are 0.9976, 0.9967, and 0.1198, respectively.

As this model is based on the paired comparison results, the absolute value of \( V \) can be used to compare the degree of visual discomfort between the stimuli, where the difference can be interpreted as the probability that one is preferred to another.

![Fig. 11. The scatter plot of the predicted scores and the BT scores.](image-url)

### Table 3

The linear regression analysis results for all stimuli.

<table>
<thead>
<tr>
<th>Motion type</th>
<th>Factor analysis</th>
<th>Coefficient</th>
<th>( p-value ) (( t )-test)</th>
<th>Model analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planar motion</td>
<td>( r_o )</td>
<td>1.4466</td>
<td>0.0000</td>
<td>Intercept = -1.1112</td>
</tr>
<tr>
<td></td>
<td>( v_p )</td>
<td>0.2759</td>
<td>0.0000</td>
<td>( R^2 = 0.9837 ), RMSE = 0.1678</td>
</tr>
<tr>
<td></td>
<td>( r_o \times v_p )</td>
<td>-0.0739</td>
<td>0.0000</td>
<td>( F = 221.474 ), ( p-value = 4.10 \times 10^{-10} )</td>
</tr>
<tr>
<td>Static stimuli</td>
<td>( r_o )</td>
<td>2.5312</td>
<td>0.0000</td>
<td>Intercept = 0.5438,</td>
</tr>
<tr>
<td></td>
<td>( d_p )</td>
<td>0.0739</td>
<td>0.0000</td>
<td>( R^2 = 0.9975 ), RMSE = 0.1517</td>
</tr>
<tr>
<td></td>
<td>( v_d )</td>
<td>0.0739</td>
<td>0.0000</td>
<td>( F = 1176.37 ), ( p-value = 5.45 \times 10^{-5} )</td>
</tr>
<tr>
<td>In-depth motion</td>
<td>( d_p )</td>
<td>0.4538</td>
<td>0.0000</td>
<td>Intercept = 2.5088</td>
</tr>
<tr>
<td></td>
<td>( v_d )</td>
<td>0.3114</td>
<td>0.0000</td>
<td>( R^2 = 0.9817 ), RMSE = 0.1120</td>
</tr>
<tr>
<td></td>
<td>( r_o \times d_p \times v_d )</td>
<td>-0.2110</td>
<td>0.0031</td>
<td>( F = 147.181 ), ( p-value = 1.8 \times 10^{-9} )</td>
</tr>
</tbody>
</table>
This model works for the condition of a single moving object. However, for natural content conditions, it remains an open issue that how to combine the visual discomfort induced by several different moving objects and how to integrate the visual discomfort scores of different scenes are still under study.

5. Conclusions

In this study, we used paired comparison method to evaluate visual discomfort induced by motion of stereoscopic 3D videos. Different motion types, i.e., static condition, planar motion and in-depth motion conditions are compared. The results showed that motion does not always induce more visual discomfort when compared with static stereoscopic images, in particular, in the condition that the object is far from the background, static objects will introduce more visual discomfort than the moving conditions. Generally, in-depth motion stimuli would generate more visual discomfort than the planar motion stimuli, which might be opposite in the condition of very small disparity offsets.

The velocity and the relative disparity are main factors of visual discomfort in stereoscopic 3D videos. The disparity amplitude did not impact visual discomfort significantly which is consistent with the conclusion of [11]. In addition, it is found that the frequency of the oscillation of the in-depth motion is not a significant factor of visual discomfort in 3DTV.

The time course effect on visual discomfort measured by paired comparison method was studied in this paper. The results indicate that observer’s discriminability on visual discomfort did not change significantly with the growth of the time. Compared with the scale-based subjective assessment method, e.g., ACR, the paired comparison method is more reliable in assessing visual discomfort.

According to the regression analysis, an objective model which can be used to compare the visual discomfort of different types of motion was proposed. In the future work, this model will be extended and evaluated by the natural content video sequences.

Acknowledgment

The participants of the subjective experiment are gratefully acknowledged. This work has been partly conducted within the scope of the Just Explore Dimension (JEDI) ITEA2 project which is supported by the French industry ministry through DGCIS and the PERSEE project which is financed by ANR (Project Reference: ANR-09-BLAN-0170).

References
