Research Article

The Effect of Applying 2D Enhancement Algorithms on 3D Video Content

Sin Lin Wu,[1,2] Jorge E. Caviedes,[2] Lina Karam,[3] and Ingrid Heynderickx[4,5]

[1] *Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands*
[2] *Intel Corporation, 5000 W Chandler Boulevard, Chandler, AZ 85284, USA*
[3] *Arizona State University, Tempe, AZ 85287, USA*
[4] *Eindhoven University of Technology, Postbus 513, 5600 MB Eindhoven, The Netherlands*
[5] *Philips Research Laboratories, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands*

Correspondence should be addressed to Sin Lin Wu; sin.lin.wu@intel.com

Enhancement algorithms are typically applied to video content to increase their appeal to viewers. Such algorithms are readily available in the literature and are already widely applied in, for example, commercially available TVs. On the contrary, not much research has been done on enhancing stereoscopic 3D video content. In this paper, we present research focused on the effect of applying enhancement algorithms used for 2D content on 3D side-by-side content. We evaluate both offline enhancement of video content based on proprietary enhancement algorithms and real-time enhancement in the TVs. This is done using stereoscopic TVs with active shutter glasses, viewed both in their 2D and 3D viewing mode. The results of this research show that 2D enhancement algorithms are a viable first approach to enhance 3D content. In addition to video quality degradation due to the loss of spatial resolution as a consequence of the 3D video format, brightness reduction inherent to polarized or shutter glasses similarly degrades video quality. We illustrate the benefit of providing brightness enhancement for stereoscopic displays.

1. Introduction

Postprocessing is nowadays very common on commercially available TVs. It is widely used to make content look more appealing to the viewers. Postprocessing algorithms can generally be classified into two categories. One category of algorithms aims at restoring the video by reducing the occurrence and/or visibility of artifacts resulting from compression and transmission errors. Deblocking and denoising algorithms typically fall in this category [1, 2]. The second category of algorithms aims at enhancing the quality of the content, typically by applying sharpness, contrast, or color enhancement [2–5]. In this paper, we study the resulting video quality when applying such enhancement algorithms originally designed for 2D video to a pair of 2D views of stereoscopic 3D video.

Sharpness enhancement is typically done with a peaking algorithm, in which mid to high frequency parts of the signal are amplified [2, 6]. As a result, all edges become

sharper. Various implementations of peaking for 2D content have been discussed in the literature [7–9]. One possible improvement over the standard peaking algorithm [7] is to make it dependent on the content by using block-based content-adaptive sharpness enhancement; as such, a different filter is applied to edges, details, or textures. Another possible extension is to integrate a noise reduction step in the peaking algorithm in order to avoid that the noise in the signal is amplified by the peaking algorithm [9]. Contrast enhancement is commonly done by performing local or global histogram equalization or correction [2, 10]. A possible side effect of this process, however, is the generation of color artifacts, resulting from the desaturation of colors in areas of the image where the histogram equalization considerably reduces the intensity. Hence, researchers developed more advanced contrast enhancement algorithms, in which color saturation and lightness are compensated for possible loss due to histogram equalization [11, 12].

The postprocessing algorithms described above are widely used on 2D displays. Currently, 3D displays can be found in consumer video products; so it is logical to evaluate to what extent postprocessing optimized for 2D displays may be reused for 3D displays. Several technologies are currently used to render 3D video content. These display technolo-gies can be divided into two classes: autostereoscopic and stereo-scopic displays [13–17]. Autostereoscopic displays use technologies that are able to show a slightly different view of an image to each eye of a viewer without using external means, such as glasses. This can be done either by creating two views by means of a parallax barrier included in the LCD display and using eye tracking to adapt the two views to the location of the eyes of a viewer or by creating multiple views using a sheet of lenticular lenses in front of an LCD display [13–16]. Autostereoscopic TVs were not yet commercially available at high quality at the time of this research, and so we evaluated the more viable approach to 3D TV at that time, being two separate views provided to a stereoscopic 3D display.

The first generations of 3D TVs consisted mainly of stereoscopic TVs, which necessitated glasses be used to see the 3D effect. These displays can make use of different types of glasses: passive ones based on either color filters or polarizing filters and active ones based on shutter glasses that open and close synchronously with the content presented on the screen [14–16]. In all cases, the glasses are not completely transparent, and so they block some of the light emitted from the TV, resulting in a loss of brightness. Generally, 3D TVs receive the 3D content signal in a side-by-side format. This means that the left- and right-eye views are subsampled to half the original resolution and packed into one frame. This may have a detrimental effect on the overall quality of the 3D content.

Whereas the effect of video enhancement algorithms is well documented for 2D video content [18], enhancement for 3D content is relatively unexplored. Papers that do discuss signal processing for 3D content focus mainly on compression and transmission formats [19–21] but hardly touch upon options or algorithms for sharpness, contrast, or color enhancement. Enhancement methods for 3D should enhance features such as sharpness, colorfulness, and contrast and offer depth adjustments. Currently, the most common way to apply enhancement algorithms to 3D content consists of simply applying known 2D enhancement algorithms in the so-called dual pipe processing for stereoscopic streams [22]. This implies that the left- and right-eye video content is independently enhanced. In some commercial 3D TVs and computer displays, additional corrections for the 3D content are added to this standard processing chain. For example, the brightness loss due to the glasses used in a stereoscopic display is compensated for by applying a brightness boost, such as the NVIDIA 3D LightBoost technology for Acer and ASUS displays and the Quattron technology for Sharp. In addition, in most commercial 3D displays, the end user may control the level of enhancement applied, but the default setting is usually the same for 2D and 3D content. In some cases specific enhancement settings are disabled when 3D content is displayed, basically because it is known that the enhancement may have an adverse effect on 3D content.

In all these cases, however, publications or technical details about the enhancement algorithms used are sparse, and so their effect cannot be consistently evaluated. As such, for optimizing the quality of 3D content, it remains important to evaluate to what extent enhancement applied in 2D carries over in 3D content without adjustments for depth.

The research reported in this paper focuses on the effect of 2D video enhancement for 3D content, as applied in consumer TVs. Standard 2D enhancement algorithms are applied to the 3D content and the effect of this enhancement is compared to the effect of the same enhancement applied to the corresponding 2D content. This comparison shows whether comparable quality improvement can be expected in 3D content as in 2D content when using 2D enhancement algorithms. The enhancement is applied in two ways: once with the real-time enhancement settings as implemented by the manufacturer of the TVs and once with offline software-based enhancement and real-time playback of the output. Since we use commercially available consumer 3D TVs, frame-compatible side-by-side content is used, and, in the comparison, we particularly look into the aforementioned issues for stereoscopic displays, that is, the resolution and brightness loss.

