CEIG - Spanish Computer Graphics Conference (2016)
Alejandro García-Alonso and Belen Masia (Editors)

Crossmodal perception in immersive environments

Marcos Allue Ana Serrano Manuel G. Bedia Belen Masia

Universidad de Zaragoza

Abstract

With the proliferation of low-cost, consumer level, head-mounted displays (HMDs) such as Oculus VR or Sony's Morpheus, we are witnessing a reappearance of virtual reality. However, there are still important stumbling blocks that hinder the development of applications and reduce the visual quality of the results. Knowledge of human perception in virtual environments can help overcome these limitations. In this paper, within the much-studied area of perception in virtual environments, we chose to look into the less explored area of crossmodal perception, that is, the interaction of different senses when perceiving the environment. In particular, we looked at the influence of sound on visual motion perception in a virtual reality scenario. We first replicated a well-known crossmodal perception experiment, carried out on a conventional 2D display, and then extended it to a 3D head-mounted display (HMD). Next, we performed an additional experiment in which we increased the complexity of the stimuli of the previous experiment, to test whether the effects observed would hold in more realistic scenes. We found that the trend which was previously observed in 2D displays is maintained in HMDs, but with an observed reduction of the crossmodal effect. With more complex stimuli the trend holds, and the crossmodal effect is further reduced, possibly due to the presence of additional visual cues.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality

1. Introduction and related work

With the proliferation of low-cost, consumer level, head-mounted displays (HMDs) such as Oculus VR or Sony's Morpheus, we are witnessing a reappearance of virtual reality. New applications are developed every day, going far beyond entertainment and gaming, and including advertising, virtual tourism, prototyping, medicine, scientific visualization, or education, to name a few. There are still important stumbling blocks that hinder the development of more applications and reduce the visual quality of the results; examples include limited spatial resolution, significant chromatic aberrations, tracking issues, limited processing capability leading to lag, subsequent motion sickness, or content generation. A relevant area which has received quite some interest but remains full of unanswered questions and open problems is how our perception is modified or altered when immersed in a virtual environment. Knowledge of human perception in virtual environments can help overcome the aforementioned current limitations; in the past, perception has helped tremendously in many computer graphics-related areas such as rendering [RFWB07], material modeling and acquisition [SRD08], or display [MWDG13]; a good review of applied perception in graphics can be found in the course by McNamara and colleagues [MMG11].

In this paper, within the much-studied area of perception in virtual environments, we chose to look into the less explored area of crossmodal perception, that is, the interaction of different senses when perceiving the environment. In particular, we looked at the *influence of sound on visual perception in a virtual reality scenario*.

Nowadays, a popular view in neuroscience holds that the human brain is structured into a large number of areas in which information is highly separated [Fod00]. This perspective assumes that mental processes such as perception -but also emotions or intentions- are limited to neural processes inside the brain and confined to particular areas. In the same way, it is often assumed that inputs coming from different perceptual modalities are processed in the brain independently and in different brain regions [Sam00].

However, the feeling of unified perceptions of objects and events is an ordinary experience. It suggests that information from different sensory modalities must somehow be bounded together in the brain in order to represent a single object or event [Pri06]. This assumption is cornerstone in most recent alternative neurodynamic views (as for example, bodily and sensorimotor approaches) in order to propose solid explanatory alternatives to traditional and internalist perspectives of brain organization [TE98,VM01]. In these alternative approaches, multisensory perception processes and different sensory modalities are understood as closely related through flexible integrations of the dynamics of brain by means of the emer-

DOI: 10.2312/ceig.20161307

gence of transient assemblies of neural synchronization when a unified perception arises [LVQ11]. Thus, a complete understanding of perception would require to know the different ways in which one sense modality is able to impact another, creating crossmodal illusions [SL10]. If we understood the interactions among perceptual modalities, we could shed light on the true mechanisms that support perceptual processes.

