EUROGRAPHICS 2010 / T. Akenine-Möller and M. Zwicker
(Guest Editors)

Volume 29 (2010), Number 2

Supplemental Material: Perceptual Study Details for Perceptually-motivated Real-time Temporal Upsampling of 3D Content for High-refresh-rate Displays

Piotr Didyk[1] Elmar Eisemann[1,2] Tobias Ritschel[1] Karol Myszkowski[1] Hans-Peter Seidel[1]

[1] MPI Informatik, Saarbrücken, Germany [2] Saarland University, Saarbrücken, Germany

The goal of this document is to provide complementary information on psychophysical experiments described in Sections 7 and 8. The experimental methods are described in the main paper, and here missing details concerning participants, materials and apparatus, experimental procedure and statistical data are given.

1. Participants

14 participants with normal or corrected-to-normal vision took part in the „Rating" and „Artifacts" experiments. Only a subset of 10 participants took part in the „Game" and „Stereo Vision" experiments, as well as additional study over the „Camel" scene. Subjects were compensated for their efforts with a small fee (14 $). Participants were recruited from the university campus and were mostly students of computer science. Subjects were naïve regarding the goal of the experiment and inexperienced in the field of computer graphics.

2. Materials and Apparatus

All stimuli were presented on a 22 inch (diagonal) Samsung 2233RZ 120 Hz display of resolution 1680×1050 that was connected to a personal computer with an NVIDIA GTX 260 running in the synchronization mode. The monitor was viewed by the subjects orthogonally at a distance of $60-80$ cm. The video sequences and images of resolution 512×512 have been used in all studies except „Game", where the full display resolution has been used. Experiments „Rating" and „Artifacts" required that three sequences are simultaneously shown next to each other in a horizontal arrangement.

3. Procedure

The participants were seated in front of a monitor running the experimental software in a room with controlled artificial lighting. They received standardized written instructions regarding the procedure of the experiment. In all experiments (except „Game") the time for each trial has been unlimited. In case of the „Game" experiment a unlimited-time practice session has been offered until the subject felt comfortable with the game.

In our study we did not have any restrictions concerning the experience of participants. They all played video games before but of course the level of experience varied. Although Green et al. [GB03] have shown that video games can modify visual selective attention, in our case, there seemed to be no direct correlation between detection of artifacts and the level of video game experience. It is important to mention that all subjects who noticed problems with artifacts, reported only slight differences between our upsampling and the original 120Hz rendering. This can be explained by the fact that our method, which upsamples 40 Hz signal, has significantly less information over time than the original 120Hz rendering.

4. Optical Flow Experiment

A clear drawback of our studies is that we did not include any practical solution relying on optical flow computation used in modern TV sets. The problem is that such algorithms are not revealed and it is currently not possible to send our output into any available TV because it cannot be externally fed with 120 Hz sequences. Therefore, we decided to experiment with one of the state-of-the-art optical flow techniques proposed in [ZPB07]. The technique is of significantly lower performance than ours (about 30 Hz on a modern GPU with a 512^2 resolution). In a precomputation, we interpolated in-between frames based on the two nearest keyframes, to obtain 120 Hz sequences. We did not include such obtained sequences in our study as visual artifacts have been readily visible. Further, such interpolation always implies a one-frame lag.

5. Comparison to other methods

In our experiment, we did not compare our method to all those described in Section 4. Comparing to black data insertion or backlight flashing is currently impossible due to technical reasons. Refresh-rates of above 200 Hz are needed and, even though such TV sets are available on market, they cannot work with more than a 60 Hz input signal. Further, some general drawbacks of these solutions exist (e. g. brightness and contrast reduction). Nevertheless, it could be interesting to combine our solution with those strategies in the future. We also did not compare to motion-compensated inverse filtering because such solutions cannot recover frequencies that are lost by the hold-type effect. Only high frequencies (unsharp masking) can be enhanced to slightly improve the perceived sharpness. We found that, nowadays, the best methods are those based on frame interpolation. For this reason, we compared our method to state-of-the-art implementations of optical flow, which are more accurate than those available in TV-sets. To make our study more challenging we compared our method to the ground truth and showed that our upsampling from 40 Hz to 120 Hz and the original 120 Hz rendering are almost indistinguishable in terms of blur.

6. Statistics

Refer to Figure 1., Tables 1. and 2.

References

[GB03] GREEN C. S., BAVELIER D.: Action video game modifies visual selective attention. *Nature 423*, 6939 (May 2003), 534–537. 1

[ZPB07] ZACH C., POCK T., BISCHOF H.: A duality based approach for realtime tv-l1 optical flow. In *DAGM-Symposium* (2007), pp. 214–223. 1

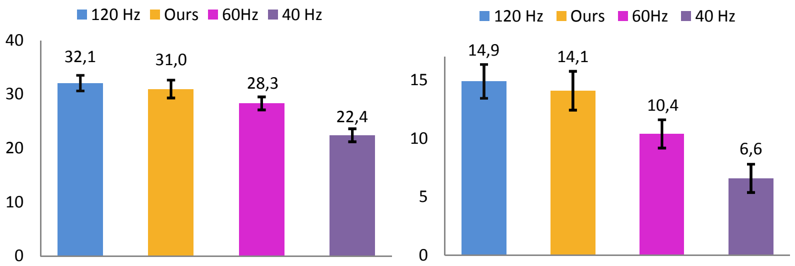

Figure 1: *Game statistics: The left plot presents the mean value of scores for all methods with ± 1 standard error of the mean (SEM). Such scores include correctly identified Landolt shapes as well as the full circles (toruses). The plot on the right shows only the mean value of the successes of Landolt shape detection with ± 1 SEM.*

	ANOVA			Our vs. 120Hz			Our vs. 40Hz		
Scene	$F(2,26)$	R^2	p	$t(13)$	p	Cohen's d	$t(13)$	p	Cohen's d
Sponza	302.07895	.93936	<.00001	-2.70153	.011988	-1.021081	16.86200	<.00001	6.373233
Tower	38.50120	.66380	<.00001	-1.54022	.135592	-.582147	5.92260	<.00001	2.238529
Fan	24.08909	.55264	<.00001	-.64210	.526431	2.130014	5.63549	<.00001	2.130014
Trees	43.10340	.68852	<.00001	.08069	.936312	.030495	10.17879	<.00001	3.847221
Camel	2.06101	.09559	.14110	—-	—-	—-	—-	—-	—-

Table 1: *"Rating" experiment: The table contains F-, p- and $R-$squared values for ANOVA applied for each scene independently to rating data for our temporal upsampling from 40 Hz to 120 Hz vs. 120 Hz and 40 Hz native rendering. $R-$squared values are computed as a ratio of the explained sum of squares to the total sum of squares. Also, t- and p- values as well as the effect size (Cohen's d) for pairwise comparison of our method with respect to 120 Hz and 40 Hz native rendering are given. Note that for the "Camel" scene already ANOVA shows that there is no main effect.*

Refresh rate	$t(9)$	p	Cohen's d
40 Hz	7.71	.00005	2.29894
60 Hz	4.25	.01000	1.02826
120 Hz	-2.18	.12062	-.54211

Table 2: *"Game" experiment: The table presents outcome of t-tests performed over the subject scores obtained for our temporal upsampling from 40 Hz to 120 Hz vs. native rendering with refresh rates 40 Hz, 60 Hz, and 120 Hz, respectively.*