First, the experimental setup, including the equipment and video sources used, is discussed in Section 2. Next, the experimental procedure of the five experiments we conducted is provided in Section 3. Then, results showing how the 2D enhancement algorithms affect 2D and 3D content are presented in Section 4 and further discussed and concluded in Section 5.

2. Experimental Setup

The five experiments carried out in this research were done with two TVs in a side-by-side setup, and the subjects were requested to score the perceived quality of the content on the two TVs at the same time. The quality scales on the evaluation form were also placed side by side; one scale for each TV. Scoring was done in a similar manner as the double stimulus continuous quality scale (DSCQS) method; however, guidelines were added to the scoring scale. These guidelines consisted of descriptive definitions of how the quality of the sequences had to be assessed next to a numerical value ranging from 1 to 5 [23]. An example of the scoring scale and the guidelines are given in Figure 1 and Table 1.

2.1. Displays. Two TVs of the same brand and model were used, namely, two Sony Bravia XBR 46HX929. Both TVs were carefully calibrated in such a way that their color and contrast matched using CalMAN calibration software and a Konica Minolta CS-200 chroma meter. To see the stereoscopic content, active shutter glasses of the type X103 Xpand Universal were used. The subjects were sitting at a distance of 2 meters from the TVs, corresponding to 3.5 times the height of the TV screen.

2.2. Postprocessing. Two types of postprocessing were used in this research: offline processing (with standard processing

TABLE 1: Guidelines used in this research.

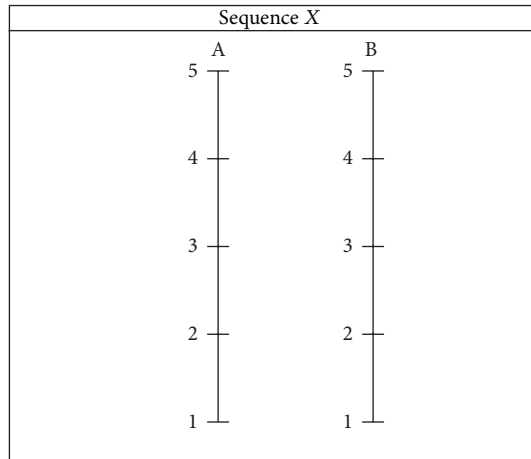| Score	Description
5	Excellent overall quality: all desirable features are at their best, including naturalness, compelling depth, and color/sharpness/contrast; no artifacts and no discomfort
4	Good overall quality and desirable features, including depth which is present to the right extent
3	Acceptable overall quality: depth is present but not compelling; artifacts if any have small or no impact
2	Fair overall quality: noticeable depth issues including unnatural/incorrect/artifacts; artifacts are noticeable and annoying
1	Poor or bad overall quality: very poor for most desirable features, serious depth issues, or other artifacts very annoying, causing serious fatigue

FIGURE 1: The scoring scales used in this research.

TABLE 2: Settings for the original and enhanced displays.

Setting	Original	Enhanced
Color	60	90
Color temperature	Neutral	Warm 2
Gamma	0	+1
Sharpness	20	50

in the TV turned off as much as possible) and real-time processing (using the standard processing chain of the TV). In both cases the images were enhanced in terms of sharpness and in terms of color and contrast. Sharpness enhancement in the offline processing was based on a peaking algorithm [7]. Color and contrast enhancement was done with the Joint Luminance Color and Contrast Enhancement (JLCCE) algorithm, as described in [11, 12]. The final output of the offline processing was subsequently played back on both TVs, while both TVs were in their matching calibrated settings. The offline processing used 3 levels with a low, medium, and high setting for both the sharpness enhancement and the joint color and contrast enhancement. For the real-time processing, the related settings on the TV were either on or off, and so, in the on-state, the processing chain as implemented in the TVs was used. Settings in the user menu of the TVs for color, color temperature, gamma, and sharpness were adjusted to the values shown in Table 2. Due to the proprietary nature of the algorithms of Sony, no further details can be given about the algorithms used in the TVs.

2.3. Stimuli. All experiments used the same three video clips "Balloon," "Mall," and "PedXing," screenshots of which are given in Figures 2(a), 2(b), and 2(c), respectively. After the first experiment, the fourth video clip "BCS," shown in Figure 2(d), was replaced by the clip "Suspension," shown in Figure 2(e). The replacement was motivated by the quality of the video clip "BCS"; since it originally was interlaced, it

had to be deinterlaced, which resulted in some artifacts. In addition, the flashing lights of the police motorcycles were disturbing in 3D.

All stimuli had a spatial resolution of 1920×1080 pixels and a temporal resolution of 60 Hz. The stereoscopic information was included in the side-by-side frame compatible format, which implied that the left- and right-eye view each had a resolution of 960×1080 pixels. 2D content was produced either by showing an upscaled version of the left view of the original video clips or by showing the left-view information of the 3D content to both the left and right eye. Each stimulus had duration of 15 seconds. The subjects saw each stimulus twice in loop mode, resulting in a total viewing time of 30 seconds per stimulus. After each stimulus, a homogeneous midgrey still image was displayed for 6 seconds.

3. Experimental Procedure

The research reported in this paper consisted of five experiments. In each experiment, the original (unprocessed) video (considered as the reference) was shown on one TV side by side with the enhanced video (hereafter referred to as stimulus) on the other TV. The stimulus could be either 2D or 3D content, in which case the reference was also 2D or 3D content, respectively. Whether the stimulus or reference content was shown on the left-hand or right-hand TV was randomized in all trials.

3.1. Experiment 1. In the first experiment, 2D original content was compared to 2D postprocessed content and 3D original content to 3D postprocessed content. To display the 2D content, the TV was set in its 2D mode, and, as a consequence, the subjects did not wear 3D glasses. The video signal to the TV was the left frame of the 3D content, upscaled offline to a full HD (1920×1080 pixels) frame. The 3D content was displayed in the TV's 3D mode and required the subjects to wear 3D glasses. The enhancement was done via

FIGURE 2: Stimuli used in the experiments: (a) Balloon, (b) Mall, (c) PedXing, and (d) BCS, which was replaced after the first experiment by (e) Suspension.

the post-processing chain of the TV (real-time), as explained in Section 2.2. Ten subjects participated in this experiment. Each subject had normal or corrected to normal vision, good stereo vision as tested with the Randot Stereotest, and no colorblindness as tested with the Ishihara colorblindness test. The subjects had to score a total of eight pairs of stimuli: 2 depth levels (i.e., 2D content and 3D content) × 4 sources.

