Half-Occluded Regions and Detection of Pseudoscopy

Jonathan Bouchard[1,2], Yasin Nazzar[1], and James J. Clark[1]
[1]Centre for Intelligent Machines, McGill University
Montreal, Canada
[2]Mokko Studio, Montreal Canada
jonathan@cim.mcgill.ca

Abstract

We propose that left- and right-half-occlusion regions contain information that can distinguish between stereoscopic and pseudoscopic display conditions. A machine vision method is presented based on this idea, which detects pseudoscopy using only the histograms of left and right half-occlusion pixel locations.

Two psychophysical experiments are described which study the ability of human viewers to detect, or be influenced by, pseudoscopic display. The results of this study show that, during free viewing of HD 3D video imagery, humans judged pseudoscopic imagery to be of lower quality than stereoscopic imagery. Subjects performed at a 72% rate in deciding whether a short 5-second video clip was presented stereoscopically or pseudoscopically. Subjects were observed to fixate on half-occlusion regions with a frequency of 15.8%, as opposed to a frequency of 7.8% indicated by random chance. Viewers more frequently (17.4%) fixated on half-occlusion regions when making correct decisions then when they were incorrect (11.9%).

1. Introduction - Stereoscopy and Pseudoscopy

Three-D stereoscopic movies are currently a staple of the commercial cinema market. They provide viewers with a sense of depth by displaying separate images to their left and right eyes, mimicking the binocular parallax that would be seen by their eyes in viewing a real scene in the world. If the illusion of depth is to be produced correctly, the image captured by the left-hand film camera must be presented to the viewer's left eye and the image captured by the right-hand camera must be presented to the viewer's right eye. If the images are reversed, or *pseudoscopically*, so that the left eye sees what the right eye should see and vice-versa, a peculiar perception is engendered where objects or people that should be in the foreground now appear as though seen through holes in the background. Clearly, when playing a 3D movie, care should be taken so that the stereo images are presented to the correct eyes.

The detection of a pseudoscopic display condition is deceptively difficult. Viewers may feel that something is amiss, but often do not sense any problem with their depth perception. A pioneer in 3D movies, Chris Condon (in [19], chapter 2, page 14) recounts a prime example of this in the early days of 3D cinema:

> ... if the projectionist is not careful, or he misthreads or missplices the film, the whole darn thing will go in reverse. I had seen that happen with *Friday the 13th Part III* in a theater in North Carolina. I walked in in the middle of the picture, and here it is in reverse. It's running pseudoscopically, and 150 people in the audience were just sitting there watching the film with the 3-D turned inside out. So I went up to the projection booth and I tried to tell the guy. And he says, "No, you're not allowed to be in here." I said, "Look. My name is Chris Condon. I'm very experienced in this field. You've got it misframed. It's in reverse. He said, "No, it looks fine to me."

Anecdotal evidence aside, it is an interesting question as to whether viewers of 3D films are, in fact, able to detect pseudoscopy. If they are, then it is of scientific interest to understand how they do it, and to identify conditions in which pseudoscopy will be fatiguing to the viewer or prevent easy assimilation of displayed information. It is also important to develop machine vision techniques for detecting pseudoscopic display conditions based on models of human perception so that problematic viewing conditions can be detected and corrected automatically. In this paper we describe a psychophysical study on human detection of pseudoscopy, and also present an automated method for pseudoscopy detection. We suggest that viewers make use of information related to *half-occluded regions* to decide whether a movie is being presented pseudoscopically, and also present a computer vision technique for detecting pseudoscopy based on analyzing half-occluded regions.

2. Half-Occluded Regions in Stereoscopy and Pseudoscopy

To help understand the reasoning behind our proposal, we will first review the concept of *half-occlusion*. In vision, an *occlusion* occurs when an opaque object lies in the line of sight between the viewer and another object. The object in front *occludes* the object behind, preventing it from being seen.

The presence of an occlusion depends on the line of sight. Thus, an object that is occluded from one viewpoint may be visible, or unoccluded, from another viewpoint. This can happen even when two viewpoints are very similar, such as in the case of the two eyes in our heads. Because each eye has a slightly different viewpoint from the other, an object could be visible in one eye, but occluded in the other. This situation is known as *half-occlusion*, since the view of only one of the eyes is occluded.

Half-occluded regions are problematic for stereo vision, since these regions generally do not have any match or corresponding points in the other view, making measurement of binocular disparity impossible. However, half occlusions can be useful, as they provide powerful cues to other information such as object boundaries, depth discontinuities, etc. In particular, they provide information which can be used to help decide whether a given stereo image pair is presented stereoscopically or pseudoscopically.

Figure 1 shows the location of half occluded regions in a simple scene consisting of a flat object in front of a backplane. In a stereoscopic presentation the left eye sees the left-half-occlusion region (which is the region which the left eye can see but the right cannot) at the left edge of the front object. The right eye sees the right-half-occlusion region (which is the region which the right eye can see but the left eye cannot) at the right edge of the front object.

