A Framework for Enhancing Video Viewing Experience with Haptic Effects of Motion

Fabien Danieau, Julien Fleureau, Audrey Cabec, Paul Kerbiriou, Philippe Guillotel, Nicolas Mollet, Marc Christie, Anatole Lécuyer

▶ To cite this version:

A Framework for Enhancing Video Viewing Experience
with Haptic Effects of Motion

Fabien Danieau *
Technicolor / INRIA

Julien Fleureau
Technicolor

Audrey Cabec
Technicolor

Paul Kerbiriou
Technicolor

Philippe Guillotel
Technicolor

Nicolas Mollet
Technicolor

Marc Christie
IRISA

Anatole Lécuyer
INRIA

Abstract

This work aims at enhancing a classical video viewing experience by introducing realistic haptic feelings in a consumer environment. More precisely, a complete framework to both produce and render the motion embedded in an audiovisual content is proposed to enhance a natural movie viewing session. We especially consider the case of a first-person point of view audiovisual content and we propose a general workflow to address this problem. This latter includes a novel approach to both capture the motion and video of the scene of interest, together with a haptic rendering system for generating a sensation of motion. A complete methodology to evaluate the relevance of our framework is finally proposed and demonstrates the interest of our approach.

Index Terms:

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Video; H.5.2 [Information Interfaces and Presentation]: Multimedia Information Systems—Haptic I/O;

1 Introduction

New technology developments allow the creation of more and more immersive multimedia systems. 3D images and sound spatialization are now present in the end-user living space. But these systems are still limited to the stimulation of two senses, sight and hearing, while researches in virtual reality have shown that haptic perception seems to be strongly connected to the feeling of immersion [12].

In line with these works, we focus here on the way to enhance a natural video viewing experience with "realistic" haptic effects in a consumer environment. More precisely, our motivation is to develop a framework to make the user feel the **motion** embedded in the multimedia (Audio/Video) content he is watching. In this work, we target an automatic way to both produce and render the motion effects embedded in an audiovisual content. We especially consider the case of a first-person point of view A/V content, for which we expect the viewer to feel what the main actor is currently feeling in terms of motion. In this context, a comprehensive framework is presented and includes:

1. An approach to both capture motion and video of the scene of interest,
2. The method necessary to send and transform the original motion information to the back-end haptic device,
3. A haptic rendering scheme for generating a motion effect on a force-feedback device,

---

*e-mail: fabien.danieau@technicolor.com

4. A methodology to evaluate the interest and the relevance of such a framework.

The remainder of this paper is as follows. Section 2 describes existing works proposing to produce or render motion effects. Section 3 presents the framework including a description of the main components of the proposed workflow. Section 4 details the methodology adopted to assess the proposed system as well as the obtained results. Finally section 5 provides conclusions and perspectives.

2 Related Work

2.1 Haptic effects for audiovisual contents

The production and the rendering of haptic effects for audiovisual contents are two typical issues identified in the pioneer theoretical work of O'Modhrain and Oakley [9].

The first issue deals with the content creation, i.e. how the haptic effects can be **produced** in order to be added to the audiovisual stream. In our context haptic effects are "motion effects". According to O'Modhrain and Oakley there are two ways to create such effects: the off-line creation and the real-time generation. In the first case haptic effects are synthesized and the editor of the content manually adds effects to the media. In the second case the effects are directly captured from physical sensors using specific device.

The second issue refers to the visualization of the content, more precisely to the **rendering** of the haptic cues. The haptic feedback should be rendered by a technology able to produce a wide range of haptic sensations. Moreover several constraints might appear if the content has to be displayed in a user's living space, potentially shared.

Only a few contributions exploring the interest of haptic feedback for audiovisual contents have been reported in the literature. Most of them rely on the use of vibrotactile devices (Rahman et al. [11], Kim et al. [8]) or force-feedback devices (Gaw et al [5], Cha et al. [2]). But the haptic effects proposed by these contributions are relatively simple : artificial vibrations patterns or abstract force-feedback. To our best knowledge, no haptic effects of motion have been introduced.

2.2 Simulation of motion

In the context of motion rendering, motion simulators are well-known devices designed to make the user feel motion. They are intensively used as driving or flight simulators for learning purposes or in amusement parks. Most of them are based on the Stewart's platform [4]. It is a 6-DOF platform moving thanks to 6 hydraulic cylinders. Motion platforms are very immersive but they remain expensive for end-user customers and are not designed to be integrated in a user's living space. In a less invasive way, the sensation of motion can be also induced by a force-feedback device. Ouarti et al. [10] apply a force on the user hand and the system is expected to generate an illusion of motion with force-feedback.

