Real time comfort enhancement in stereoscopic displays by disparity and content-adapted blur

Laure Leroy\(^{(1)}\), Philippe Fuchs\(^{(1)}\), Guillaume Moreau\(^{(2)}\)

\(^{(1)}\)Robotics Center – Mines Paritech, \(^{(2)}\)Ecole Centrale de Nantes - CERMA
laure.leroy@ensmp.fr, philippe.fuchs@ensmp.fr, guillaume.moreau@ec-nantes.fr

Abstract - Stereoscopic devices become widely used (immersion-based working environments, stereoscopically viewed movies, auto-stereoscopic screens,...). In some instances, exposure to stereoscopic immersion techniques can be lengthy, and eyestrain sets in. We propose a method for reducing eyestrain induced by stereoscopic vision. After reviewing sources of eyestrain linked to stereoscopic vision, we will focus on one of these sources: images with high frequencies contents associated with large disparities. We will put forward an algorithm for removing the irritating high frequencies in high disparity zones (i.e. for virtual objects appearing far from the real screen level). We will elaborate on our testing protocol to establish that our processing reduces eyestrain caused by stereoscopic vision, both objectively and subjectively. We will subsequently quantify the positive effects of our algorithm on the relief of eyestrain.

I. INTRODUCTION

Stereoscopic immersion has a reputation for producing visual fatigue in long exposures. As many visual immersion devices make use of stereoscopy, it appears important to study the so-called visual fatigue in stereoscopic environments. Visual fatigue is a generic expression which comprises a range of reversible symptoms, usually in the form of ocular and visual problems, as well as more commonly found, less specific ones, such as stinging or heavy eyes, misty or double-outlined vision, persistent headaches, dizziness, etc…

In order to make stereo screening possible we endeavor to reproduce binocular vision and we make use of its basic principles, but in the end a number of differences between the real world and the replicated one cannot be eliminated, and these differences are known to generate eyestrain. Sometimes, people have to work longtime in such immersive rooms, we will try to minimize their eye strain

In this article we will, more to the point, study visual fatigue caused by high spatial frequencies in the presence of large disparities occurring in stereo screening. We will propose algorithms for reducing this visual fatigue as well as the results of our experiments, in order to demonstrate the effectiveness of our algorithms.

We will begin by setting out the sources of visual fatigue in relation to high frequencies. We will subsequently put forward the algorithm of our choice for reducing this fatigue and we will end by presenting fatigue test results to verify the effectiveness of our processing techniques and an analysis of our findings.

II. VISUAL FATIGUE AND HIGH SPATIAL FREQUENCIES LINKED TO LARGE DISPARITIES

Stereoscopic vision consists of directing two slightly different images to each eye.

![Fig. 1. Panum’s Area](image)

When the horizontal disparity \(\alpha-\beta\), in other words the difference between the two images (also called horizontal parallaxes), is too significant, it can no longer be fused by the brain and is perceived as double-outlined (diplopia). The Panum’s area is known as the zones around the null-disparity zone (\(\alpha-\beta=0\)) where stereo images can be fused (Figure 1).

In the following sections, we will show that fusion limits are variable with respect to the frequency content of the images but that high frequencies also have an influence on visual comfort.

A. Fusion limit

Several studies have established that the fusion threshold is higher when we look at a grid of low spatial frequencies [1]. Schor used a vertical luminance gradient based upon a difference of Gaussians. The fusion threshold proved then to get lower as the spatial frequency was being increased, such as shown in Figure 2. The fusion limit in the vertical axis was clearly lower than the one in the horizontal axis [2]. When the spatial frequency was less than 1.5 cpd (cycle per degree), the
fusion limit practically corresponded to a 90° phase shift (limit beyond which the subject no longer fuses the matching strips, also known as the Rayleigh limit). It was from this inferred that at this frequency the stereoscopic acuity was only limited by the monoscopic acuity. With respect to spatial frequencies higher than 2.4 cpd, the fusion limit ranges from 5 to 10 min arc. As a consequence, the Rayleigh limit no longer corresponds to the stereoscopic fusion limit with respect to higher frequencies. The fusion limit for high frequencies lies between 6 and 8 times the width of the DOG’s center.

These experiments were repeated with a vertical grid no longer following a DOG pattern, but with rectangles with sharp edges whose width and luminance equaled those of the corresponding DOG’s. It can be seen from Figure 2 that the fusion limit for those sharp lines clearly resemble that of the narrowest DOG’s. This implies that the people subject to the test fundamentally based their perception on the outline of the objects, and thus on higher frequencies [3].