In the case shown in figure 1 of an object in front of the background, it can be seen that, in the case of stereoscopic display, the right-half-occlusions are to the right side of the foreground object, while the leftt-half-occlusions are to the left side. This pattern is reversed in the pseudoscopic display. If we knew we were looking at this type of a simple scene then we could determine whether the display is stereoscopic or pseudoscopic just by taking the difference of the x-coordinate of the right-half-occlusions and the left-half-occlusions. If this difference is positive then the display is stereoscopic. If it is negative then the display is pseudoscopic.

However, one could have a scene such as that shown in figure 2, in which the background plane is being viewed through a small hole in the foreground plane. In this case the locations of the left and right half occlusions are reversed from what they were in the case of a the foreground object occluding the background. Thus it would seem that a simple

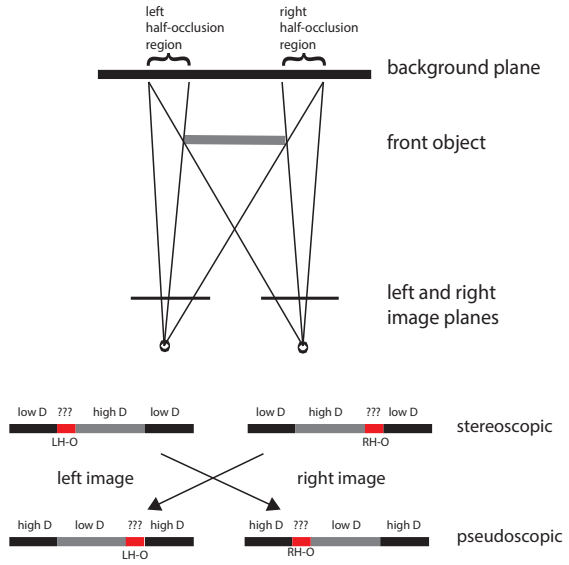

Figure 1. The left and right half-occluded regions in a simple scene of a small planar object in front of a planar background. Shown at the bottom are depictions of the left and right images in stereoscopic (top) and pseudoscopic (bottom) display. The black regions correspond to images of the background and the grey regions to the images of the foreground object. The red regions correspond to the half-occlusions. The disparity levels for the image regions are also shown. The disparity of the near object are higher than that of the background. There is no disparity defined in the half-occlusion regions.

differencing of the right and left half occlusion locations would not help in detecting pseudoscopy.

However, in most scenes encountered in practice during the filming of 3D movies, the number of foreground objects occluding background objects is generally much higher than the number of holes in the foreground objects. Thus, the situation in figure 1 dominates over that of figure 2, and we can reliably use the left-right occlusion location differences to determine whether the imagery is presented normally or pseudoscopically. In practice we would compute the centroids of all of the left and right half-occlusion pixels in the imagery and use the difference of the centroids in deciding whether the display is pseudoscopic.

It is interesting to note that, when a scene is viewed pseudoscopically, the occluding foreground objects appear to be background objects viewed through holes the exact shape of these objects. It may be the unnatural preponderance of such "holes" perceived in pseudoscopic imagery that leads to the feeling on the part of the viewer that something is wrong with the imagery.

Even in naturally occuring scenes, there may be roughly

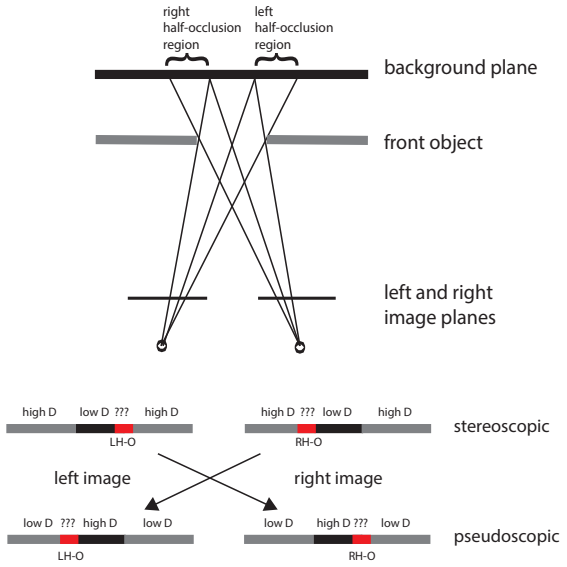

Figure 2. The left and right half-occluded regions in a simple scene of a planar background being viewed through a hole in the foreground plane. Shown at the bottom are depictions of the left and right images in stereoscopic (top) and pseudoscopic (bottom) display. The disparity levels for the image regions are also shown.

equal numbers of "holes" and occluding objects. In this case our right-left occlusion difference method will not work. This type of situation would reveal itself by a relatively low difference value. For such cases we need to use another approach. One such approach suggested by Akimov *et al.* [1] is to observe the appearance (say of the intensity, color, or the texture) of the visible part of the half-occlusion regions as compared with the foreground and background regions. While it is true that we cannot compute a meaningful disparity value for the half-occlusion regions, we can see that the visible half of the half-occluded region generally looks similar to the background region rather than to the foreground region. This is because it is the background that is being occluded and hence not visible in the other image. In the pseudoscopic display this effect is reversed and the visible part of the half-occluded regions appears similar to that of the nearer object (the one with higher disparity). This is the case no matter whether we have the situation of a foreground occluding object or that of a hole. Thus, this method of pseudoscopy detection will work in situations where we have an equal preponderance of occluding objects and holes.

