Universität des Saarlandes
Max-Planck-Institut für Informatik

Thesis Tytple

Masterarbeit im Fach Informatik
Master's Thesis in Computer Science
von / by

Author Name

angefertigt unter der Leitung von / supervised by

betreut von / advised by

begutachtet von / reviewers

Saarbrücken, August 2015

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Statement in Lieu of an Oath

I hereby confirm that I have written this thesis on my own and that I have not used any other media or materials than the ones referred to in this thesis.

Einverständniserklärung

Ich bin damit einverstanden, dass meine (bestandene) Arbeit in beiden Versionen in die Bibliothek der Informatik aufgenommen und damit veröffentlicht wird.

Declaration of Consent

I agree to make both versions of my thesis (with a passing grade) accessible to the public by having them added to the library of the Computer Science Department.

Saarbrücken, August 2015 — Mohamed Gad-Elrab

Abstract

Stereoscopic and automultiscopic displays suffer from crosstalk. An effect which greatly reduces image quality, viewer comfort and distort the perception of depth. Previously, only a limited work has been done on understanding the relation between crosstalk and the perceived depth with respect to the nature of the stimuli. Moreover most of the previous work is carried on simple monochromatic scenes. Since the human visual system uses numerous other cues than disparity to estimate the depth of an object in a stereo scene, monochromatic scenes are poor choice for understanding the above mentioned relation. Moreover, the model for depth resolution via disparity as provided by the current literature fails to justify why and how the perceived depth is affected by the crosstalk. In this work, we improved and performed more generalized experimentation to see how the depth perception is affected by the crosstalk for different kinds of stimuli. Based on the result of these experiments, we derived a model for human visual system's resolution of depth from disparity that accurately measures the depth of a stimulus as perceived by the human in presence of cross-talk. Finally some improved algorithms for removal/compensation of crosstalk in automultiscopic are developed.

Acknowledgements

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Preliminaries

Discuss General idea behind crosstalk and why is it so bad. How it affects the depth and what current literature think about it.

1.2 Contributions of this Thesis

Experimentation, mitigation, HVS model.

test

1.3 Structure

1.4 List of Commonly used abbreviations

Chapter 2

Relevant Background

2.1 Depth Perception

Depth perception is the ability of the Human Visual System to visualize the three dimensional world as well as measuring the distance of an object based on two dimensional images obtained from the eyes. Depth perception is imperative for performing basic everyday tasks such as avoiding obstacles without bumping into them or interacting with the world with relative ease. In animals (specially predators), it is critical to estimate the distance of a prey for an efficient attack. Depth sensation is the term used for animals as it is not known whether they sense the depth in the same way as humans do or not[5].

Human visual system uses several monocular and binocular cues to determine the depth of objects in the view. These cues can be categorized into two categories i.e. cues extracted from a single image (Monocular Cues) and cues extracted from two images (Binocular cues)[4][5]. Figure 2.1 gives an outlook of the depth cues used by the HVS. These cues are then dynamically weighted according to their robustness by the HVS in order to estimate a depth value for each object in the view [3](Write details of those cues in Appendix).

2.2 Binocular Vision and Stereopsis

Generally speaking, all the animals with two eyes have binocular vision and they can integrate the information from two eyes based on the binocular overlap. But the term Binocular vision is usually used for the animals that have a large area of binocular overlap (Human and most other predators) and use it to get the depth information of the world around them. In addition to calculating depth, binocular vision also has advantages in