These experiments also clearly indicated that contrast and luminance play no role in fusion, which was confirmed by Heckmann et al [4]. This allow to assume that removing high frequencies will allow to increase the Panum’s area. We can, for each object not perceived in the Panum’s area, remove the high frequencies until the Panum’s area includes the object or the problematic point.

Fig. 2. Relation between spatial frequencies and Panum’s area radius [2].

A. Visual comfort and high frequencies

Wöpking carried out a study on visual comfort in relation with the spatial frequencies occurring in a pair of stereoscopic images. Twelve people were presented a pair of stereoscopic images with calibrated spatial frequencies and disparities.

They were asked to try to quantify their level of discomfort, ranging from -2 (very irritating) to +2 (imperceptible) [5].

Perrin [5] designed a comfort function which was basically an interpolation from Wöpking’s data. It establishes a relationship between comfort (“C(d,f)”), horizontal disparities (“d”) and spatial frequencies (“f”). As was the case with Wöpking, the level of discomfort is expressed by a value ranging from -2 (very irritating) to +2 (imperceptible). The comfort function based upon the interpolation of the Wöpking’s data is shown hereunder (Figure 3)[6].

\[
C(d, f) = a(d - d_0 - kf^2)
\]  

Fig. 3. Perrin’s comfort function [6]

III. PROPOSITION

Since high frequencies, where large disparities occur, are a strain on the vision, we propose to define an algorithm which will locate the higher frequencies prior to removing them. This reduction will be related to local disparities, which means that it will be more significant in the large disparity areas.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Wavelet</th>
<th>Box filter</th>
<th>GPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frames per second</td>
<td>0,5</td>
<td>4,6</td>
<td>97</td>
</tr>
<tr>
<td>real time</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>decomposition</td>
<td>complete image</td>
<td>complete image</td>
<td>by object</td>
</tr>
<tr>
<td>recomposition</td>
<td>complete image</td>
<td>zone</td>
<td>by object</td>
</tr>
<tr>
<td>blur choice</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>direct link with frequencies</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

Fig. 4. Recapitulative table.

Previous studies have been conducted by means of a wavelet transform [7]. Even though the link with high frequencies is more intuitive, such a process is however too slow to be carried out in real time [6]. Therefore, we have chosen to work on other techniques, such as the one provided by the so-called BoxFilter or other blurring methods [8][9]. However, our preference went to the blur method, based upon the computing of a sliding average on a computer graphics card (GPU), as it is considerably
faster. A brief comparison of the various algorithms is provided in Table 4. The GPU method is faster despite its higher complexity because it has the advantage of being completely parallel (pixel-based) whereas the Box filter algorithm is based on the notion of integral image and is therefore totally sequential.

So, in order to make a real time processing possible, we have retained as a solution the blur algorithm obtained through computation on a graphics card of a sliding average. From equation (1), we deduce the upper limit frequency for a given comfort level:

\[ f = \frac{1}{14} \sqrt{\frac{4}{k}} - \frac{k}{4} - C(f, d) \]  

(2)

Let us assume that the comfort level must be better than 0. In order to compute a satisfactory frequency by means of equation (1), we want the d-disparity or the horizontal parallax. The latter is obtained by correlating the location of the subject and that of the virtual point. Through the relationships for similar triangles, we can write:

\[ HP = \frac{OD \times IPD}{OD - ED} \]  

(3)

Now we know the acceptable upper limit frequency at a given point. However, spatial frequencies are irrelevant to graphics card soft/hardware’s, which only take pixels into account. We will therefore have to translate our spatial frequency (in cycle per degree of visual angle) into pixels (Figure 5).

We take the visual half-angle as a basis and we apply trigonometry rules to a right-angled triangle:

\[ \tan(\frac{1}{2}) = \frac{SD}{ED} \]

Therefore

\[ SD = 2 \times \tan(\frac{1}{2}) \times ED \]

(4)

And as our pixels are of width \( N \), we can say that the number of pixels per visual angle equals:

\[ NPPDV = \frac{2 \times \tan(0.5) \times ED}{N} \]  

(5)

The averages will consequently be computed on that number or pixels, divided by the spatial frequency considered for a given disparity. The generated blur is quite progressive, and is a function of disparity.

IV. EFFECTIVENESS OF THE PROCESS

To assess the effectiveness of our process, we have to measure visual fatigue. This will be done by experimenting on a set of active observers who will have to perform a task in a virtual work. It is hereafter described. The task is repeated twice: one day with our image processing, one day without, in random order.

A. Protocol

Our virtual world is made up of 5 vertical cylinders (10 cm in diameter and 20 cm high) located 80 cm from the ground and physically distributed on both sides of the real screen. The environment also contains a 5 cm sphere located at the end of a 60 cm long tube ‘connected to’ the WiiMote that more or less represents a hammer. A red square (10 cm) is used to designate the target of the task. The remainder of the virtual environment is composed of a screen background and 3 prisms. All these items (except for the red square) are covered with textures rich in high frequencies (see Figure 6).