While the interface is pulling the hand, the user feels moving forward. In our work we consider using a force-feedback device to render the effect of motion instead of a motion simulator. Such a technique takes advantage of its low cost and of its convenient size to be compatible with consumer applications.

In the context of motion capture, motion effects can be produced thanks to external sensors such as i) Inertial Measurement Units (see [1] or [15] for examples in the context of radio-controlled cars and actor modeling respectively) or ii) the camera used to record the scene combined with motion extraction techniques (see Hu et al. [6] for a survey). Extraction algorithms can be helped by adding markers into the filmed scene (Sigal et al. [14] relied on this technique to perform motion capture) or by using a camera enhanced with infra-red capabilities (Microsoft Kinect device for instance). Thus several motion capture techniques exist but they are not designed to enhance a video viewing session. They target other applications such as human behavior modeling or human computer interaction.

3 Outline of the system

The system we propose is a comprehensive framework designed from the production of motion effects to their rendering on dedicated devices. Three main steps compose this framework as depicted in Figure 1.

Figure 1: Workflow Overview. Data are simultaneously captured by a camera and a motion sensor. Then motion data are converted into a signal suitable for the dedicated haptic renderer. Finally both video and motion are rendered simultaneously.

3.1 Motion capture with physical sensors

The first step consists in i) capturing the motion effects for the renderer and ii) recording the audiovisual content that should be displayed simultaneously with the renderer. In the context of our work, a combined system making use of, on one hand, an Inertial Measurement Units (IMU) and, on the other hand, of a High Definition (HD) camera dedicated to sportive activities has to be designed.

The IMU we chose is the Ultimate IMU board which combines an ADXL345 accelerometer, an ITG-3200 gyroscope and a HMC-5843 electronic compass (cf. Figure 2-Left,B). The first component records the 3-axis accelerations of the board, the second one quantifies the rotation of the board around its 3 axes and the last component allows a geocentric orientation by giving an estimation of the local magnetic field. An additional micro-SD memory card may be embedded on the board and allows the recording of the three raw signals. A dedicated middleware has been developed and uploaded to the Ultimate IMU to set the recording process to 30Hz.

Complementary, a Camsports HDS-720p (see Figure 2-Left,A) was selected to record the scene corresponding to the current point of view of the actor. The camera is a HD bullet camera. It uses a 120 degrees wide-angled lens and integrates a built-in 4GB memory chipset. The spatial resolution of this device is 1280 x 720p at a frequency of 30fps. It is water-proof and is able to handle harsh environments. It finally integrates a mono-channel microphone.
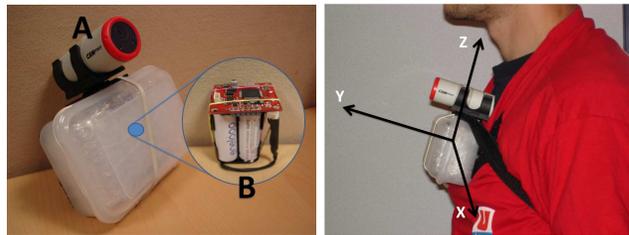

Figure 2: Overview of the device capturing both video and motion. Left: The integrated prototype is composed by a (A) Camsports HDS-720p and a (B) Ultimate IMU board with its battery. Right: The prototype is fixed on an actor's chest and records his motion on 3 axes.

A complete integrated prototype combining the IMU, its battery and the camera has been developed. As the system is designed to be fixed on an actor (first-person point of view recording), it is enough robust to resist to different conditions of recording (cf. Figure 2-Right).

To synchronise the IMU and the camera, which are obviously independent, and do not offer possibilities of external synchronisation, a mechanical trick is used (very similar to the A/V synchronisation techniques traditionally used in movie making). Before each record, three little pats are given on the prototype which cause a fast and big peak in both the acceleration signals of the IMU and the audio stream of the camera. Basic signal processing techniques are then used to make those peaks match in both signals (variance-based threshold).