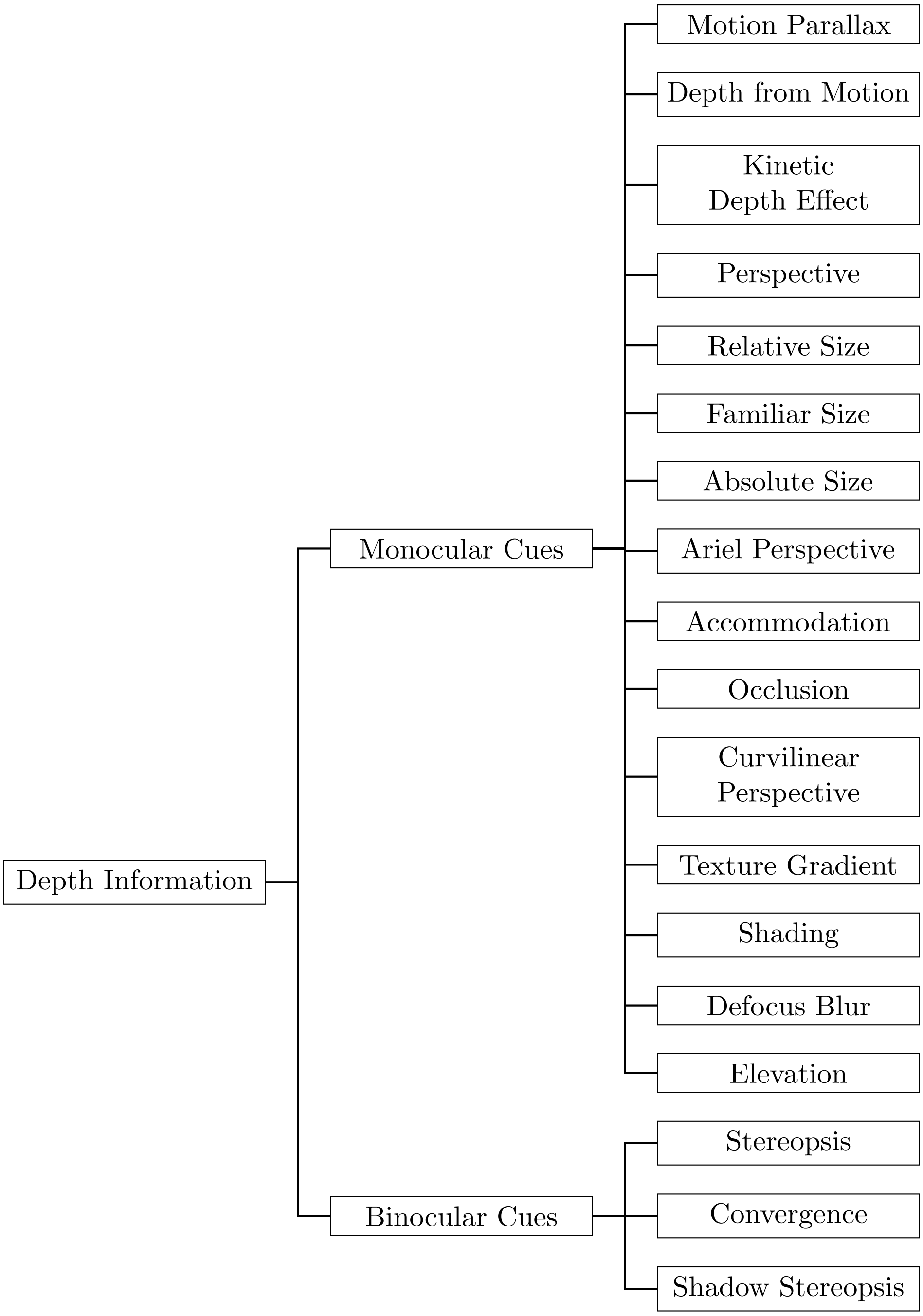


FIGURE 2.1: HVS Depth Cues

performing other tasks such as detection, discrimination, detecting camouflaged objects or eye-hand coordination. Even the resolution observed world is increased with binocular vision[2]. Among all the depth cues discussed in the section above, Stereopsis is the most influential of them all. Since the human eyes are located at different lateral positions on the head, the images formed on the retinas of these two eyes are slightly different. The difference is mainly the horizontal positions of the objects[7]. The process of obtaining a fused (Binocular fusion) image (Cyclopean image) and obtaining a depth map based on the horizontal disparities of the objects in these two images is known as stereopsis.