3.2. Experiment 2. In the second experiment, again 2D original content was compared to 2D postprocessed content and 3D original content to 3D postprocessed content, while, now, the TV was in 3D mode (subjects wore glasses) for both the 2D and 3D content. The 2D stimuli were provided by displaying twice the left-view information, once for the right eye and once for the left eye. In addition, to gain more control on the postprocessing, the postprocessing was done offline, as explained in Section 2.2. Postprocessing was applied at three different levels (i.e., low, medium, and high) independently for the two enhancement algorithms (sharpness enhancement and joint color and contrast enhancement), resulting in 6 enhancement levels (i.e., Sharpness High, Sharpness Medium, Sharpness Low, JLCCE High, JLCCE Medium, and JLCCE Low). Twenty subjects participated in this experiment; they all had normal or corrected to normal vision, good stereo vision, and no colorblindness, tested in the same way as for experiment 1. To check consistency in the scoring behavior of the subjects, the originals were also added as stimuli. As a consequence, the total number of stimuli each subject had to score was 52 pairs: 2 depth levels (i.e., 2D content and 3D content) × 4 sources × 6 enhancements (i.e., 3 sharpness levels + 3 color and contrast levels) + 4 originals.

3.3. Experiment 3. In experiment 2, it was noticed that some subjects rated the scenes on both TVs differently, even when exactly the same content was displayed. These (often

substantial) differences in scores could be considered as measurement error given that both TVs had been previously calibrated. Therefore, only subjects from experiment 2 who were able to recognize the same scene quality on both TVs were selected for the third experiment. A more detailed description of how we selected these six subjects is given in Section 4.1. The third experiment was performed to test the accuracy of the selected subjects. To do so, we tested their consistency in scoring for the exact same comparisons as used in experiment 1 (except that the "BCS" video clip was replaced with the "Suspension" video clip as indicated in Section 2.3). Therefore, all TV settings in this experiment were exactly the same as in experiment 1, and the subjects were asked to score eight stimuli pairs in total: 2 depth levels (i.e., 2D content and 3D content) × 4 sources.

3.4. Experiment 4. To investigate a possible effect of a different way of postprocessing between real-time (by the TVs) and offline, we repeated experiment 2, but now with real-time enhancements performed by the TVs. Again, as in experiment 2, 2D original content is compared to 2D postprocessed content and 3D original content to 3D postprocessed content, but with the postprocessing implemented in the TVs. Additionally, as in experiment 2, both the 2D and 3D content were displayed on the TVs in their 3D mode; so the subjects were required to wear the glasses during the whole experiment. In order to measure possible effects with the most sensitive and reliable group of participants, the same six subjects as selected in experiment 3 were used. They had to score eight stimuli pairs in total: 2 depth levels (i.e., 2D content and 3D content) × 4 sources.

3.5. Experiment 5. The main aim of the fifth experiment was to investigate the effect of spatial resolution and brightness reduction (as imposed by the stereoscopic display format and

TABLE 3: Overview of the experiments and their content.

	Comparison		Processing
Experiment 1	2Din2D	3Din3D	Real-time processing
Experiment 2	2Din3D	3Din3D	Offline processing
Experiment 3	2Din2D	3Din3D	Real-time processing
Experiment 4	2Din3D	3Din3D	Real-time processing
Experiment 5			
Resolution experiment	2Din2D	2Din3D but without glasses	Real-time processing
Brightness experiment	2Din3D	2Din3D but without glasses	Real-time processing

glasses) on the perceived quality. To do so, we used only 2D content and showed it on the TVs either in their 2D mode or in their 3D mode. To measure the effect of spatial resolution without an effect of brightness reduction, we used 2D content in the 2D and 3D viewing mode, but in the 3D viewing mode the content was viewed without glasses. Hence, the 2D content was assessed in full (2D mode) and half (3D mode) spatial resolution at the same brightness. To investigate the effect of reduced brightness at the same spatial resolution, we used 2D content in the 3D viewing mode, but the content was viewed once in the 3D viewing mode with glasses and once in the same mode without glasses. It has to be noted that, for the latter comparison, the horizontal spatial resolution of the input was half the original resolution (linked to the side-by-side frame compatible 3D mode); this reduction in spatial resolution was unavoidable in testing the effect of brightness. In total, five of the six subjects who participated in experiments 3 and 4 participated in this experiment. Each subject had to score 12 stimuli pairs of 2D content, consisting of 4 pairs (i.e., one for each original content) in 2D viewing mode + 4 pairs in 3D viewing mode without glasses + 4 pairs in 3D viewing mode with glasses.

An overview of all five experiments is given in Table 3. Note that in this table the notation 2Din2D refers to the situation in which 2D video content is shown on the TVs in their 2D viewing mode (so at full spatial resolution and in the absence of glasses). The notation 2Din3D refers to the situation in which 2D content is shown on the TVs in their 3D viewing mode (hence, at half the spatial resolution and with 3D glasses). Obviously, the 3D content was displayed in the TVs' 3D viewing mode (hereafter referred to as 3Din3D).

4. Results

The initial subject pool of 20 subjects was larger than the six subjects that participated in the later experiments. How these subjects were selected is discussed first; this selection turned out to be critical to obtain reliable results for the experiments. After having discussed the subject selection procedure, the results of the various experiments are combined in order to be able to draw some conclusions. To give an overview of the results of the experiments, all the mean scores and their standard deviation are given in Table 4. In the next sections, these results are analyzed thoroughly.

All statistical analyses are performed with the software package IBM SPSS Statistics 19. Although we report means

and 95% confidence intervals, we performed nonparametric statistical tests, because of the small number of data points in some cases and because of the lack of normality in most of the data.