The computational complexity of texture similarity that is needed for the half-occlusion texture comparison method is generally much greater than that of the left-right half-occlusion difference method. As real-time operation is important for detecting pseudoscopy, we will primarily use the latter method.

3. A Computational Technique for Detecting Pseudoscopy

There have not been many computational techniques presented in the literature for the detection of pseudoscopy. Notable methods include Lee *et al.* [11] and Akimov *et al.* [1]. The approach of Lee *et al.* is based on doing a foreground/background decomposition of the images, followed by a disparity comparison. The idea is that the disparity of the foreground objects should be less (more negative) than the disparity of the background objects. If this comparison is reversed then the display is judged to be pseudoscopic. This approach can work well as long as the foreground and background can be effectively separated, and that the background is actually behind the foreground (for example, a tilted background may have some parts that lie in front of foreground objects). The foreground/background segmentation can be computationally expensive, in addition to the required disparity measurement. The method proposed by Akimov *et al.* is based on an analysis of the occluding regions. They generate a weighted probability measure which combines the width of the occlusion region, and the strength of edges that are closest to the points in the occlusion region, weighted by the distance of the occlusion point to the edge segment. The idea is that the closest edges are assumed to be those of the occluding object, and hence the direction to this edge will identify whether the occlusion is a left or right occlusion. This method will work well when the edge assumption is true, but will fail when the background has strong edges, such as when the background is highly textured.

In this section we describe a novel computational method for detecting pseudoscopic image presentation. Our approach is similar to the Akimov *et al.* method in that it analyzes the half occlusion regions, but uses a simpler computation - that of computing centroids of the region points, rather than computing edge maps and the weighted distance measures.

Our technique involves four main steps: (1) Left-Right and Right-Left Disparity Estimation; (2) Half-Occluded Region Detection; (3) Blur-based Masking of Erroneous Half-Occluded Regions; (4) Right-Left Half-Occlusion Region Centroid Difference Measurement. In practice, steps (1) and (2) are combined in an iterative process, where the computed disparity is used to find the occlusions, which are then fed back to improve the disparity computation. The overall combined process is depicted in Figure 3.

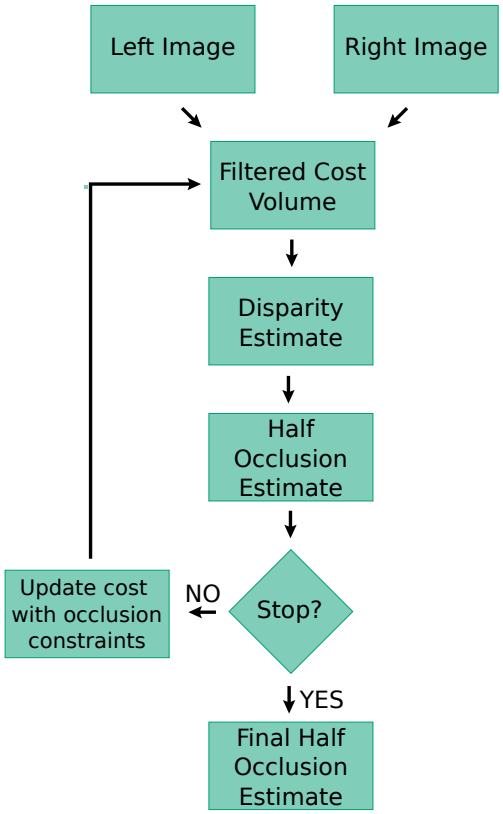

Figure 3. Flowchart outlining the process of half occlusion detection

3.1. Disparity estimation

There are many disparity estimation algorithms to choose from in the computer vision literature. The overall goal of our work is to provide real-time on-set assistance for cinematographers or people working on 2D-3D image conversion. Therefore, along with the accuracy of the depth map we would also like our method to be as computationally efficient as possible. Local disparity estimation methods are generally simpler than global methods, requiring minimal effort in the optimization phase and are generally highly parallelizable. But the quality of their estimates are usually far from that of globally optimized methods. The results tend to be either "blocky" or "unsmooth" based on the window size as cost functions are defined and aggregated over rectangular windows. An aggregation or filtering operation that does not uniformly filter across the rectangular window would be an ideal methodology. This would allow us to retain the simple and parallelizable approach of local methods as well as achieve better looking results. He *et al.* introduced a novel filtering operation called the *guided filter* in [7]. In [15] this filter was used, among other tasks, to compute disparities. Hosni *et al.* [8] further explored utilizing this filter to implement a local stereo correspondence algorithm. The real-time performance of this method makes it suitable for our purposes.