When the eyes verge in order to focus on some object (or point) in space, that object is projected at identical corresponding points in the retinas. It means that the difference between their horizontal positions is zero. This process is called fixation of the eyes and the distance of the object (point) at which the eyes are fixated is called the fixation distance. The locus of all the points in space that is projected on identical retinal points is called the horoptor[6]. Theoretically, via geometrical principles, the horoptor is a circular segment in the plane of fixation point. However, Wheatstone in 1938 observed that the actual/emperical horopter is much larger than that. Figure 2.2 shows both the theoretical and empirical horoptor. Any object that is farther away from the horoptor has uncrossed disparity in the retinas i.e. the eyes need to be diverged (uncrossed) in order to fixate on that object. Similarly, any object that is closer than the horoptor has crossed disparity in the retinas i.e. the eyes need to be converged (further crossed) in order to fixate at it cite and make the diagram for crossed uncrossed disparities. Figure ?? shows a graphical representation of crossed and uncrossed disparities add the dolphin images from wikipedia article..

Stereopsis is believed to be processed in the binocular neurons of the visual cortex of mammals. The binocular neurons have receptive fields in different horizontal positions in each eye. These cells are active only when the object of interest is in certain range of disparity in one eye relative to the other i.e. there is a maximum disparity limit. add the dolphin images from wikipedia article.. As the objects in the images formed at the retinas of both eyes are slightly shifted horizontally, presenting two different images with shifted object two both eyes can fool the HVS into perceiving depth. This process is called stereoscopy. The first stereoscope was invented by Sir Charles Wheatstone in 1838[8]. It used two mirror both tilted at 45 degrees to the eyes that reflected two different images from the sides. Currently all the stereoscopic screen present a different perspective image two both eyes with different technologies that will be discussed in later sections.

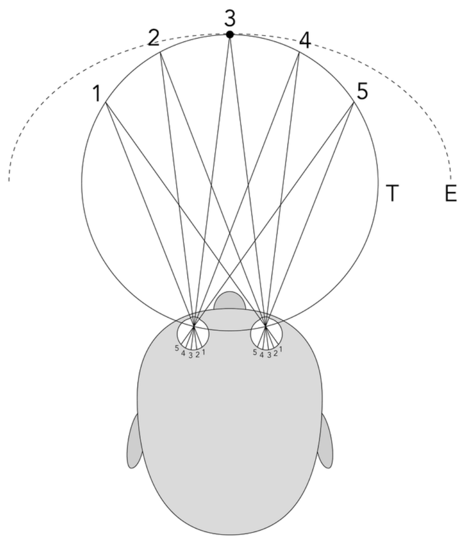


FIGURE 2.2: Representation of theoretical (T) and empirical (E) horoptor

2.3 Limits of Stereopsis

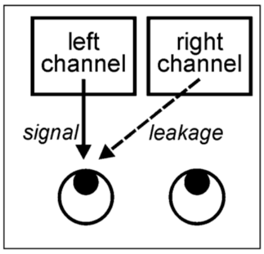
2.4 Crosstalk

As discussed in the previous section, stereoscopy includes the process of displaying different perspective images to each eye in order to mimic the effect of depth in a scene. However, it is critical that the perspective of these two images should be segregated completely from each other. Currently, all the commercially available 3D displays (with exception of head mounted displays or Wheatstone setups) fail to isolate the two images completely. That means that a percentage of the image of one eye leaks to the other eye as well. This unintended leakage is called the crosstalk in stereoscopic screens[10]. Figure 2.3 Shows one such example where the screen has a simulated crosstalk of 14%. This means that 14% image intensity of the right eye image is leaked into the left eye image and vice versa.

"System crosstalk" is the term used to define the amount of light leakage that occurs between two views and is independent of the image contents. In simplest form, it can be mathematically defined as:

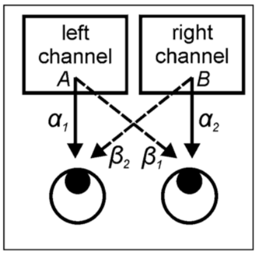
$$Crosstalk(\%) = \frac{leakage}{signal} * 100 \tag{2.1}$$

(A) Original Image



(B) Original Image

(C) Original Image



(D) Viewers Crosstalk

FIGURE 2.3: (a-c) A simulation of 14% crosstalk. An Image intended for the left eye containing 14% of image intended for right eye. (d) Illustration of Viewers crosstalk with the transfer functions as defined by Andrew J. Woods [10]