4.1. Subject Selection: Rejection Criteria for Experiment 3. In experiment 2 (all pairs of videos were shown in 3D mode with glasses), we included eight pairs of original versus original; that is, each of the four original videos was shown once in 2D and once in 3D on both TVs simultaneously. For these pairs, we calculated the absolute difference in scores between the two TVs. This absolute difference was expected to be zero, as the content on both (calibrated) TVs was equal. The actual spread found in the absolute difference is shown in the boxplot of Figure 3(a). Clearly, some subjects were not able to recognize that the two TVs were showing the same original content. Absolute differences in scores on exactly the same videos were as large as 3 or 4 units on the 5-point scoring scale as indicated by individual points in Figure 3(a). These large differences cast doubt on the reliability and accuracy of the other scores as obtained in experiments 1 and 2. Therefore, we decided to select a subset of more experienced participants by using the following criteria: for the selected participant, the mean of the absolute difference between the scores of original versus original had to be lower than 0.25 or the median had to be lower than 0.15. Six participants (out of the 20 subjects that participated in experiment 2) fulfilled these criteria. The resulting boxplot of these selected subjects for the comparison of original versus original as obtained in experiment 2 is shown in Figure 3(b).

Clearly, all outliers are removed (as compared to Figure 3(a)), and, for most video pairs, the spread in the absolute difference is reduced. For all of the considered video pairs absolute differences smaller than 1 unit on the scoring scale are found. The consequences of this subject selection procedure are further discussed in Section 5.

With the subset of selected subjects, we repeated experiment 1 in the so-called experiment 3. Note that experiment 3 used only three original video sequences that were also used in experiment 1 (since the fourth scene was changed after being used in experiment 1 only). Hence, we here present in the comparison only the results of the scenes that were common between experiments 1 and 3. To show that the six selected subjects in experiment 3 were indeed more consistent than the subjects of experiment 1, we compared

TABLE 4: Overview of all the mean scores and standard deviations for all clips and all experiments listed in Table 3.

	Processing	Mode	Balloon		Mall		PedXing		Suspension	
			μ	σ	μ	σ	μ	σ	μ	σ
Experiment 1	Original	2Din2D	3.45	0.53	3.37	0.57	3.63	0.50	—	—
		3Din3D	3.53	0.82	3.16	0.68	3.71	0.56	—	—
	Enhanced	2Din2D	4.03	0.64	3.93	0.71	4.21	0.61	—	—
		3Din3D	3.79	0.93	3.37	0.83	4.03	0.78	—	—
Experiment 2	Original	2Din3D	3.15	0.79	2.85	0.86	3.42	0.81	3.09	0.84
		3Din3D	3.41	0.59	3.11	0.78	3.61	0.68	3.35	0.62
	Enhanced	2Din3D	2.86	0.70	2.82	0.89	3.11	0.85	2.88	0.90
		3Din3D	3.03	0.77	3.05	0.66	3.23	0.69	3.56	0.62
Experiment 3	Original	2Din2D	3.25	0.33	3.02	0.47	3.27	0.36	3.23	0.53
		3Din3D	2.63	0.53	2.49	0.47	3.06	0.45	3.31	0.58
	Enhanced	2Din2D	4.01	0.49	3.83	0.42	4.10	0.58	4.12	0.46
		3Din3D	3.65	0.46	3.38	0.81	4.07	0.34	4.09	0.45
Experiment 4	Original	2Din3D	2.50	0.68	2.43	1.02	2.55	0.83	2.40	0.75
		3Din3D	2.56	0.78	2.67	0.66	3.03	0.37	3.27	0.53
	Enhanced	2Din3D	3.36	0.88	2.95	0.87	3.28	1.26	3.36	1.02
		3Din3D	3.43	0.74	3.48	0.64	4.05	0.33	4.19	0.49
Experiment 5	Original	2Din2D	2.92	0.58	2.87	0.76	3.09	0.59	2.88	0.66
		2Din3D	2.47	0.63	2.58	0.78	2.76	0.85	2.73	0.93
		2Din3D*	2.57	0.61	2.53	0.78	2.83	0.45	3.01	0.80
	Enhanced	2Din2D	3.71	0.44	3.62	0.62	3.91	0.56	3.55	0.52
		2Din3D	3.35	0.80	3.27	0.78	3.57	0.86	3.46	1.10
		2Din3D*	3.39	0.67	3.35	0.73	3.89	0.53	3.67	0.83

2Din3D*: 2D content shown in 3D mode without the viewer wearing glasses.

both results, showing in Figure 4, for all six pairs used in experiments 1 and 3, the mean and 95% confidence interval. As the size of the latter depends on the number of subjects used, we randomly selected six subjects of experiment 1 and calculated their mean score and 95% confidence interval. To reduce chance in the results, we repeated this procedure six times and calculated the average on the means and 95% confidence intervals. These results are shown in Figure 4(a), whereas Figure 4(b) shows the mean and 95% confidence interval of the results of experiment 3. Comparing both figures shows that the subjects of experiment 3 score 3D content more consistently and explicitly higher than the subjects of experiment 1. Surprisingly, the 95% confidence intervals are somewhat smaller for the results of experiment 1 (mean size of confidence intervals is 0.51 units on the 5-point scoring scale) than for the results of experiment 3 (mean size of the confidence intervals is 0.60).

4.2. Real-Time versus Offline Processing. In experiment 2, the content was processed offline and, in experiment 4, the content was processed in real-time. Hence, to compare the effect of processing on the perceived quality for both 2D and 3D content, we could directly compare the results of experiments 2 and 4. Obviously, for experiment 2, we only used the results of the selected subset of participants, in order to make the scores comparable to those of experiment 4.

We first tested for (the limited set of results of) experiment 2 the effect of stimulus on the overall quality scores with a Friedman test. The results showed a significant effect (χ^2 = 135.222, df = 95, P = 0.004) of stimulus on perceived quality, and, thus, additional analyses were necessary to determine whether the significant effect was attributed to video content (i.e., 4 source levels), depth (i.e., 2 levels; 2D content and 3D content), and enhancement (i.e., 7 levels, resulting from the three levels in sharpness enhancement and three levels in contrast enhancement + the originals). A Friedman test showed that the video source did not have a significant effect on perceived quality (χ^2 = 7.400, df = 3, P = 0.060), although there was some trend (i.e., P < 0.1). Also the depth mode did not significantly affect perceived quality as tested with a Wilcoxon signed ranks test (Z = −1.153, P = 0.249). Finally, also the enhancement level of the video did not significantly affect perceived quality as tested with a Friedman test (χ^2 = 7.964, df = 6, P = 0.241). Hence, neither of the underlying factors in itself sufficiently explained the effect of stimulus on perceived quality, but the three factors may have reinforced each other to generate an overall effect of stimulus on perceived quality. Nonetheless, enhancement, albeit not significantly, degraded rather than improved the perceived quality. Indeed, the high sharpness enhancement setting scored lowest (μ = 2.868, σ = 0.753), while the low JLCCE enhancement scored highest (μ = 3.246, σ = 0.451) but was very close to the mean score of the original (unprocessed) content (μ = 3.205, σ = 0.476).