As our pseudoscopy detection algorithm needs both the left and right half-occlusion regions, we need to compute the disparities of both the left and right images.

3.2. Incorporating Occlusion Constraints in disparity estimation

The second stage of our method is the detection of half-occlusion regions. There are many methods available for finding occlusions. Humayun *et al.* [9] provide an effective machine-learning based method which learns to detect points which become occluded from frame-to-frame within a video. Our need, however, is for a method which can find half-occlusions which are distinct from dynamic occlusions/disocclusions. The approach follows the philosophy of Belhumeur and Mumford [2], who showed how half-occlusion information can be integrated with the disparity estimation process.

We obtained the half occlusions from our initial estimation of the disparity map. We can adjust the cost function used in the disparity estimate optimization and obtain an updated disparity map that accounts for the half occlusions. We can iteratively update the disparity estimates and the half occlusion estimates till the change becomes minimal. Even with the iterative approach the half occlusion estimation will not always be perfect due to the presence of noise in the disparity estimate. Therefore some post processing steps are applied to the half occlusion estimates from the iterative process. The first post processing step involves applying a morphological operator to the estimate. Specifically a morphological closing (dilation followed by an erosion) operation is performed to close up holes in the estimates of half occlusion. Following this, a size filter is applied to the output to remove small noisy specks in the estimate. The last step of post processing involves utilizing some information regarding half occlusions as opposed to general image processing techniques. Half occlusions are guided by the location of the camera and the difference in depth between the occluding and occluded surfaces. The width of the half occlusion region is dictated by these factors. We can utilize this information to enforce the width of the half occlusion in terms of pixels in our own estimate and improve the results. The stereo rectification process yields a pair of images that mimic the geometry of images captured using parallel configuration whereas the disparity serves as a representation of the depth. Hence, the width of the half occlusions, in terms of pixels, can be represented by the disparity difference between the occluding surface and the occluded surface. But these occluded surfaces do not have a reliable disparity estimate as they do not have any matching points. The closest pixel that is not a half occlusion is searched for to compute this disparity difference. In the

left half occlusion regions, this search for reliable pixels is conducted to the left of the half occlusion region and to the right of the half occlusion regions in the right image. This is because half occlusions exists on the left side of depth boundaries on the left image and on the right side of the depth boundaries in the right image as has been touched upon previously. Utilizing these final steps we obtain our estimation for half occlusion regions in both the left and right images, while obtaining a refined disparity estimate in the process.

An example of the application of our half-occluded region finding algorithm is shown in Figure 5, which shows the half-occluded regions associated with the stereo image shown in Figure 4.

Figure 4. A single stereo frame from the 3D video test set.

Figure 5. The half occlusion regions detected by our approach. Black regions are the left-half-occlusions and white regions are the right-half-occlusions. The two half-occlusion maps are overlaid onto a single image for ease of display.

3.3. Removing Spurious Half-Occlusions using a Blur Mask

It may be seen in examining multiple frames of a video that some of the half-occluded regions found by our approach are spurious, in that they do not persist from frame-to-frame, or change shape radically over time. Such spurious regions tend to arise in uniform intensity, or low-texture, regions of the image. The disparity computation in these regions gives spurious results, mainly responding to image-noise induced structures. These spurious half-occlusion regions should be filtered out before making use of the half-occlusion information in detecting pseudoscopy.

The standard approach to detecting spurious disparities is to detect, and mask out, image regions of uniform intensity or low texture. We do this, but with a slightly different interpretation of the problem. That is, we look for image regions that are *blurred*, or out-of-focus. Image blur is a specific mechanism by which texture is reduced, and intensity becomes more uniform. We use image blur, or de-focus, as it is intentionally present in many 3D movies. Focus is a means by which cinematographers draw attention of the viewers to different parts of the scene. It is frequently used to add to the storytelling aspect of a movie. By determining the regions that are blurry and the extent of that blur, we essentially have an estimation of focus areas as well. Some methods explored in literature aim to classify an overall image as blurry. Along with the classification an extent of the blur is also stated. Such methods are seen in [17] and [13]. Each paper follows its own approach, [13] taking a probabilistic approach to blur detection whereas [17] analyses the edges in the image using Haar wavelet transform. These works perform relatively well in determining whether the image is blurry or not. However, classifying an entire image as being blurred or sharp along with a measure to report the extent of this blur is not useful for our purpose. We need to determine which regions within the image are out of focus and which regions are in focus. To that end we make use of the technique of Su *et. al* [16]. They do a singular value decomposition (SVD) analysis of the image.