Here, "signal" is the luminance of the original image intended for an eye and "leakage" is the luminance of the light that leaks from the unintended image. Usually, on a stereoscopic scree, the amount of crosstalk is measured by observing the luminance of the left channel while a black (minimum luminance) image is displayed in the left channel and a white (maximum luminance) image is displayed on the right channel and vice versa. This definition is accurate for the screens that can manage to display a true black i.e. zero luminance. Almost all of the LCDs can not produce zero luminance as their minimum luminance hence resulting in a non-zero black level at its minimum. Another mathematical representation of crosstalk take into effect this non-zero black level and is defined as:

$$Crosstalk(\%) = \frac{leakage - blacklevel}{signal - blacklevel} * 100 \tag{2.2}$$

Throughout the rest of the thesis, we will be using crosstalk as mentioned in eq 2.2.

Viewers crosstalk on the other hand is the amount of crosstalk that can be perceived by the viewer as ghosts. It is dependent on the image contents i.e the contrast at the ghosting point and the parallax of objects in the scene. If the system crosstalk is defined as

$$System\ Crosstalk\ (left\ eye) = \frac{\beta_2}{\alpha_1} \tag{2.3}$$

Where α_1 denotes the percentage part of the left-eye image at position (x,y) as observed by the left eye and β_2 denotes the percentage amount of right eye image at the same location (x,y) leaked into the left eye. Then the viewers crosstalk is defined as

$$Viewers\ Crosstalk\ (left\ eye) = \frac{B\beta_2}{A\alpha_1}. \tag{2.4}$$

where "A" is the luminance of that particular point in left-eye image and "B" is the luminance of that particular point in right-eye image (as described in Fig 2.3d). The Variables α and β are characterizing the transfer functions from the displayed image to the observed image i.e. the amount of light reaching the eyes after being displayed on the screen and going through the glasses or any other medium that resides between the screen and the eyes. In most displays, crosstalk is an additive process and is roughly linear. This means that to simulate crosstalk, adding a desired amount of unintended image to an intended image should be sufficient. The simulation of crosstalk in our experiments was carried out in the same manner.

The perception of crosstalk obeys the Webber's law which means that leaked light will be greatly perceivable by the viewer on the dark image areas rather than bright areas. Also, the percieved crosstalk is Dependant upon the contrast of the image and the binocular parallax of the stimuli i.e. crosstalk perception will increase with increase in contrast or increase in binocular parallaxMaybe I need to show why?. It is commonly believed that ghosting plays the most critical role in determining the image quality. Wilcox and Stewart [9] observed that over 75% of the observers in their experiments reported the ghosting to be the key feature that deteriorated the image quality. Apart from reduction of image quality, perceivable ghosting is also responsible for loss of depth, viewer's discomfort, reduction of sharpness and contrast, decreased fusion limits, and difficulty in fusion. As mentioned earlier, two objects can not be fused together if the angular separation between them is smaller than the disparity[1]. In fact since the ghost and the stimulus lie in the same (x, y) position but at different depth, the angular separation becomes zero and the disparity gradient of the ghost and the stimulus becomes infinity. Hence both of them can not be fused together. In this case the more luminant object i.e. the stimulus will be fused while the ghosts will remain in diplopic stateadd the image showing infinity DG and cite the source. Crosstalk in still images is perceived to a larger extant than crosstalk in dynamic scenes (e.g. a movie). This means that motion of objects in a scene can mask the perception of crosstalk.

The stereoscopic literature provides a lot of advices for the acceptable and unacceptable crosstalk for the viewers. Woods[10] points some of these advices as:

- Crosstalk between 2% to 6% significantly affects the visual quality and increase the viewers discomfort.
- In order to produce accurate depth range between 40 arcmin, the crosstalk should be as low as 0.3%.
- Crosstalk is visible even at 1% to 2%.
- 5% of crosstalk is enough to induce discomfort.
- JND for crosstalk is 1%.
- 2-4% of crosstalk can significantly decrease the amount of perceived depth.