It is worth highlighting that, until very recently, the neural principles of multisensory integration and crossmodal illusions have remained unexplored. The modular view of the brain has been so strong with respect to the visual stimuli that it has been considered in the past as independent from other modalities. However, in recent years the interest in understanding crossmodal phenomena and illusions has increased substantially [Shi01]. Some of the deeper studies are those involved in alterations between auditory and visual senses. The best known example amongst these is the *ventriloquism effect* which refers to the perception of speech sounds as coming from a different direction than its real source, forced by the influence of visual stimuli from an apparent speaker [HHBW06]. Another well-known example is *the McGurk effect* [MH76] where lip movements of a subject are integrated with different but similar speech sounds.

The aim of the present paper is to investigate the effect of auditory spatial information on the perception of moving visual stimuli. We focus on the case of *motion perception* because previous studies have suggested that there should exist common neural substrates between the visual and auditory modalities [SL00]. The work is inspired in a classical experiment developed in the 90s where sound influenced ambiguous visual motion perception as proposed by [SSL97]. The authors found that when two objects -in a virtual and ambiguous simulation- moving along crossing trajectories reached the same position and then moved apart, they would be sometimes perceived by participants in the study as if they followed the same trajectory. However, in other cases, they reported that the objects reversed their direction as they would do following a collision. Sekuler et al. [SSL97] discovered that this ambiguity was solved when a sound emerged at the moment of coincidence of the objects, as this would show that the sensory information perceived in one modality (audition) could modulate the perception of events occurring in another modality (visual motion perception). Although the crossmodal effect reported by Sekuler and collaborators was accused of simply showing a cognitive limit rather than a genuine crossmodal perceptual effect, the authors opened the debate regarding the perceptual nature of many other crossmodal illusions between visual and auditory stimuli. For instance, the effect known as sound-induced flash illusion [SL00,SKS02] showed how the perception of a brief visual stimuli could be altered by concurrent brief sounds. When a single flash of light was showed together with two beeps, the perception changed from a single flash to two flashes. The reverse illusion could also occur when two flashes were accompanied by a single beep (which would be then perceived as a single flash).

These results revealed that unified and integrated perceptual constructs cannot be simply an assemblage of modality-specific components and that the traditional conception of perceptual experiences as an aggregate of ingredients (different type of sensorial stimuli) is not accurate. No particular modality of sensorial perception can be characterized entirely in isolation from the others. We consider that these conclusions on how crossmodal illusions are constituted could contribute significantly to current progress in perceptual research. Here, we take steps towards further understanding this phenomenon by performing two experiments: First, we reproduced the experiment of Sekuler et al. [SSL97] on a conventional, 2D display; and conducted, the same experiment on a HMD, with the aim of discovering whether the same trends in crossmodal perception are observed in HMDs as the ones observed on conventional displays (Experiment 1). Next, in Experiment 2, we extended the original experiment by modifying the original stimuli, to check whether the effects observed by Sekuler et al. still hold in the presence of more complex stimuli. With the exception of the replication of Sekuler et al.'s experiment in a conventional display (first part of Experiment 1), it is the first time, to our knowledge, that these experiments are performed with a virtual reality scenario. We describe both experiments in Section 2 and analyze and discuss the results in Section 3; final conclusions are drawn in Section 4.

2. Experimental procedure

We have performed two experiments in order to determine how much an immersive environment interferes with the crossmodal interaction between the visual and auditive systems. Our experiment is based in the work of Sekuler et al. [SSL97], where they explore the perceptual consequences of sound altering visual motion perception. In their experiments, they showed two identical disks that moved steadily towards each other, coincided, and then continued in the same direction. This scenario is consistent with two different interpretations: either the two spheres did not collide and continued in their original directions (they *streamed*), or they collided and *bounced*, changing their traveling direction. The goal of the experiment is to analyze whether a sound at the moment of the impact can affect the interpretation of the scenario.