We performed the same analyses on the results of experiment 4 and found again an effect of stimulus on the quality
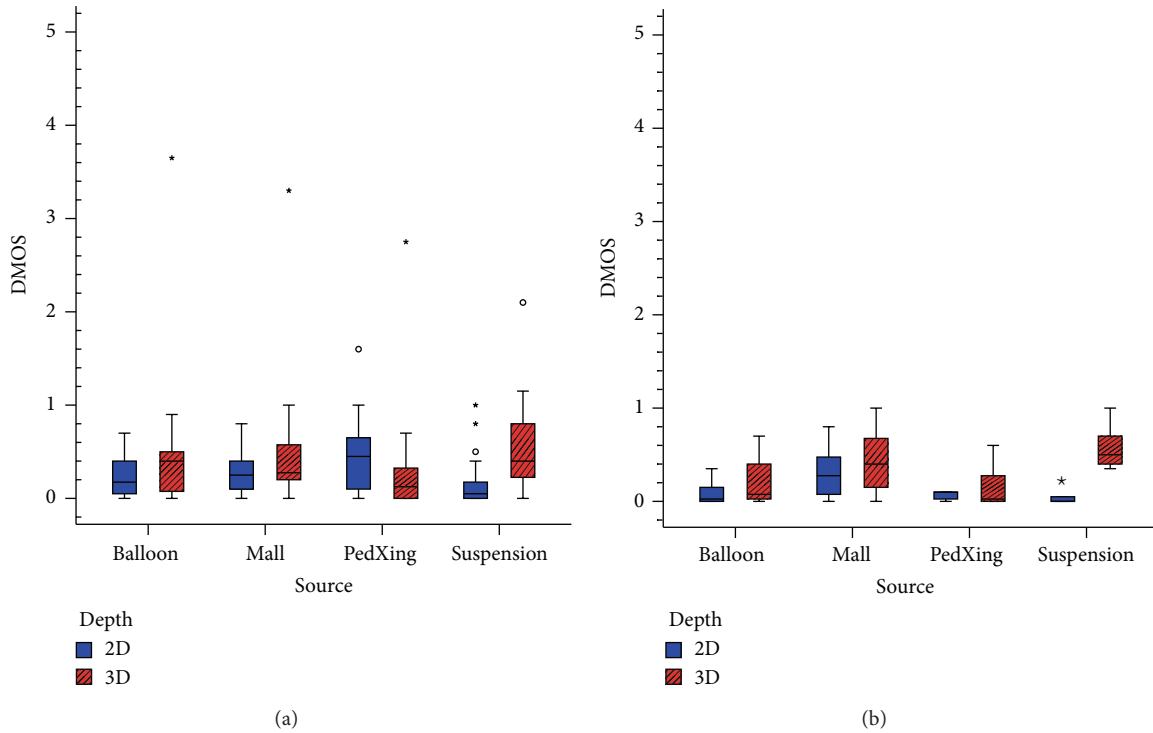
(a)

(b)

FIGURE 3: Boxplots of the absolute difference in scores for the eight pairs in experiment 2 comparing original versus original (either for 2D content or for 3D content): (a) the boxplot showing the results of all subjects of experiment 2 and (b) the boxplot showing the results after selecting a subset of the subjects.

scores with the Friedman test (χ^2 = 54.481, df = 15, P < 0.001). Subsequent analyses showed that the video source significantly affected quality (χ^2 = 8.085, df = 3, P = 0.044). The additional two Wilcoxon signed ranks tests (i.e., one for the effect of depth and one for the effect of enhancement, the latter now only having two levels, with enhancement being on or off) showed no significant effect of depth (Z = −1.363, P = 0.173), but a significant effect of enhancement (Z = −2.201, P = 0.028) on perceived quality. The mean scores illustrated that the source video "Mall" (μ = 2.881, σ = 0.612) scored on average lowest, while the source video "Suspension" (μ = 3.304, σ = 0.674) scored highest. Enhanced content (μ = 3.513, σ = 0.581) had a higher perceived quality than original content (μ = 2.674, σ = 0.570), and—although not significant—3D content (μ = 3.333, σ = 0.427) scored higher than 2D content (μ = 2.853, σ = 0.867).

Because of the difference in enhancement levels between the applied offline processing and real-time processing, it is inappropriate to directly compare the effect of both ways of processing on perceived quality. As might be expected, the effect of real-time postprocessing on quality for both the 2D and 3D content was significantly higher than the effect of the offline postprocessing. The latter mainly had a negative impact on the observed quality for both the 2D and 3D content. The lowest setting of offline processing was hardly visible to most participants, while the highest setting was not appreciated. Apparently, the level of enhancement applied in the real-time processing was chosen more appropriately. The issue is further discussed in Section 5.

