Generation of Vibrotactile Feedback for Enhancing Collision Awareness and Quality of Experience in First-person View Videos

著者	Gongora Flores Daniel Marcell
学位授与機関	Tohoku University
学位授与番号	11301甲第18949号
URL	http://hdl.handle.net/10097/00130226

氏　　　　　　　　名	ゴンゴラ　フロレス　ダニエル　マルセル
研究科，専攻の名称	東北大学大学院情報科学研究科（博士課程）応用情報科学専攻
学 位 論 文 題 目	Generation of Vibrotactile Feedback for Enhancing Collision Awareness and Quality of Experience in First-person View Videos (一人称視点ビデオにおける接触検知と体験の質を高める触覚フィードバックの生成)
論 文 審 査 委 員	主査 東北大学准教授　昆陽　雅司　東北大学教　授　田所　諭 東北大学教　授　岡谷　貴之　東北大学教　授　平田　泰久 （工学研究科）

論文内容要約

Our interaction with information is largely mediated by graphical user interfaces, but in certain tasks our sense of sight alone is either not enough or not suitable for understanding the information shown on the screen. For instance, two issues that arise during our interaction with First-Person View (FPV) videos are an unreliable capacity to detect short events, and a poor quality of experience. In the first case, a visual-only representation of events is insufficient because it does not guarantee an immediate response. In the second case, a visual-only representation is not adequate because it might prompt feelings of discomfort, especially when the video depicts intense camera trembling. When a visual representation is insufficient or inadequate in the real world, we recourse to other senses. However, traditional visual-only information displays do not afford such natural mechanism to deal with the perceptual limitations of the sense of sight. The consequences of a visual-only representation of information can be as trivial as preventing people from watching videos depicting extreme sports, or as serious as limiting the success in search-and-rescue missions supported by robots where a FPV video feed is used to control the robot.

Vibrotactile feedback has shown great potential supporting visual tasks but its adoption remains limited. The benefits of adding vibrotactile feedback to a visual information display range from faster reaction times to visual alerts, faster completion times in visual search tasks, better performance in tasks where detecting transitions is important, etc. Nevertheless, vibrotactile feedback is yet to become widely adopted. One of the problems interfering with a more prevalent adoption of vibrotactile feedback has to do with the necessity of specialized knowledge and tools for its generation. For instance, people considering the addition of vibrotactile feedback to an information display might be dissuaded by unavoidable considerations of factors like the number and placement of vibrotactile actuators, signal processing that accounts for the perceptual characteristics of the mechanoreceptors involved in the perception of vibrotactile stimuli, etc. Furthermore, although there is no lack of applications showcasing the potential of vibrotactile feedback, the benefits of incorporating vibrations to information displays are not easily capitalized due to the dependence of existing implementations on actuators arranged in ways not readily available. Potential adopters of vibrotactile feedback are hence required to fully commit with a given approach instead of having the freedom to assess its benefits first. Therefore, methods for automatically generating vibrotactile feedback that do not rely on elaborated arrangement of actuators would increase the adoption of vibrotactile feedback.

The goal of this dissertation is to investigate vibrotactile methods to enhance event-awareness and Quality of Experience (QoE) in FPV videos with intense camera trembling. In recent years FPV videos have been gaining popularity but limitations regarding event-detection and QoE diminish the information and entertaining value viewers get from them. For enhancing event-detection, we consider two tasks. First, we focus on a collision detection task during teleoperation of a mobile robot. Then, we focus on a collision detection task during fast video playback of FPV videos obtained during teleoperation. In both tasks, collisions are hard to perceive visually for two main reasons. First, collisions have low visual salience due to their short duration. Second, viewers are prone to inattentional blindness because they are engaged in either controlling the robot or controlling the video playback interface. As for enhancing QoE, we focus on a task consisting of watching action videos recorded from a FPV perspective. In this task, intense camera trembling in the videos is likely to induce discomfort or visual fatigue after prolonged exposure.