The singular values of the SVD image decomposition are arranged from largest to smallest. So by using the first k weights, the small weights at the end are discarded and an approximation is obtained without losing too much detail. This is similar to what happens during image blurring. The large scale details of an image are retained (such as rough shapes) while smaller scale details are discarded. Interpreting this in terms of eigen-images, the smaller singular values that relate to small scale details bear smaller weights for blurred images. This leads to the conclusion that the first few most significant eigen-images therefore have higher weight for a blurry image compared to that of a clear image. This can be extended to finding the amount of blur in regions of a single image. The image can be analysed in local patches around each pixel and SVD is used to calculate the singular value for the patches. The focus (as opposed to blur) level at a pixel is given by the following ratio

$$F_p = 1 - \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{j=1}^{n} \lambda_j} \tag{1}$$

This equation is applied to each pixel in the image to obtain pixel-wise dense estimates of focus.

Figure 6 shows the effect of applying the blur mask to the half-occlusion regions of Figure 5. It is evident that the spurious half-occlusion regions have been greatly reduced.

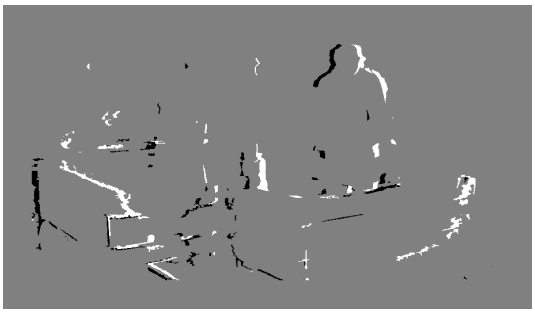

Figure 6. The half-occlusion regions passed by the blur mask. Spurious occlusions have been filtered out.

4. Detecting Pseudoscopy from Half-Occluded Region Centroid Differences

Once we have detected the left- and right-half-occluded regions for a stereo image pair, we form the histogram of horizontal coordinates of the pixels in the half-occluded regions. From the histogram we can compute the centroids of the right- and left-half-occlusion regions, and compute the difference between them.

An example of such a histogram is shown in Figure 7. The peaks in the histograms indicate the locations of the largest half-occluded regions. The centroid of the left-half-occlusion regions is left (lower x-coordinate) of the centroid of the right-half-occlusion regions, which under our assumption implies a stereoscopic presentation.

In this example, the difference between the centroids is 95.07. The Recall/Precision curve shown in Figure 8 shows the effect of the threshold on the centroid differences on the performance of the pseudoscopy detection.

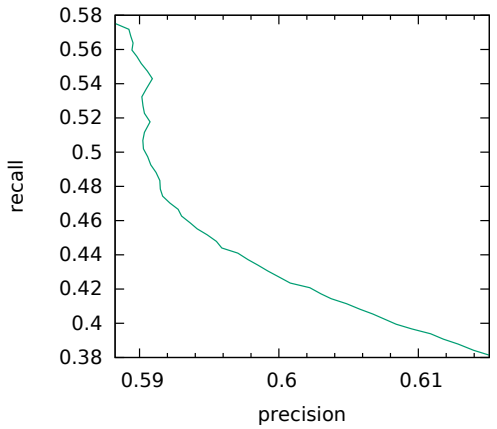

Figure 8. Precision vs. Recall of the centroid-difference pseudoscopy detection method as a function of the centroid difference threshold value (varied from 0 to 50). This curve is computed over the frames from all 52 video clips.

5. Detection of Pseudoscopy by Viewers of 3D Video

We have seen in the previous section that half-occlusion regions do have information which permits the detection of pseudoscopic display conditions. In this section we describe two psychophysical experiments that study whether humans can detect pseudoscopy and whether they make use of half-occlusion regions to do so.

5.1. General Experimental Setup

In the experiments subjects viewed 3D movies displayed on a 23 inch LG D2343p passive 3D display. The viewing distance was set to 85 cm. This distance was selected following the the optimal viewing distance from the specifications [12] of the monitor (90 cm), the ITU recommendation for such experiments [10] of 3.1H (88.97 cm). At the prescribed distance of 85 cm, a display pixel (0.2652×0.2652 mm) forms a visual angle of ≈ 64 seconds of arc.

The video stimuli used in both experiments was selected from five different publicly available 3D video databases assembled by European and Australian research groups [3, 4, 5, 6, 18]. Out of 135 short videos, 52 clips were selected. The clips were selected from the lot of clips that did not present any stereo defect to our knowledge. For instance, videos with stereo window violations were removed. Each video was full-color HD resolution. A frame of one of the selected videos is shown in figure 4.

The binocular gaze directions of the viewers' eyes were also tracked during the experiment. This was done using a Tobii X-120 eye-tracker, which is placed directly below the display screen. This provides gaze data for each eye synchronized to each video frame. Special purpose 3D video player software was created, which maintains synchronization with the display refresh cycle and the eye-tracker data stream. For 3D viewing, no shutter glasses could be used since they block the sight of the eye-tracker cameras. A passive 3D display was selected, employing circularly polarized glasses as a mean of separating left from right, identical to the ones used in RealD theatres. To maintain a stable geometric relationship between the screen and the eye-tracker throughout the experimentation, an adjustable stabilization jig was machined by combining a copy stand, a VESA monitor mount, workshop clamps and various hardware. Calibration of the eye tracker was performed before the start of each subject's experiment.