4.3. Enhancement in the Various TV Modes and the Effect of Reduction in Resolution and Brightness.

Two viewing modes were used to display 2D content in this research. In experiment 3, 2D content was shown in 2D mode (viewing mode 1: 2Din2D, so at full spatial resolution and without glasses), while, in experiment 4, 2D content was shown in 3D mode (viewing mode 2: 2Din3D, so at half spatial resolution and with glasses). In addition to the 2D content, the 3D content was scored in both experiments (i.e., experiments 3 and 4). First, the effect of stimulus on the perceived overall quality was determined by performing a Friedman test. The stimuli consisted of combinations of source, viewing mode, and enhancement. The result showed that there was a significant effect of stimulus on the overall quality scores (χ^2 = 120.027, df = 31, P < 0.001). Additional analyses were done to determine which variable—video content (i.e., 4 source levels), depth (i.e., 2 levels; 2D content and 3D content), or processing (i.e., 2 levels; original versus enhanced content)— contributed to the significant effect. A Friedman test showed that the source had a significant effect on the quality scores (χ^2 = 10.627, df = 3, P = 0.014). The source "Mall" scored on average the lowest (μ = 3.03, σ = 0.51) and "Suspension" scored on average the highest (μ = 3.50, σ = 0.53). The variable depth had no significant effect as tested with a Wilcoxon signed ranks test (Z = −0.943, P = 0.345). Because the 2D content had two different viewing modes (i.e., 2Din2D versus 2Din3D), an additional analysis was performed on the effect of this variable. A Wilcoxon signed ranks test showed that there was a significant effect between the two viewing modes
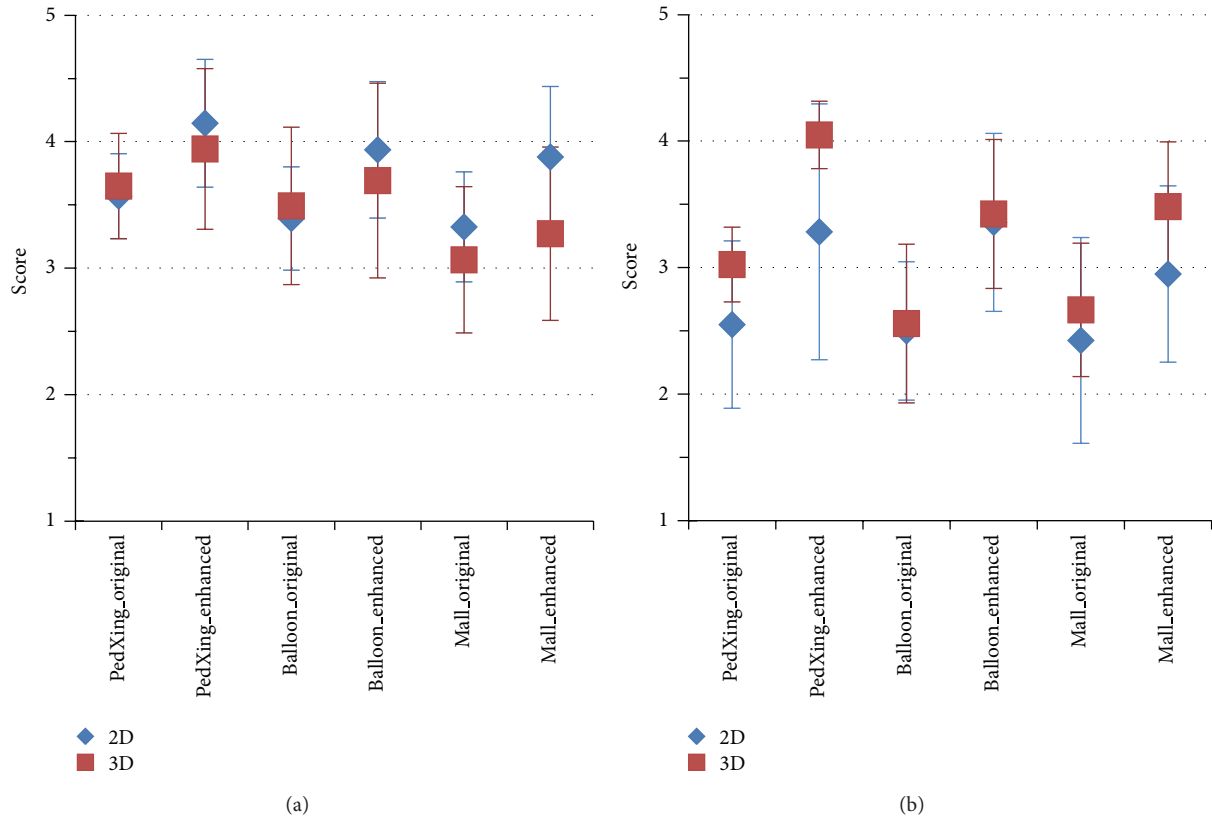
(a)

(b)

Figure 4: A comparison of the 95% confidence intervals between (a) experiment 1 and (b) experiment 3. Since the number of participants was different in experiments 1 and 3, we corrected for this difference by six times randomly selecting six participants of experiment 1 and using the average of the corresponding 95% confidence intervals.

($Z = -2.201$, $P = 0.028$). The 2Din2D viewing mode scored higher ($\mu = 3.60$, $\sigma = 0.35$) than the 2Din3D viewing mode ($\mu = 2.85$, $\sigma = 0.87$). Figure 5 shows the difference in mean quality between the 2D content displayed in 2D mode (2Din2D) and in 3D mode (2Din3D), including the 95% confidence interval, and illustrates that the 2Din2D scores are higher than 2Din3D scores.

Furthermore, Figure 5 illustrates that the mean scores and confidence intervals measured at different times (i.e., in experiments 3 and 4) for the 3D content are practically the same. Another Wilcoxon signed ranks test showed that there was a significant effect of the processing on the scores ($Z = -2.201$, $P = 0.028$). The enhanced content scored higher ($\mu = 3.71$, $\sigma = 0.43$) than the unprocessed content ($\mu = 2.85$, $\sigma = 0.47$).

Since the difference in quality scores for 2D content in the two viewing modes may be explained by a reduction in spatial resolution and brightness, experiment 5 was performed. The main purpose of this experiment was to test whether the reduction in spatial resolution or in brightness had the biggest effect on quality. Figure 6 shows the results of this experiment for the original content and enhanced content for the three different 2D viewing modes, that is, the normal 2D viewing mode (2Din2D, i.e., full spatial resolution and with the original brightness), 2D viewing in 3D mode without glasses (2Din3D*, i.e., half the spatial resolution but at the

original brightness), and 2D viewing in 3D mode with glasses (2Din3D, i.e., half the spatial resolution and at about half the brightness). A Friedman test was performed on the effect of the stimulus on the overall quality score and showed that there was a significant effect ($\chi^2 = 74.226$, df = 23, $P < 0.001$). The stimuli consisted of combinations of the variables, video content (i.e., 4 source levels), viewing mode (i.e., 3 levels: 2Din2D, 2Din3D*, and 2Din3D), and processing (i.e., 2 levels; original or enhanced content). A closer look into the variables by performing Friedman tests showed that source ($\chi^2 = 7.800$, df = 3, $P = 0.050$) had a significant effect on perceived quality, while viewing mode ($\chi^2 = 1.600$, df = 2, $P = 0.449$) did not. With a Wilcoxon signed ranks test a significant effect was found for the processing on the overall quality scores ($Z = -2.023$, $P = 0.043$). The results showed that the enhanced content ($\mu = 3.56$, $\sigma = 0.57$) scored higher than the original content ($\mu = 2.77$, $\sigma = 0.64$), which is also obvious from Figure 6.