We built upon Sekuler's work, and extended his experiment to virtual reality, aiming to explore the consequences on crossmodal interactions of introducing the user inside a more realistic and complex environment presented with a *head mounted display* (HMD). In order to ensure that the replication is accurate, we performed both the original 2D experiment on a conventional display, and the experiment with a 3D environment on an HMD. So that our 3D environment reproduced the 2D one faithfully, we projected our 3D scene on the screen, and ensured that all the measurements of the original 2D scene are maintained. We show the transformation between the two spaces in Figure 1.

2.1. Experiment 1

We first reproduced the experiment described in Sekuler's work both in a regular screen and in a HMD (*Oculus Rift DK2*) to determine if introducing a sound at the moment of the collision between the two spheres would promote the perception of bouncing. We sought to analyze the influence of presenting the action on an HMD on perception.

Stimuli The visual stimuli were rendered with *Unity*. They consisted of two spheres with radius *0.5 degrees*, placed over a white

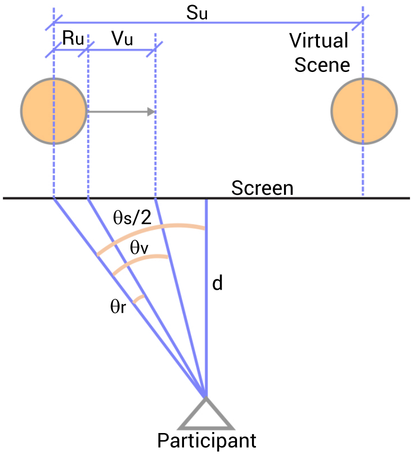

Figure 1: *Projection of our 3D scene on the 2D screen plane for reproducing the experiment described in the work of Sekuler et al. [SSL97].* θ_r *is the angle between the center and the radius of the sphere (*0.5 degrees*),* θ_v *represents the speed of sphere (*6 degrees per second*),* θ_s *is the angle between the centers of the spheres, and d is the distance to the screen. These angles are translated to distances in our 3D scene:* R_u *is the radius of the spheres,* S_u *is the distance between the spheres, and* V_u *is their speed when moving.*

plane. The material of the spheres was brown and very diffuse to avoid introducing additional visual cues. The two spheres were initially separated by a distance of *4.2 degrees*, and moved towards each other at a constant speed of *6 degrees per second*. After they coincided, they continued moving without changing their original direction. We show in Figure 2 the initial layout of the scene. In this scenario we presented three different visual conditions: the spheres moved continuously, paused one frame at the point of their coincidence, or paused two frames at the point of their coincidence[†]. These three visual conditions were presented together with one of the four following auditory conditions: no sound, accompanied by a brief click sound (frequency of *2000 Hz*, duration of *3 milliseconds*) triggered *150 milliseconds* before or after the coincidence, or accompanied by a brief click sound at the point of coincidence.

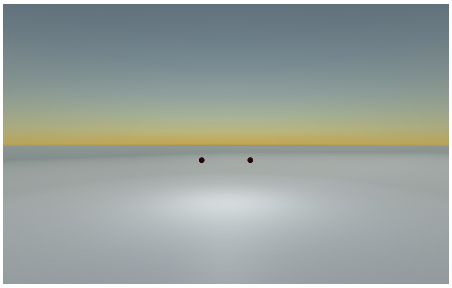

Figure 2: *Initial layout of the scene for Experiment 1.*

Participants Thirteen participants took part in the experiment, three women and ten men, with ages ranging from 18 to 28 years. They all had normal or corrected-to-normal vision.

Procedure During the experiment we presented a total of twelve different conditions to each participant, three visual (continuous movement, pause one or two frames at the coincidence) and four auditory (no sound, sound at, before, or after the coincidence). Each of these conditions was presented *ten* times, making a total of 120 trials that appeared in a random order. We performed two blocks of the same experiment ordered randomly: one displayed on a regular screen (*Acer AL2216W TFT 22"*), and the other one displayed on an HMD (*Oculus Rift DK2*).