5.2. Methodology - Experiment 1

The goal of the first experiment was to determine whether pseudscopic display influenced viewers' perceived image quality.

A group of 10 naive subjects viewed each of the 52 clips for a total viewing time of 20 minutes. Half of the videos

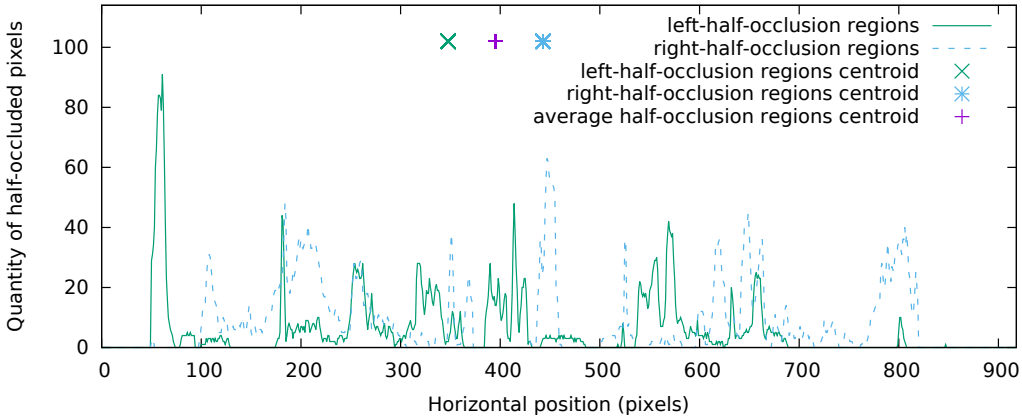

Figure 7. The histograms of the pixel x-coordinates in the left- (red) and right-half-occlusion (blue) regions.

were presented stereoscopically, and the other half, pseudoscopically. The stereo/pseudo halves were interchanged between each participant. For each video, 5 participants saw the stereo version and the 5 others viewed the pseudo version. Also, the video presentation order was shuffled to avoid fatigue induced bias in the results. The viewers were not given any specific task to carry out while viewing the movie clips, and they were free to focus on whatever aspects of the movie they wished.

After each video, the participants were asked to report their impression of the depth quality on an integer valued scale ranging from 1 to 5, labelled "Bad", "Poor", "Fair", "Good" and "Excellent" as suggested in [10]. Apart from providing valuable data, this task helps to ensure participant focus throughout the video presentation.

5.3. Results - Experiment 1

The opinion scores on the quality of depth lead to the distribution shown in the histogram of figure 9. There is some overlap in the two distributions, but overall, with an optimal threshold ($DQ < 4 \Rightarrow pseudoscopic$), learned through a random 10-fold cross validation, we can correctly detect pseudoscopy from raw[1] score $\approx 76\%$ (394/520) of the time. In most cases, a good "depth quality" ($DQ \geq 4$) is perceived in the stereoscopic class of videos. What is more interesting though, is what happens for $\approx 24\%$ (126/520) of the samples where the quality score cannot be used to predict pseudoscopy.

Out of the incorrectly classified 126 samples, 46 ($\approx 37\%$) are false positives. These samples were wrongly classified as pseudoscopic according to the depth quality score. From these errors, three videos were mainly concerned. The videos come from the Sigmedia Stereo Video Database[4] which is characterised by a soft focus. Moreover, the concerned videos are poorly illuminated. Both image characteristics significantly reduce the photometric contrast, which may explain the confusion.

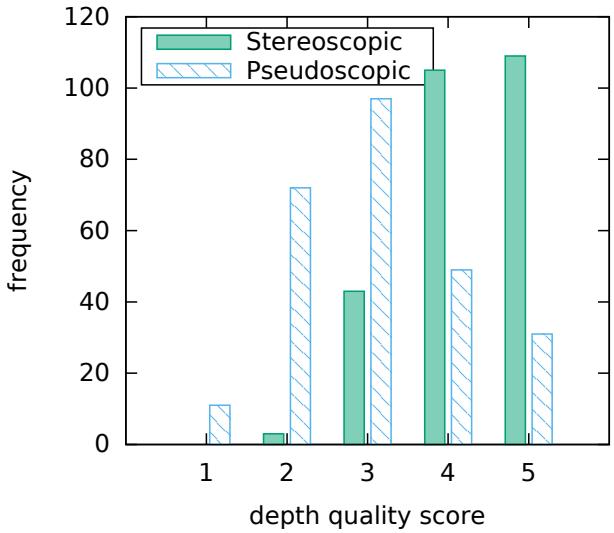

Figure 9. *Distribution of depth quality opinion scores for each class of videos.*

For 38 out of the 52 videos some participants gave a high depth quality score of 4, or 5, even if the video were shown pseudoscopically. These false negative events account for $\approx 63\%$ (80/126) of the errors of the classifier. This data confirms the observation of [19] cited in introduction: it is often hard to detect pseudoscopy.