5. Discussion and Conclusions

To get more accurate results, a subset of subjects was selected from the initially larger pool of subjects based on their ability to match scores for identical content. Using trained subjects for perception experiments was discussed in the literature
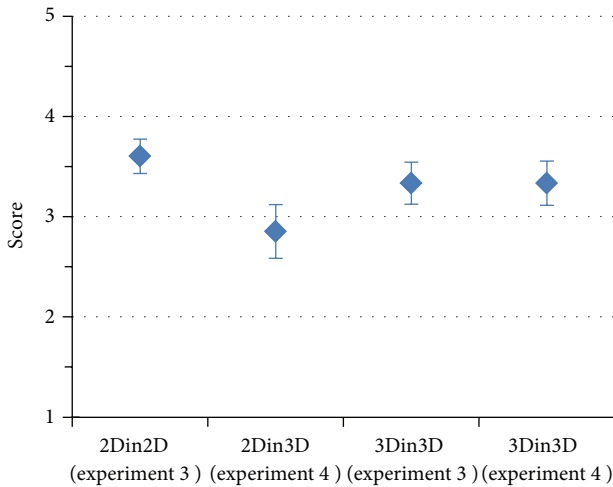
FIGURE 5: Comparison of the mean quality scores of experiments 3 and 4, in which a different viewing mode was used for the 2D content (i.e., experiment 3 used the 2D mode at full spatial resolution and no glasses, while experiment 4 used the 3D mode at half spatial resolution and with glasses).
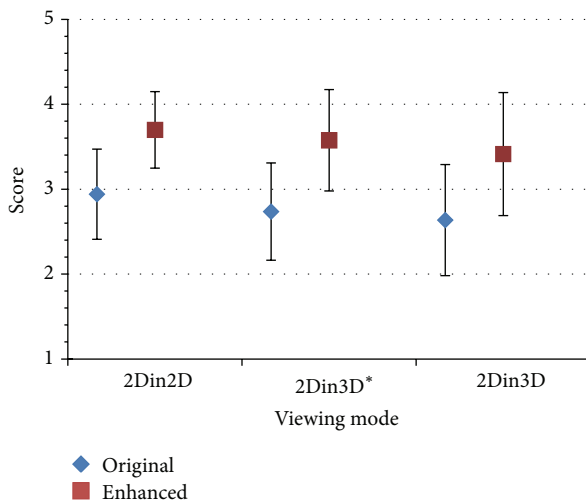

FIGURE 6: The mean scores and their respective 95% confidence intervals for the reduction of brightness and resolution are given for the original and enhanced content.

before; more particularly selection and training of subjects have been proven to have added value for subjective audio experiments [24]. For our experiments 3, 4, and 5 and also for the further use of part of the results of experiment 2, we used a similar concept; we selected the subset of subjects that were able to produce consistent scores. An advantage of using these selected subjects is that their scores are more reliable and vary less, but, on the other hand, the subject pool is reduced and doing statistical analysis is not reliable for small populations. Nonetheless, even when doing most of the experiments with only six subjects, we were able to find a substantial number of significant main effects on the quality scores. We even found more significant effects with

the reduced number of participants than with the larger pool of participants, indicating that most participants were adding noise to the quality scores. The relevant question that arises is how representative these selected subjects are to the average consumer. If only 30% of the population is able to consistently perceive quality differences between 2D and 3D original and enhanced content, we still need to find more powerful algorithms to improve the visual quality of 3D content. On the other hand, we should not forget that most selected participants were, through their profession, experienced in judging quality of postprocessing and 3D content. Therefore, the lack of scoring accuracy in the overall pool of subjects may be simply a matter of 3D viewing experience that may improve as the public becomes more accustomed to watching 3D content.

In some experiments reported in this paper, we found a significant effect of source video on perceived quality. The mean scores for the different experiments showed that some videos consistently scored low on quality, while others scored high on quality. Typically, the video "Mall" mostly scored lowest, while the videos "PedXing" and "Suspension" had a higher quality score. We even discarded one of the original videos after the first experiment, removing it from the rest of the experiments, because its quality was too low to evaluate effects of depth and enhancement. In addition, our results seem to suggest that only for some videos did adding stereoscopic depth improve the overall quality perception. As such, these results might imply that whether adding stereoscopic information improves the perceived overall quality of the content depends on specific video characteristics, a hypothesis that requires additional research to test.

The comparison of results of experiments 2 and 4 showed that offline enhancement as used in our experiment 2 did not improve the perceived quality of the content. When applying sharpness and contrast-color enhancement separately, their lowest setting was hardly noticeable to the subjects, while their highest setting was not appreciated. The real-time processing, on the other hand, did show a positive contribution to the perceived overall quality. This result may have been achieved by a carefully tuned combination of sharpness enhancement and contrast-color enhancement or by additional processing steps for the 3D mode as compared to the 2D mode. Since we used commercially available Sony TVs for which the exact processing chain is not disclosed, we cannot comment on whether the real-time video enhancement in 3D viewing mode added specific processing steps.

One possible processing step might be adding depth information in the 2D enhancement algorithms. In this research, we just applied 2D enhancement algorithms to each view of the 3D content and did not take into account the stereoscopic depth information. But the latter may be included, for example, by making use of a depth map. This depth map may, for example, help to target certain areas in an image for enhancement, for example, applying sharpness enhancement only to objects in the foreground, while keeping the natural blur for objects in the background. Although of potential merit, generating a depth map from a stereo pair does not guarantee accurate depth information; quite often generated depth maps still contain objects with improperly

assigned depth values. In the latter cases, enhancing the sharpness or color/contrast of these objects may be more harmful than beneficial.

When focusing solely on the real-time processing, and assuming that known 2D enhancement algorithms are applied to the 3D content, our experiments show that these known 2D enhancement algorithms provide a first viable approach to enhance 3D content. Nonetheless, the weak effects found with the built-in processing and the absence of quality improvement for the offline processing suggest that improved video enhancement algorithms are needed. As mentioned above, improvements may result from depth-based processing, of which first attempts are published in literature [25–29]. But our results imply that the development of 3D-specific enhancement algorithms would need to be justified by a considerable further improvement in visual quality at limited costs.

With the technologies now available on the consumer market, it is important to keep in mind that, with the use of glasses to view stereoscopic content, the brightness and spatial resolution are reduced. One of our experiments shows that the loss in brightness and the loss in spatial resolution affect the overall quality score; however, this is not significant. This might be due to the small group of subjects. Some commercial products already address the brightness loss by implementing different technologies to obtain a brightness boost. In addition, our experiments were limited to the use of active glasses to display stereoscopic content. We expect our results to be valuable for the use of passive glasses as well, but since autostereoscopic 3D displays differ considerably from stereoscopic displays, different enhancement approaches may be needed for autostereoscopic displays.