Subjects have to point a randomly designated (by the red square) cylinder with a "hammer" very precisely 300 times.

We measure task effectiveness (positioning error on the cylinders), ease of accommodation before and after (speed of accommodation after moving on to a new target), amplitude of accommodation (this is the minimum focus distance) and stereoscopic acuity (this is the smallest discriminatory perception on the subject’s retina) [10].

All the 20 subjects are between 18 and 40 year old. According to [10], older people could have disturbed the ease of accommodation measure.

B. Physical device

During vision tests (accommodative or stereoscopic acuity), the targets were precisely at a 40 cm distance from the subject’s eyes. A music stand mechanically interdependent of a car seat was used to ensure compliance with this prerequisite (left of Figure 7).

The screening surface is provided for by a 3.10 m by 1.74 m LUMIN screen. Our projector is a Christie Mirage 3. The screen and the projector’s optics are 3.5 m apart. Our pixels are 1.61 mm wide. The frame per second frequency is 60. The tracking system is a millimeter-accurate ART2 device also operating at a 60 Hz frequency (right of Figure 7).
C. Results

1. Amplitude of accommodation

The measurement method (Donder's Push-up Test) was chosen for its aptness to detect minute variations in visual fatigue [11]. Amplitude of accommodation is measured with a target approaching the eyes of the subject. He tells us when he can reach to read.

Figure 8 shows the differences in amplitudes of accommodation measured before and after each task. It can be seen that the amplitude of accommodation is hardly affected when spatial frequencies are gradually removed by our processing. Without this processing, a loss of 1.21 cm is experienced by the subjects. Their eyes are thus more readily tired when our image processing does not take place. The probability associated with the significance of average differences amounts to 99.6% which excludes the possibility of a mere statistical aberration.

2. Ease of accommodation

Ease of accommodation is measured by means of a test known as the “Flipper Lens Test”. The purpose of this test is to determine the shortest period of time necessary for our eyes to repeatedly adjust to a new stimulus [10]. It is performed with two pair of lens (2 and -2 of dioptre). We change (half-cycle) the lens when the subject reaches to read the target.

Figure 9 shows the differences in comparative deviations in the numbers of half-cycles before and after the task. As can be seen, the average difference is 1.77 cycles per minute. This is a significant difference, especially since the significance probability associated thereto is 99.45%.

3. Stereoscopic acuity

From Figure 10, we can see that the difference before and after the test in stereoscopic acuity is quite significant. The stereoscopic acuity appears to be clearly further reduced without our processing and, conversely, it does not seem to be affected when our algorithm alters the image.

It should however be noted that this distribution is - regretfully - not normal. The difference between two standard deviations is significant and we have less than 30 subjects. It should nonetheless be further noted that whenever we compute the significance by means of these tests, a risk probability less than 0.00001 is observed.

4. Effectiveness when the task is performed

Figure 11 shows a chart in which the blue bars indicate the average of errors made without processing (expressed in meters) while the red ones represent the averages of errors made with a processing applied, with respect to each of the cylinders. The error bars represent standard deviations. No deviation from the average is however significant, but the differences in standard deviations increase significantly when no processing takes place, except for cylinder 3 which is tied to the screen and for which images are never corrected.
5. Subjective measures

It can be inferred from Figure 12 that the people subject to the test indicate a preference for a world which has not been cosmetically altered. This does not really come as a surprise since we’ve intentionally blurred it locally.

It should be further noted that almost one third of the people interviewed did not experience any difference in visual perception from one day to another. Figure 13 shows that twice as many subjects consider that a world which we did not process is more tiring than the one we did.

This would tend to lend a lot of credibility to the results of the visual fatigue tests we carried out in the previous paragraphs.

On the other hand, the processed images are perceived as less attractive and appeal less to the subjects, even though there seems to be a consensus on them being less vision-straining. This might prove to be an issue whenever aesthetic detail plays a major role. This should obviously be a point of consideration when dealing with works of art, public exhibitions, and whenever focusing on detail is an issue, for example, when inspecting the finish of a manufactured product, etc.

We have seen no degradation in the task effectiveness. Its precision error remained unaltered. On the other hand, its standard deviation is increased. This could mean that on average the number of errors made by the subjects did not rise, but that compared with an unprocessed virtual world, some of the subjects make more errors and some other less. Furthermore, many subjects claim that processing the image to suppress visual fatigue relieves the strain off the task.

In future work, we will try to determine the biases of our image processing on shape perception and to see whether our blurring method should be applied on an image basis or on an object basis.
VI. REFERENCES