For four of the videos, a large variability of quality scores is observed whether they are presented stereoscopically or pseudoscopically. We found that these videos are characterised by a low photometric and stereoscopic contrast. Also it was found that two of these videos have a significant camera alignment problem that was undetected pre-

[1]It was observed that a classifier based on a per-subject standarised dataset leads to very similar results.

viously. These videos alone account for 21 out of the 126 errors ($\approx 17\%$) made by our threshold classifier. On the other hand, a group of 16 videos is attributed a high score when viewed stereoscopically and scored low when viewed pseudoscopically. These video were found to bear large and horizontally aligned disparities. Spurious low quality scores for these videos account for $\approx 7\%$ of the errors in our classifier.

5.4. Methodology - Experiment 2

We have seen that half-occluded regions can be used in a computer vision algorithm to detect pseudoscopy. In the second experiment we analyzed the eye movements of viewers that were captured when viewers were specifically instructed to detect the occurence of pseudoscopy (as opposed to the free viewing task of experiment 1). In this experiment only 4 subjects were studied. Each subject viewed short (5-seconds long) extracts of the video clips used in experiment 1. Each subject viewed a total of 217 of the short video clips. Half of the clips were presented normally, while the other half were presented pseudoscopically. The presentation mode was randomized. The subjects were instructed to indicate at the end of each trial whether they thought the images were presented pseudoscopically or stereoscopically.

The half-occluded regions for each video frame was extracted using the techniques described earlier in the paper. Extracting the disparity maps for HD videos is quite computationally expensive. For convenience, the computing cluster of a local video FX company was used for this purpose, with the computation of the nearly 30,000 HD video frames distributed among 32 machines, totalling 768 3.4 GHz processing cores and 1.0 terabyte of memory. Even with such hardware, and reducing the HD resolution by half, the computation took around 8 hours. The selected algorithms for stereo matching and focus estimation can be efficiently ported to GPU-based machines for a significant speedup.

The eye movement data was analyzed and filtered to extract fixations. We used the Tobii I-VT algorithm [14] to detect fixations in our gaze data. This merges fixations that are close in time and space (within 50ms and 0.01 rad / 0.57 degrees). A fixation was deemed to be at a half-occlusion if more than 50% of the gaze samples within a fixation are within 5 pixels of the detection half-occlusion regions. The number of such fixations was tabulated for each trial.

5.5. Results - Experiment 2

The results of experiment 2, pooled across the 4 subjects (total of 868 trials), are summarized in Tables 1 and 2.

The fraction of correct responses was 0.72, which is very close to the performance observed in the free viewing case (experiment 1) when assuming that the differences in the viewer's quality judgements were related to detection of stereoscopy/pseudoscopy.

TP: 243	FP: 57
TN: 379	FN: 189

Table 1. Numbers of True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) for detection of pseudoscopy in experiment 2. Results are pooled across all 4 subjects, for a total of 868 trials.

The Recall for the 868 trials was 0.56 while the Precision was 0.81. While better automated results might be obtainable using a method that employed information beyond the half-occlusions (e.g. [11]), it is interesting that the human performance is comparable to that obtained by the machine vision approach based on half-occlusion regions.

The question arises whether there is any evidence that humans use the half-occlusion regions in making their decision regarding pseudoscopy. One way to judge this is to look at whether viewer's eye movements are preferentially directed towards half-occlusion regions. The results summarized in Table 2 suggests that this is indeed the case. First, 15.8% of eye fixations are made to half-occlusion regions over all trials, as compared to just 7.8% if fixations were made to random locations. Secondly, the number of fixations made to half-occluded regions is similar whether the display is stereoscopic or pseudoscopic. However, it is seen that in trials where the viewer correctly decided on pseudoscopy versus stereoscopy, the fraction of fixations on half-occluded regions rises to 17.4% whereas it falls to 11.9% when the viewer answered incorrectly. The latter number is approaching the value (7.8%) that would be obtained with a uniform random fixation location. This suggests that viewers are better able to distinguish between stereoscopic and pseudoscopic display if they look at the half-occlusion regions.

All: 0.158
Stereo: 0.160
Pseudo: 0.156
Correct: 0.174 (S.E.: 0.0016)
Incorrect: 0.119 (S.E.: 0.0018)
Image: 0.078

Table 2. Fraction of fixations made to half-occlusion regions. All: all trials. Stereo: all stereoscopic trials. Pseudo: all pseudoscopic trials. Correct: all trials in which the subject answered correctly. (S.E. is the standard error). Incorrect: all trials in which the subject answered incorrectly. Image: the ratio of pixels within 5 pixels of a half-occlusion region to the total number of pixels in the image.