Before the HMD block, the lenses of the *Oculus Rift DK2* were adjusted to the participant eyes. We additionally introduced a training session before this block, where we showed two spheres at different depths (see Figure 3) and the participant had to choose which one was closer. We presented ten trials of the training with spheres at random depths. With this training the user gets used to the device, setup, and answering procedure.

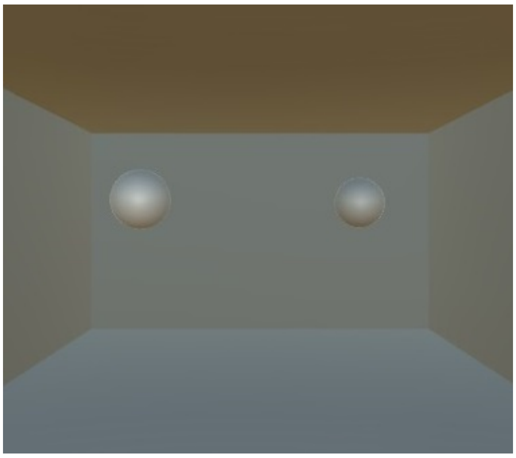

Figure 3: *Training for the HMD experiment, in the scene the participants saw two spheres at different depths, and had to indicate which sphere was closer, clicking the corresponding (right or left) mouse button.*

We guided the participants through the test by showing several slides with descriptions of each phase of the experiment. After each trial, a slide was displayed with the question *"Did the spheres bounce or stream?"*, and a visual aid indicating the participant to answer with a mouse click (right or left). The supplementary material shows screenshots of the test, including explanation and questions.

[†] The original experiment [SSL97] reported frames in a regular analog screen whose typical framerate is 25 frames per second. Since the framerate of our screen and the HMD (*Oculus Rift*) were very different, we adjusted the pause to last $1/25$ seconds. Therefore, throughout the paper the terminology is as follows: one frame is equivalent to $1/25$ seconds, and two frames are equivalent to $2/25$ seconds.

Data processing We first processed the collected data by rejecting those users with stereo vision problems. In order to do this, we discarded a user if during the training the percentage of successful answers was equal or under 70%. We further processed the data by rejecting outliers. To do this, we first calculated for each participant and for each of the twelve conditions the percentage of *bouncing* answers over the ten trials. Then we used the first and third quartiles (Q_1 and Q_3), and the interquartile difference (Q_d) to find outliers for each condition [HI87]. We discarded a condition if it fulfilled any of the following inequalities:

$$\begin{aligned} condition &< (Q_1 - K_d * Q_d) \\ condition &> (Q_3 + K_d * Q_d) \end{aligned} \tag{1}$$

with $Q_d = Q_3 - Q_1$ and $K_d = 1.5$. Additionally, if a participant was marked as an outlier for more than one condition, all the answers of the participant were discarded.

2.2. Experiment 2

For the second experiment we sought to further analyze the effect of a more realistic environment in the crossmodal interaction between the visual and auditory systems. In order to do this, we increased the realism of the scene in three different ways (we term them three *blocks*) while keeping the proportions between distances and speed of the spheres of the original experiment.

Stimuli The visual stimuli were rendered once again with *Unity*. We designed a new scene where the spheres are placed on a white table, inside a furnished room, and with a more realistic illumination. With respect to the first experiment we also increased the size of the spheres to *1 degree* of radius, and the distance between them to *8.4 degrees*, to make them more visible. In order to keep proportions, we needed to increase the speed accordingly. This increase of speed is shown in Figure 4 and described by Equation 2.