3.2 Processing of the captured motion signals

Once the different motion information have been recorded, they have to be processed to be compatible with the input of the haptic device that is used. As stated previously, Ouarti et al. [10] relied on a force-feedback device to make the user feel a sensation of self-motion. A force in the user hand was applied while the subject was watching a visual stimulus. The force (orientation and direction) is highly correlated to the motion embedded in the visual content and creates an illusion of self-motion even though the application point (hand) is different from the stimulating point where the accelerations have been recorded (chest).

In our context, we propose to replace the single acceleration of the approach of Ouarti et al. [10] by using the three raw accelerations $\mathbf{a}[k] = \{a_x[k], a_y[k], a_z[k]\}^T$ recorded during the capture step at each sample k to modulate the 3-axis effort applied to the subject's hand. The methodology described here is generic and may be applied to other kinds of rendering devices. Though this approach is quite direct and requires very few processing, several points require specific attention (see Figure 3).

Figure 3: Processing Overview. The gravity is removed from the accelerations motion signals captured. Optionally extra processing can be performed such as low-pass filtering.

3.2.1 Gravity removal

The main processing to apply is linked to the gravitational component **g** included in the raw acceleration. This latter is quite important regarding the other external sources of acceleration and can mask some useful information needed to render a motion feeling. Our empirical observations showed that removing this specific contribution enhance the user experience. A two-step methodology is therefore applied to remove this component from the original signal. In a first step, the board orientation is estimated using the approach described by Sabatini [13]. This latter especially combines the use of i) a quaternion-based representation of the board attitude and ii) a dedicated extended Kalman filter to estimate the board orientation by fusioning the information coming from the three sensors (gyroscope, accelerometer, magnetometer). This operation allows to estimate the direction of the gravity (vertical) $\mathbf{n}_g[k]$ in the accelerometer frame (frame **A**) at each time sample k. The raw acceleration vector is therefore updated by removing the quantity, $\|\mathbf{g}\|\mathbf{n}_g[k]$, from each sample a$[k]$.

The processing output signal $\mathbf{p}[k] = \{p_x[k], p_y[k], p_z[k]\}^T$ may be formalized, at each time sample k, by:

$$\mathbf{p}[k] = \mathbf{a}[k] - \|\mathbf{g}\|\mathbf{n}_g[k] \qquad (1)$$

3.2.2 Filtering and Additional effects

A filtering step is necessary to reduce the noise of the original signal. For practical reasons, the filtering is actually performed on the IMU. More precisely, the 3 sensors data were natively sampled at 200Hz but due to limitations with the writing speed on the embedded micro-SD, samples were averaged and down-sampled at 30Hz. This averaging step contributes to a low-pass filtering of the raw signal.

Extra operations enhancing the signal for a better rendering for the end-user may be optionally addressed. They may be simple operations to remove artifacts or artificial modulation (reduction or amplification) of some parts of interest in the signal **a** to underline specific haptic events.

3.3 Haptic rendering of the motion effect

An open-loop rendering system was introduced to display the force-feedback. This way the signal **p** was simply rendered as a force vector **F**, defined as $\mathbf{F}[k] = \{F_x[k], F_y[k], F_z[k]\}^T$. However a few transformations of the signal **p** should be performed in order to render a force suitable for the haptic device (Figure 4).
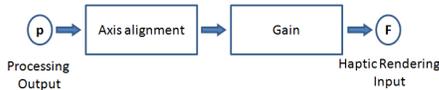

Figure 4: Pre-Rendering Overview. The signal from the processing step is aligned to the axes of the haptic device and then amplified.

The processed signal **p** is sent to a haptic rendering algorithm controlling a haptic device which is, in our framework, a Novint Falcon device. This latter is a 3-DOF force-feedback device, able to apply a force along three axes.

3.3.1 Axis alignment

To be rendered on the haptic device, an axis permutation of the signal **p** has to be performed to align the axes of the accelerometer (frame **A**) with the axis of the device (frame **D**). The associated permutation matrix is termed $\mathcal{M}_\mathbf{D}^\mathbf{A}$.

In our context, the matrix simply switches the y and z-axis of **A** in **D**. A complementary step reverses the z-axis as the force-feedback device is placed in front of the user and it is supposed to pull the user's hand when the recorded acceleration is positive on the z-axis.