In conclusion, the research on stereoscopic video processing presented in this paper addresses fundamental questions on the enhancement required for proper visualization of stereoscopic 3D content and the proper evaluation of the resulting visual quality. The ability to discern 3D quality differences is a skill that may be improved through training and more exposure to 3D content. Participants that were able to consistently score quality reported improved quality for real-time enhanced 3D video, suggesting that standard 2D enhancement on the two views of the stereoscopic content provides a first option for 3D video enhancement. Nonetheless, the small improvements in perceived quality that were found in this study suggest that further progress in quality enhancement should be possible. A better understanding of the specifics of 3D processing and their quality assessment are essential for this progress.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Signal Processing Magazine*, vol. 14, no. 2, pp. 24–41, 1997.

[2] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, Prentice Hall Press, 2002.

[3] R. N. Strickland, C.-S. Kim, and W. F. McDonnell, "Digital color image enhancement based on the saturation component," *Optical Engineering*, vol. 26, no. 7, pp. 609–616, 1987.

[4] J. S. Lim, "Image enhancement," in *Digital Image Processing Techniques*, M. P. Ekstrom, Ed., pp. 1–51, Academic Press, 1984.

[5] G. M. Johnson and M. D. Fairchild, "Sharpness rules," in *Proceedings of the 8th IS and T/SID Color Imaging Conference: Color Science, Systems and Applications*, pp. 24–30, Scottsdale, Ariz, USA, November 2000.

[6] F. Oberti and I. Heynderickx, "Evaluation of optimal sharpness enhancement for different image content and different display technologies," in *Human Vision and Electronic Imaging IX*, Proceedings of SPIE, pp. 234–242, January 2004.

[7] P. Li, L. Shao, and D. Znamenskiy, "Block-based content-adaptive sharpness enhancement," in *Proceedings of the 13th IEEE International Symposium on Consumer Electronics (ISCE '09)*, pp. 18–21, May 2009.

[8] J. E. Caviedes, "Intelligent sharpness enhancement for video post-processing," in *Proceedings of the 14th European Signal Processing Conference*, September 2006.

[9] O. A. Ojo and T. G. Kwaaitaal-Spassova, "An algorithm for integrated noise reduction and sharpness enhancement," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 3, pp. 474–480, 2000.

[10] G. E. Hague, A. R. Weeks, and H. R. Myler, "Histogram equalization of the saturation component for true-color images using the C-Y color space," in *Applications of Digital Image Processing XVII*, vol. 2298 of *Proceedings of SPIE*, pp. 236–247, July 1994.

[11] A. Sarkar, M. D. Fairchild, J. Caviedes, and M. Subedar, "A comparative study of color and contrast enhancement for still images and consumer video applications," in *Proceedings of 16th Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications*, pp. 170–175, November 2008.

[12] A. Sarkar, J. E. Caviedes, and M. Subedar, "Joint enhancement of lightness, color and contrast of images and video," US 2010/0085487 A1, 2010.

[13] N. A. Dodgson, "Autostereoscopic 3D displays," *IEEE Computer*, vol. 38, no. 8, pp. 31–36, 2005.

[14] L. Hill and A. Jacobs, "3-D liquid crystal displays and their applications," *Proceedings of the IEEE*, vol. 94, no. 3, pp. 575–589, 2006.

[15] I. Sexton and P. Surman, "Stereoscopic and autostereoscopic display systems," *IEEE Signal Processing Magazine*, vol. 16, no. 3, pp. 85–99, 1999.

[16] J. Konrad and M. Halle, "3-D displays and signal processing," *IEEE Signal Processing Magazine*, vol. 24, no. 6, pp. 97–111, 2007.

[17] J.-Y. Son and B. Javidi, "Three-dimensional imaging methods Based on multiview images," *IEEE/OSA Journal of Display Technology*, vol. 1, no. 1, pp. 125–140, 2005.

[18] R. L. Lagendijk, P. M. B. van Roosmalen, and J. Biemond, "Video enhancement and restoration," in *Handbook of Image and Video Processing*, A. Bovik, Ed., pp. 227–241, 2000.

[19] L. Stelmach, W. J. Tam, D. Meegan, and A. Vincent, "Stereo image quality: effects of mixed spatio-temporal resolution," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 2, pp. 188–193, 2000.

[20] A. Smolic, K. Mueller, P. Merkle, P. Kauff, and T. Wiegand, "An overview of available and emerging 3D video formats and depth enhanced stereo as efficient generic solution," in *Proceedings of the Picture Coding Symposium (PCS '09)*, pp. 1–4, May 2009.

[21] P. Aflakia, M. M. Hannuksela, J. Häkkinen, P. Lindroos, and M. Gabbouj, "Subjective study on compressed asymmetric stereoscopic video," in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP '10)*, pp. 4021–4024, September 2010.

[22] L. Blondé, D. Doyen, and T. Borel, "3D stereo rendering new processing and perception challenges," Technicolor Whitepaper, 2010, http://3dathome.org/resources-white-papers.aspx.

[23] S. Wu, J. Caviedes, L. Karam, and I. Heynderickx, "Development of a practical methodology to evaluate the 3D visual experience," *Journal of Imaging Science and Technology*. In press.

[24] S. Bech, "Selection and training of subjects for listening tests on sound-reproducing equipment," *Journal of the Audio Engineering Society*, vol. 40, no. 7-8, pp. 590–610, 1992.

[25] C. C. Cheng, C. T. Li, Y. M. Tsai, and L. G. Chen, "A quality-scalable depth-aware video processing system," *SID Symposium Digest of Technical Papers*, vol. 40, no. 1, pp. 123–126, 2009.

[26] M. M. Subedar and L. J. Karam, "Increased depth perception with sharpness enhancement for stereo video," in *Stereoscopic Displays and Applications XXI*, vol. 7524 of *Proceedings of SPIE*, February 2010.

[27] W. Hachicha, A. Beghdadi, and F. A. Cheikh, "Combining depth information and local edge detection for stereo image enhancement," in *Proceedings of the 20th European Signal Processing Conference (EUSIPCO '12)*, pp. 250–254, August 2012.

[28] S. W. Jung and O. Choi, "Color image enhancement using depth and intensity measurement of a time-of-flight depth camera," *Optical Engineering*, vol. 52, no. 10, Article ID 103104, 2013.

[29] S.-W. Jung and S.-J. Ko, "Depth map based image enhancement using color stereopsis," *IEEE Signal Processing Letters*, vol. 19, no. 5, pp. 303–306, 2012.