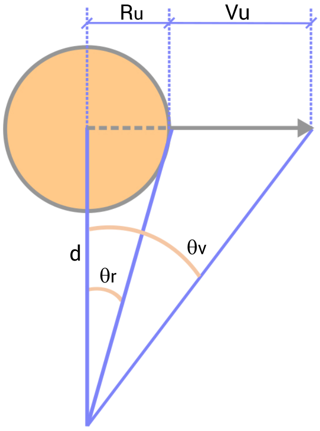

Figure 4: *Scheme showing the increase of* θ_v *when increasing* r. *For a small* θ_v, *the increase of* θ_v *is linear when increasing* r, *as shown in Equation 2.*

$$\begin{gathered} tan(\theta_v) = \frac{r+v}{d} \\ \theta_v \approx 0 \implies tan(\theta_v) \approx \theta_v \end{gathered} \tag{2}$$

For a small θ_v, the increase of θ_v is linear when increasing r. A screenshot of the initial layout of the scene for the first block of the experiment is shown in Figure 5. For the second block of the experiment, starting from the scene in the first block, we additionally introduced two more visual cues to the spheres. First, we increased the glossiness of the material of the spheres, and second, we slightly lifted the spheres over the table in order to have more visible shadows (see Figure 6). Finally, for the third block of the experiment, starting from the scene in the first block, we also rotated the plane of the collision between the spheres. We show a screenshot of the initial layout for this block in Figure 7.

Figure 5: *Initial layout of the scene with increased radius of the spheres (block 1) for Experiment 2.*

Figure 6: *Initial layout of the scene with increased radius of the spheres and additional visual cues (block 2) for Experiment 2.*

Figure 7: *Initial layout of the scene with increased radius of the spheres and rotated plane of the collision (block 3) for Experiment 2.*

Participants Twenty seven participants took part in the experiment, two women and twenty five men with ages ranging from 19 to 32 years. They all had normal or corrected-to-normal vision.

Procedure During the experiment we presented a total of six different conditions, two visual (continuous movement, pause two frames at the coincidence), and three auditory (no sound, *click sound* at, or after the coincidence). Based on the results of the first experiment we removed the visual condition with a pause of one frame because the percentage of bouncing perceived was similar to the one perceived with the pause of two frames, and the auditory condition corresponding to the sound before the coincidence, also because of its similarity with the sound after the coincidence. Each of these conditions was presented *ten* times, making a total of 60 trials that appeared in a random order. All the blocks of the experiment were presented in the *HMD*, and each participant performed three randomly ordered blocks that corresponded to the three scenes described in the *Stimuli* section, totalling 180 trials per subject. Before starting the test, the participants performed the same training described in Experiment 1.

Finally, in this experiment the slides with instructions about the test were shown on a frame on the back of the room striving to preserve as much as possible the realism of the environment.

Data processing We followed the same methodology as in Experiment 1 for rejecting outliers.

3. Results and analysis

In this section we analyze the results collected in our experiments, in particular we calculate for every user and condition the percentage of times subjects observed the spheres *bouncing* in the *ten* trials. We use repeated measures ANOVA to test whether each of the conditions have influence in the observed percentage of bounce responses. We need the repeated measures scheme because we measure the same independent variables (e.g., frames paused) under different conditions performed by the same subjects. We fix a significance value (p-value) of 0.05 in all the tests, and in those cases in which results from Mauchly's test of sphericity indicate that variances and covariances are not uniform, we use adjusted measures (Greenhouse Geisser correction [CW11]). Previous to the analysis, we perform outlier rejection as detailed in Section 2.

