

Résumé

Peu d'études ont été menées sur l'influence de la stéréoscopie sur la perception d'un mixage audio au cinéma. Les témoignages de mixeurs ou les articles scientifiques montrent pourtant une grande diversité d'opinions à ce sujet. Certains estiment que cette influence est négligeable, d'autres affirment qu'il faut totalement revoir notre conception de la bande-son, aussi bien au niveau du mixage que de la diffusion.

Une première série d'expériences s'est intéressée à la perception des sons d'ambiance. 8 séquences, dans leurs versions stéréoscopique (3D-s) et non-stéréoscopique (2D), ont été diffusées dans un cinéma à des sujets avec plusieurs mixages différents. Pour chaque présentation, les sujets devaient évaluer à quel point le mixage proposé leur paraissait trop frontal ou au contraire trop «*surround*», le but étant de mettre en évidence une éventuelle influence de la stéréoscopie sur la perception de la balance avant/arrière d'un mixage audio. Les résultats obtenus ont rejoint ceux d'une expérience préliminaire menée dans un auditorium de mixage, où les sujets se trouvaient en situation de mixeur et devaient eux-mêmes régler la balance avant/arrière : l'influence de la stéréoscopie était faible et n'apparaissait que pour quelques séquences. Une troisième expérience fut conduite pour vérifier si les séquences pour lesquelles la perception de la balance avant/arrière était significativement impactée par la stéréoscopie étaient celles dont les différences entre versions 2D et 3D-s étaient les plus importantes en termes de profondeur visuelle perçue. Cependant, aucune corrélation n'a pu être trouvée.

Des études ont ensuite été menées sur la perception des objets sonores tels que dialogues et effets. Une quatrième expérience s'est intéressée à l'effet ventriloque en élévation : lorsque l'on présente à un sujet des stimuli audio et visuel temporellement coïncidents mais spatialement disparates, les sujets perçoivent parfois le stimulus sonore au même endroit que le stimulus visuel. On appelle ce phénomène l'*effet ventriloque* car il rappelle l'illusion créée par le ventriloque lorsque sa voix semble plutôt provenir de sa marionnette que de sa propre bouche. Ce phénomène a été très largement étudié dans le plan horizontal, et dans une moindre mesure en distance. Par contre, très peu d'études se sont intéressées à l'élévation. Dans cette expérience, nous avons présenté à des sujets des séquences audiovisuelles montrant un homme en train de parler. Sa voix pouvait être reproduite sur différents haut-parleurs, qui créaient des disparités plus ou moins grandes en azimut et en élévation entre le son et l'image. Pour chaque présentation, les sujets devaient indiquer si la voix semblait ou non provenir de la même direction que la bouche de l'acteur. Les résultats ont montré que l'effet ventriloque était très efficace en élévation, ce qui suggère qu'il n'est peut-être pas nécessaire de rechercher la cohérence audiovisuelle en élévation au cinéma.

Une cinquième et dernière expérience a permis d'étudier l'influence de la stéréoscopie sur les attentes des spectateurs en termes de cohérence audiovisuelle spatiale. Au cinéma, les objets sonores sont en général diffusés sur l'enceinte centrale, indépendamment de la position à l'écran des sources visuelles associées. Cependant, certains ingénieurs du son et chercheurs ont suggéré que la cohérence audiovisuelle spatiale pouvait améliorer significativement l'expérience des spectateurs, surtout dans le cas de films en relief. Dans cette expérience, les sujets devaient évaluer à quel point la bande-son leur paraissait « adaptée » à l'image pour 8 séquences projetées en 2D et en 3D-s. Selon la bande-son, les sources sonores pouvaient être plus ou moins cohérentes en azimut et en profondeur avec la position de leur source visuelle respective sur l'écran (la cohérence en élévation avait été mise de côté au vu des résultats de l'expérience 4). Les résultats ont montré que la cohérence en azimut pouvait améliorer significativement l'adéquation du son à l'image. En profondeur, une amélioration a pu être constatée, mais seulement pour une séquence. Par contre, la stéréoscopie n'a eu aucune influence sur les jugements des sujets, en accord avec les résultats des premières expériences sur la perception des sons d'ambiance.

Abstract

Few psychoacoustic studies have been carried out about the influence of stereoscopy on the sound mixing of movies. Yet very different opinions can be found in the cinema industry and in scientific papers. Some argue that sound needs to be mixed differently for stereoscopic movies while others pretend that this influence is negligible.

A first set of experiments was conducted, which focused on the perception of ambiance. Eight sequences - in their stereoscopic (s-3D) and non-stereoscopic (2D) versions, with several different sound mixes - were presented to subjects. For each presentation, subjects had to judge to what extent the mix sounded frontal or « surround. » The goal was to verify whether stereoscopy had an influence on the perception of the front/rear balance of ambiance. Results showed that this influence was weak, which was consistent with a preliminary experiment conducted in a mixing auditorium where subjects had to mix the front/rear balance of several sequences themselves.

A third experiment was conducted to verify if the sequences that were significantly influenced by stereoscopy corresponded to the sequences whose differences between s-3D and 2D versions were the most important in terms of perceived visual depth, yet no correlation could be found.

Studies were then conducted on the perception of sound objects such as dialogs or on-screen effects. A fourth experiment focused on ventriloquism in elevation: when presented with a spatially discordant auditory-visual stimulus, subjects sometimes perceive the sound and the visual stimuli as coming from the same location. Such a phenomenon is often referred to as *ventriloquism*, because it evokes the illusion created by a ventriloquist when his voice seems to emanate from his puppet rather than from his mouth. While this effect has been extensively examined in the horizontal plane and to a lesser extent in distance, few psychoacoustic studies have focused on elevation. In this experiment, sequences of a man talking were presented to subjects. His voice could be reproduced on different loudspeakers, which created disparities in both azimuth and elevation between the sound and the visual stimuli. For each presentation, subjects had to indicate whether or not the voice seemed to emanate from the mouth of the actor. Ventriloquism was found to be highly effective in elevation, which suggests that audiovisual coherence in elevation might be unnecessary in theaters.

In a fifth experiment, the influence of stereoscopy on subjects' expectations regarding audiovisual spatial coherence was investigated. In theaters, sound objects are most of the time reproduced on the central loudspeaker, regardless of the position on screen of their related visual sources. Yet, some sound engineers and researchers have suggested that a spatial audiovisual coherence could improve the experience of the audience significantly, especially for s-3D movies. In this experiment, subjects were asked to judge the suitability of several soundtracks for 8 sequences, which were presented in their s-3D and 2D versions. Depending on the soundtrack, sound sources could be more or less coherent in azimuth and in depth to their related visual sources (coherence in elevation was not investigated because of the results of the fourth experiment). Results showed that sound suitability could be significantly improved for most of the sequences when coherence in azimuth was achieved. In depth, improvement was only observed with one sequence. However, no significant effect of stereoscopy on subjects' judgments could be found, which is consistent with the previous experiments on the perception of ambiance.