6. Conclusion

This paper shows that there is information present in the left- and right-half-occlusion regions that can distinguish between stereoscopic and pseudoscopic display conditions. A machine vision method is presented that detects pseudoscopy using only the histograms of left and right half-occlusion pixel locations.

Two psychophysical experiments were done that investigated the ability of human viewers to detect, or be influenced by, pseudoscopic display. It was found that, during free viewing of HD 3D video imagery, humans judged pseudoscopic imagery to be of lower quality than stereoscopic imagery. It was also found that humans performed at a 72% rate in deciding whether a short 5-second video clip was presented stereoscopically or pseudoscopically. During these detection tasks, viewers fixated on half-occlusion regions at a much higher frequency than would be indicated by random chance. It was observed that viewers more frequently fixated on half-occlusion regions when making correct decisions then when they were incorrect.

Our conclusion is that humans are able to distinguish pseudoscopic display from stereoscopic display, and that they fixate on half-occlusion regions in doing so.

7. Acknowledgements

The authors would like to thank Mokko Studio for their financial support and for the use of their computing cluster. We would also like to thank the Consortium en innovation numérique du Québec (CINQ) for their grant supporting our research.

References

[1] D. Akimov, A. Shestov, A. Voronov, and D. Vatolin. Automatic left-right channel swap detection. In *3D Imaging (IC3D), 2012 International Conference on*, pages 1–6. IEEE, 2012.

[2] P. N. Belhumeur and D. Mumford. A bayesian treatment of the stereo correspondence problem using half-occluded regions. In *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92., 1992 IEEE Computer Society Conference on*, pages 506–512. IEEE, 1992.

[3] E. Cheng, P. Burton, J. Burton, A. Joseski, and I. Burnett. Rmit3dv: Pre-announcement of a creative commons uncompressed hd 3d video database. In *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*, pages 212–217. IEEE, July 2012.

[4] D. Corrigan, F. Pitie, V. Morris, A. Rankin, M. Linnane, G. Kearney, M. Gorzel, M. O'Dea, C. Lee, and A. Kokaram. A video database for the development of stereo-3d post-production algorithms. In *Visual Media Production (CVMP), 2010 Conference on*, pages 64–73, Nov 2010.

[5] European Broadcasting Union. 3dtv public test set, June 2012.

[6] L. Goldmann, F. D. Simone, and T. Ebrahimi. A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video. In *Proc. SPIE*, volume 7526, pages 75260S–75260S–11, San Jose, USA, 2010. International Society for Optics and Photonics. Available at http://mmspg.epfl.ch/3dvqa.

[7] K. He, J. Sun, and X. Tang. Guided image filtering. In *Computer Vision–ECCV 2010*, pages 1–14. Springer, 2010.

[8] A. Hosni, M. Bleyer, C. Rhemann, M. Gelautz, and C. Rother. Real-time local stereo matching using guided image filtering. In *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, pages 1–6, July 2011.

[9] A. Humayun, O. Mac Aodha, and G. Brostow. Learning to find occlusion regions. In *Proceedings of CVPR*, pages 2161–2168, 2011.

[10] ITU. Subjective methods for the assessment of stereoscopic 3dtv systems, October 2012. Available at http://www.itu.int/rec/R-REC-BT.1438.

[11] J. Lee, C. Jung, C. Kim, and A. Said. Content-based pseudoscopic view detection. *Journal of Signal Processing Systems*, 68(2):261–271, 2012.

[12] LG Electronics. *Owner's Manual: Cinema 3D Monitors*, September 2013.

[13] N. D. Narvekar and L. J. Karam. A no-reference image blur metric based on the cumulative probability of blur detection (cpbd). *Image Processing, IEEE Transactions on*, 20(9):2678–2683, 2011.

[14] A. Olsen. The tobii i-vt fixation filter. Technical report, Tobii Technology, March 2012.

[15] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz. Fast cost-volume filtering for visual correspondence and beyond. In *CVPR*, pages 3017–3024, june 2011.

[16] B. Su, S. Lu, and C. L. Tan. Blurred image region detection and classification. In *Proceedings of the 19th ACM International Conference on Multimedia*, MM '11, pages 1397–1400, New York, NY, USA, 2011. ACM.

[17] H. Tong, M. Li, H. Zhang, and C. Zhang. Blur detection for digital images using wavelet transform. In *Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference on*, volume 1, pages 17–20 Vol.1, June 2004.

[18] M. Urvoy, M. Barkowsky, R. Cousseau, Y. Koudota, V. Ricorde, P. Le Callet, J. Gutierrez, and N. Garcia. Nama3ds1-cospad1: Subjective video quality assessment database on coding conditions introducing freely available high quality 3d stereoscopic sequences. In *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*, pages 109–114, July 2012.

[19] R. Zone. *3-D filmmakers: Conversations with creators of stereoscopic motion pictures*. Number 119. Scarecrow Press, 2005.