3.1. Experiment 1

The goal of this experiment was to test whether the effect of sound altering visual motion perception as reported in the experiments carried out by Sekuler et al. [SSL97] is also observed when reproduced in a virtual environment with an HMD. We wanted to test three factors: (i) the overall influence of the display (2D scene presented on a *screen*, or 3D environment presented on an *HMD*); (ii) the influence of the *sound* when the spheres collide; and (iii) the influence of the length of the *pause* at the point of coincidence between the spheres. We aggregate the percentages of all trials for every condition and perform a repeated measures ANOVA; results are presented in Table 1. We can conclude that all three factors have a significant effect in the percentage of bounce responses, since all the p-values are below 0.05. We show in Figure 8 the mean percentages of bounce responses for the tested factors (error bars represent the standard error of the mean). We observe that the percentage of bounce responses decreases when using the HMD display. However, the main findings of Sekuler's work hold: a sound at the moment of coincidence, and a pause of two frames at the point of coincidence promote the perception of bouncing. We believe that the decrease in perceived bouncing in the tests with the HMD comes from the increase in the amount of visual cues due to the stereoscopic view. We additionally show in Figure 9 simple (non-aggregated) mean percentages of bouncing and error bars for each condition. Sound promotes perception of bouncing when compared with the absence of sound; however, it has significantly less effect when reproduced after the point of coincidence. Still, there is a high tolerance for asynchrony between the sound and the visual input: even when the sound is delayed, the percentage of bounce responses increases. Also, as reported previously by Sekuler and others [SSL97, BBB93, SSB95], the overall percentage of bounce responses increases with the duration of the pause.

Table 1: *Results (*F-test *and* significance*) of the analysis of the data with repeated measures ANOVA for Experiment 1. We test the influence of three factors in the perceived percentages of bounce responses.*

	F	**Sig.**
Sound vs percent. bounce	83.664	0.000
Pause vs percent. bounce	63.528	0.000
Display vs percent. bounce	13.176	0.000

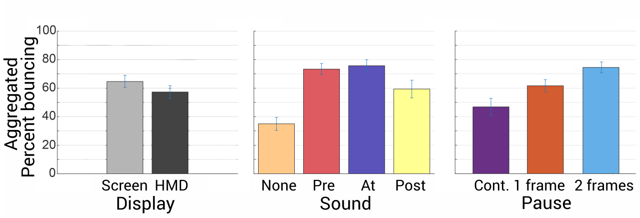

Figure 8: *Aggregated percentages of bounce responses and corresponding error bars (standard error of the mean) for the Experiment 1. From left to right: Percentages for two display conditions (screen or HMD), percentages for four auditory conditions (no sound, sound at, before, or after the moment of coincidence of the spheres), and percentages for three visual conditions (continuous movement, pause one, or two frames at the point of coincidence of the spheres).*

3.2. Experiment 2

The goal of this experiment was to test whether a more complex scene could influence the effect of sound altering visual motion perception. Again, we wanted to test three factors: the influence of each of the three scenes (three blocks) described in Section 2, the influence of the *sound* when the spheres collide, and the influence of the *pause* at the point of coincidence between the spheres. We aggregate the percentages for every condition and perform a repeated

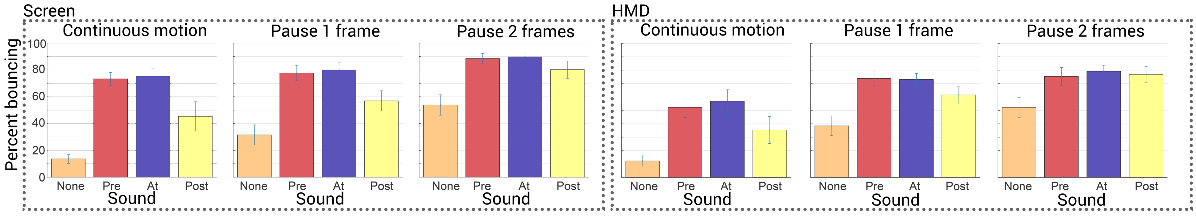

Figure 9: *Percentages (non-aggregated) and error bars (standard error of the mean) for Experiment 1. Left: Results for display on a 2D screen. Right: Results for display on an HMD. In both cases, from left to right: Plots for each of the three conditions of the pause factor (continuous movement, pause one, or two frames at the point of coincidence of the spheres). Each of the three plots shows the percentage of bounce responses for each of the four auditory conditions (no sound, sound at, before, or after the moment of coincidence of the spheres).*