As one can observe, there is a variability in these guidelines. The reason for this variability might be because of the different setups and types of stimuli that were used by the researcher in their experiments. We observed that currently the literature is not quite thorough on how the depth perception is affected when it comes to different kinds of stimuli and relatively more complex scenes. Which is why we performed some experiments of our own to verify and expand the current knowledge in this area. This is also one of the main contributions of this thesis.

2.5 Stereoscopic/Automultiscopic Screens and its cross-talk

2.5.1 CRT Screens

2.5.2 LCD Screens

2.5.3 Anaglyph Stereo

2.5.4 Active/Time Sequential Stereo

2.5.5 Passive/ Space Multiplexed Stereo

2.5.6 Automultiscopic Screens

2.6 Crosstalk Quality Metrics

2.7 Lightfields

Chapter 3

Related Work

3.1 Effects of Crosstalk on Perceived Depth

3.1.1 Thirsins's work

3.1.2 Systematic Distortion

3.1.3 Visibility Threshold and Fusion Limit

3.2 Migigation/Compensation of Cross-talk

3.2.1 Stereoscopic Screens

3.2.1.1 Subtractive Approaches

3.2.1.2 Perceptual Optimization

3.2.1.3 Temporal Approach

3.2.2 Automultiscopic Screens

3.2.2.1 Inverse Filtering

3.2.2.2 Subtractive Reduction

3.2.2.3 Sub-pixel Optimization

3.2.2.4 Low Pass Filtering

Chapter 4

Contribution

4.1 Crosstalk Experiments

4.1.1 Apparatus setup

4.1.2 Experimentation Procedure

4.1.3 Stereoscopic Experimentations

4.1.3.1 Initial Hypothesis

4.1.3.2 stimuli

4.1.3.3 Simulation of Cross-talk

4.1.3.4 Experimentation procedure

4.1.3.5 Results

4.1.4 Automultiscopic Experimentations

4.1.4.1 Initial Hypothesis

4.1.4.2 stimuli

4.1.4.3 Simulation of Cross-talk

4.1.4.4 Experimentation procedure

4.1.4.5 Results

4.1.5 Conclusion and Discussion

Chapter 5

Applications

5.1 Depth Adjustment for Depth Critical viewing Applications

5.2 Efficient Preprocessing for Crosstalkfree Images

Chapter 6

Conclusion

6.1 Summary

6.2 Future Work

6.3 Open Questions

Bibliography

[1] BURT, P., AND JULESZ, B. A disparity gradient limit for binocular fusion. *Science 208*, 4444 (1980), 615–617.

[2] HOWARD, I. *Binocular vision and stereopsis.* Oxford University Press, New York, 1995.

[3] LANDY, M. S., MALONEY, L. T., JOHNSTON, E. B., AND YOUNG, M. Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision research 35*, 3 (1995), 389–412.

[4] REICHELT, S., HÄUSSLER, R., FÜTTERER, G., AND LEISTER, N. Depth cues in human visual perception and their realization in 3d displays. In *SPIE Defense, Security, and Sensing* (2010), International Society for Optics and Photonics, pp. 76900B–76900B.

[5] WIKIPEDIA. Depth perception — wikipedia, the free encyclopedia, 2015. [Online; accessed 20-August-2015].

[6] WIKIPEDIA. Horopter — wikipedia, the free encyclopedia, 2015. [Online; accessed 21-August-2015].

[7] WIKIPEDIA. Stereopsis — wikipedia, the free encyclopedia. [Online; accessed 21-August-2015].

[8] WIKIPEDIA. Stereoscope — wikipedia, the free encyclopedia, 2015. [Online; accessed 23-August-2015].

[9] WILCOX, L. M., AND STEWART, J. A. Determinants of perceived image quality: ghosting vs. brightness. In *Electronic Imaging 2003* (2003), International Society for Optics and Photonics, pp. 263–268.

[10] WOODS, A. J. Crosstalk in stereoscopic displays: a review. *Journal of Electronic Imaging 21*, 4 (2012), 040902–040902.

Index