To improve collision awareness during teleoperation, we propose a vibrotactile stimulation method to represent frontal collisions that is based on the way people perceive impacts on a bar held with both hands. This method not only represents the occurrence of a collision but also its approximate impact point. To create vibrations that are informative but at the same time easy to understand, we first investigate bimanual cues to impact localization using vibration measurements of impacts on three types of bars. Then, to apply these results to a teleoperation interface, we obtained a psychophysical function that relates impact points to the vibration parameters of two vibrotactile actuators worn on the forearms or held with both hands. Finally, we evaluated our method in a simplified teleoperation task using a differential drive mobile robot equipped with a high speed tactile sensor on the front bumper. We observed that to estimate the impact point in a bimanual impact localization task people relied on amplitude and duration differences of the impact vibrations delivered to their hands. We also observed in a pilot study that participants completed the teleoperation task in less time when vibrotactile feedback was available.

To improve detection of short events during fast video playback, we propose a haptic exploration method considering that event-related information in vibrotactile feedback can be preserved during fast playback using Time Scale Modification (TSM) methods developed for audio. This method enables an interaction with videos inspired by the way we use our hands to explore a texture with strokes of varying speed, where changes to playback speed correspond to changes in stroke speed and where feeling collisions is analogous to feeling a ridge on the surface under exploration. We evaluate our proposal in two collision detection experiments using First-Person View (FPV) videos. In the first experiment, viewers watched at a fixed playback speed, i.e., 1× or 2×, videos recorded with a camera mounted on a platform cart. In this experiment, event-related vibrations were measured at the back of the camera mounted on the cart. In the second experiment, viewers used a media controller to adjust the playback speed in videos simulating an exploration with a mobile robot. In this experiment, event-related vibrations were generated using as reference the measurements obtained in the first experiment. We show that a haptic exploration improves collision awareness under either constant or adjustable playback speed. In both experiments, the number of collisions reported without vibrotactile feedback deviated the greatest from the actual number of collisions in a video. Moreover, collision detection performance with vibrations time-scaled without Time Scale Modification (TSM) methods was not significantly different from performance without vibrations. These results demonstrate the feasibility of a haptic exploration approach based on event-related vibrations, show the potential of a haptic exploration to improve event detection performance during fast playback.

Finally, to improve the QoE in FPV videos, we propose a vibrotactile rendering method for the motion of the camera in FPV videos. This method enables people to feel the movement of the camera with both hands, where panning movements on the horizontal axis are experienced as vibrations that move from hand to hand, and sudden vertical displacements of the camera are experienced as transient exponentially-decaying vibrations on both hands. First, we evaluate the effects on QoE of vibrations generated by our method. Then, we evaluate the potential to preserve the feeling of motion in stabilized FPV videos with vibrations generated by our method. In the first experiment, viewers reported the quality of their experience on four dimensions: realism, comfort, sensory, and satisfaction. In the second experiment, viewers reported their perceived motion intensity, synchrony between video and vibrations, comfort and satisfaction. We observed favorable effects of vibrations generated by our method on the perceived realism and satisfaction associated with the experience of watching a video. We also observed that although vibrations generated by our method did not preserve the feeling of motion intensity in stabilized FPV videos, the experience of watching a video with them was mostly described as comfortable and associated with feelings of satisfaction. Moreover, video stabilization did not affect the perceived synchrony between video and vibrations generated from the original unstable video.

Overall, the main contributions of this dissertation are threefold. First, we propose and validate a novel method to represent frontal collisions of a mobile robot using two vibrators. Second, we propose and validate a novel method to detect short events during fast video playback using vibrations. Third, we propose and validate a novel method to represent camera motion in FPV videos using two vibrators. These contributions address issues related to poor event-awareness and poor QoE when watching FPV videos. Research in these areas is justified given the prevalent use of visual feedback and visual media as sources of information and entertainment.