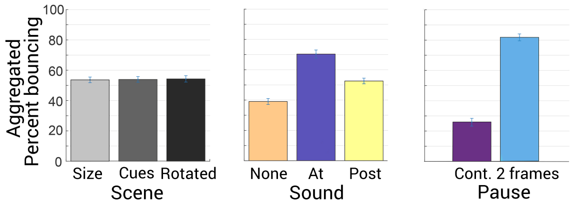

Figure 10: *Aggregated percentages and error bars (standard error of the mean) for the Experiment 2. From left to right: Percentages for the three different scenes or blocks (increase in the size of the spheres, additional visual cues in the spheres, or rotated plane of the movement); percentages for three auditory conditions (no sound, sound at, or after the moment of coincidence of the spheres); and percentages for two visual conditions (continuous movement, or pause two frames at the point of coincidence of the spheres).*

measures ANOVA; results are presented in Table 2. In Figure 10 we show the mean percentages of bounce responses for the tested factors, and the associated error bars representing the standard error of the mean. The analysis with the ANOVA reveals that, as before, there is a significant effect of the *sound*, and the *pause* in the perceived percentage of bounces. However, the p-value for the test with different scenes is very high, therefore we cannot draw any significant conclusion about the relationship between the three different scenes and the observed percentage of bouncing. When comparing Experiments 1 and 2 we can see that even when increasing the level of realism of the scene, the crossmodal effect of the sound altering the perceived motion still holds, although there is a general shift downwards of the percentage of bounce responses which can be observed by comparing the corresponding percentages of Figures 8 and 10. This shift downwards is possibly due to the presence of additional cues; however the high p-value of the scene factor, further indicates that there is no significant difference on the effect on crossmodal interaction between the three scenes (blocks) tested (i.e., no cue has proven to be significantly stronger or weaker in the detection of bouncing).

Table 2: *Results (*F-test *and* significance*) of the analysis of the data with repeated measures ANOVA for Experiment 2. We test the influence of three factors in the perceived percentages of bounces.*

	F	Sig.
Sound vs percent. bounce	124.137	0.000
Pause vs percent. bounce	845.386	0.000
Scene vs percent. bounce	0.220	0.977

4. Discussion and conclusions

In this paper, we have performed an exploration of crossmodal perception in virtual reality scenarios, in particular using an HMD. We have studied the influence of auditory signals in the perception of visual motion. To do so, we first replicated an existing experiment which demonstrated the existence of a crossmodal interaction between both senses with simple stimuli on a 2D conventional display. We were able to successfully replicate it, obtaining the same trends in the results, and then extended it to virtual reality with a HMD. We found that the same trends hold on an HMD (i.e., the factors explored had the same influence on the crossmodal effect), but that there is a reduction in the crossmodal effect. This reduction essentially means that there is a shift in the results towards a better accuracy of subjects in performing the tasks assigned in the HMD setup. This can be due to the presence of additional cues, in particular depth cues including binocular disparity and possibly motion parallax. A similar conclusion can be drawn in our second experiment: We repeated the first experiment (only on the HMD), with new subjects, and with more complex stimuli (we had three different variations of the initial stimulus) to see whether the effect would still hold with more realistic scenery. We once again observed a reduction of the crossmodal effect (subjects were better at detecting the correct behavior of the stimuli), which we hypothesize is due to the presence of additional cues, in this case pictorial cues (shading, perspective, texture).

Overall, we believe these are just a few steps in the exploration of crossmodal perception in virtual reality. In the future, we would like to expand these experiments by including other potentially influencing factors or effects, and by further increasing the complexity of the stimuli, which is required for the conclusions to be usable in a real virtual reality application. Additionally, further analysis of the first-order interactions of the factors studied is required.

5. Acknowledgments

The authors would like to thank Diego Gutierrez for fruitful insights and discussion. Ana Serrano was supported by an FPI grant from the Spanish Ministry of Economy and Competitiveness (project Lightslice).

References

[BBB93] BERTENTHAL B. I., BANTON T., BRADBURY A.: Directional bias in the perception of translating patterns. *Perception 22*, 2 (1993), 193–207. doi:10.1068/p220193. 5

[CW11] CUNNINGHAM D., WALLRAVEN C.: *Experimental Design: From User Studies to Psychophysics*, 1st ed. A. K. Peters, Ltd., Natick, MA, USA, 2011. 5

[Fod00] FODOR J.: *The mind doesn't work that way: The scope and limits of computational psychology.* Cambridge, MA: MIT Press, 2000. 1

[HHBW06] HAIRSTON D. W., HODGES D. A., BURDETTE J. H., WALLACE M. T.: Auditory enhancement of visual temporal order judgment. *NeuroReport 17*, 8 (2006), 791–5. 2

[HI87] HOAGLIN D. C., IGLEWICZ B.: Fine-tunning some resistant rules for outlier labeling. *Journal of the American Statistical Association 82*, 400 (1987), 1147–1149. 4

[LVQ11] LE VAN QUYEN M.: The brainweb of cross-scale interactions. *New Ideas in Psychology 29* (2011), 57–63. 2

[MH76] MCGURK H M. J.: Hearing lips and seeing voices. *Nature 264* (1976), 746–8. 2

[MMG11] MCNAMARA A., MANIA K., GUTIERREZ D.: Perception in graphics, visualization, virtual environments and animation. SIGGRAPH Asia Courses, 2011. 1

[MWDG13] MASIA B., WETZSTEIN G., DIDYK P., GUTIERREZ D.: A Survey on Computational Displays: Pushing the Boundaries of Optics, Computation, and Perception. *Computers & Graphics 37*, 8 (2013), 1012 – 1038. 1

[Pri06] PRINZ J.: *Is the Mind Really Modular?.* Stainton, Robert J. (Ed), (2006). Contemporary debates in cognitive science. Contemporary debates in philosophy. Malden: Blackwell Publishing, 2006. 1

[RFWB07] RAMANARAYANAN G., FERWERDA J., WALTER B., BALA K.: Visual equivalence: Towards a new standard for image fidelity. *ACM Trans. Graph. 26*, 3 (July 2007). 1

[Sam00] SAMUELS R.: *Massively modular minds: Evolutionary psychology and cognitive architecture.* P. Carruthers and A. Chamberlain, eds., Evolution and the Human Mind. Cambridge: Cambridge University Press, 2000. 1

[Shi01] SHIMOJO S. S. C. N. R. S. L. K. Y. . W. K.: Beyond perceptual modality: Auditory effects on visual perception. *Acoustical Science and Technology 22*, 2 (2001), 61–67. 2

[SKS02] SHAMS L., KAMITANI Y., SHIMOJO S.: Visual illusion induced by sound. *Cognitive brain research 14* (2002), 147–152. 2

[SL00] SHAMS L KAMITANI Y S. S.: What you see is what you hear. *Nature 408* (2000), 788. 2

[SL10] SHAMS L K. R.: Crossmodal influences on visual perception. *Physics of Life Reviews* (2010). 2

[SRD08] SILLION F. X., RUSHMEIER H., DORSEY J.: *Digital Modeling of Material Appearance*. Morgan Kaufmann/Elsevier, 2008. 1

[SSB95] SEKULER R., SEKULER A., BRACKETT T.: When visual objects collide: Repulsion and streaming. *Investigative Ophthalmology and Visual Science 36*, 50 (1995). 5

[SSL97] SEKULER R., SEKULER A. B., LAU R.: Sound alters visual motion perception. *Nature 385*, 6614 (1997), 308. 2, 3, 5

[TE98] TONONI G., EDELMAN G. M.: Consciousness and complexity. *Science 282* (1998), 1846–1851. 1

[VM01] VARELA F. L. J. P. R. E., MARTINERIE J.: The brainweb: phase synchronization and large-scale integration. *Nature reviews. Neuroscience 2* (2001), 229–239. 1