3.3.2 Gain

Then a scaling of the raw data is necessary to adapt the amplitude of the signal **p** to the renderer input range. The scaling factors s_x, s_y and s_z for each axis are assumed to be constant (independent of the time sample) and empirically set according to experimental feedbacks. The associated diagonal scaling matrix is termed $\mathbf{diag}(s_x, s_y, s_z)$ and in the context of our work $s_x = s_y = s_z$.

The force rendered by the haptic device may be finally formalized by:

$$\mathbf{F}[k] = \mathbf{diag}(s_x, s_y, s_z)\mathcal{M}_\mathbf{D}^\mathbf{A}(\mathbf{p}[k]) \qquad (2)$$

3.4 Multimedia player with haptic effects

Once the force **F** is computed, it has to be rendered by a haptic device. A force **F** is computed for each IMU sample and oversampled (point sampling, i.e. piecewise constant interpolation) to meet the requirements of the 1kHz haptic rendering loop frequency.

The rendering algorithm was integrated in a home-made multimedia player allowing the haptic rendering on a force-feedback device and the A/V rendering in a synchronized way. The HAPI [1] was used to control the Novint Falcon. This high-level C++ library allows to simply create force vectors which are then rendered on a force-feedback device.

The framework presented here allows to produce and render realistic haptic effects of motion. The rendering stage relies on a force-feedback device to generate a sensation of motion. Accelerations of the camera are rendered on this device. But the framework is not limited to one kind of haptic device. The positions of the IMU, and so of the prototype, are computed in the processing step in order to remove the gravity. These data could be used to render motion effect on a classical motion simulator which will reproduce the position of the camera.

4 User Evaluation

A dedicated experimental protocol has been developed to assess the impact of the haptic feedback on the user's Quality of Experience (QoE). QoE relates to the subjective user experience with a service or an application [7]. The protocol designed to evaluate the QoE is presented hereafter as well as the associated results.

4.1 Experimental Setup

4.1.1 Capturing test sequences

Our motion capture prototype was used to create several samples of audiovisual contents enriched with haptic data. We identified 4 scenarios to represent different kinds of motion feelings (Figure 5). The prototype was placed on an actor's chest and we obtained the following videos sequences as well as the corresponding haptic data:

- **Bike**. The objective of this scenario is to capture low-amplitude movements. The actor is performing outdoor cycling and a succession of vertical movements with small amplitude are captured. (duration 61s).

- **Horse**. In this case the actor is riding a galloping horse and feels recurrent top-down movements. High-amplitude vertical movements are captured. (duration 60s).
- **Car turning**. In this scenario, the actor is inside a car engaged in a roundabout. The centrifugal force makes him feel pushed on a side. The captured motion is felt as strong and long. (duration 45s).
- **Car Braking**. This last scenario aims to capture a strong punctual movement. The actor is in a car strongly braking and he feels a strong force pushing him forward during few seconds. (duration 75s).

Duration of each sequence is around one minute.

Figure 5: Tests Scenarios. Top Left - Outdoor cycling. Top Right - Horse riding. Bottom Left - Car engaged in a roundabout. Bottom Right - Car strongly braking.

4.1.2 Haptic setup

The playback system was composed of a 15" laptop and a Novint Falcon device (Figure 6). The user is comfortably seated in front of the computer where one of the captured video sequence is displayed and experiences haptic feedback by holding the Novint Falcon device in his dominant hand. Our home-made Haptic/Audio/Video Player was used to play the video in fullscreen mode.

Figure 6: A participant experiences an audiovisual content enriched with haptic feedback.

4.2 Experimental Protocol

4.2.1 Variables

In order to evaluate the user's QoE for each sequence we defined three types of haptic feedback to be rendered with the video:

- **Realistic Feedback**. The captured haptic feedback, consistent with the sequences.
- **No Feedback**. Only the audiovisual content is displayed. The goal of this condition is to measure the QoE of a classical A/V content. This will be used as a reference to evaluate the interest of a haptic feedback for a video.
- **Random Feedback**. A random haptic feedback made of a low-pass filtered white noise (cutoff frequency $F_c = 0.5Hz$) of the same length and amplitude of the consistent haptic feedback. This feedback is not consistent with the video and will be used to evaluate the interest of providing a realistic haptic feedback.

Combining the whole set of possibilities we get **12 conditions** (4 videos sequences x 3 haptic feedbacks) which were tested in each experiment in order to evaluate the QoE, our independent variable. These conditions will be presented in a random order to the participants.

4.2.2 Measures

A questionnaire was designed to evaluate the QoE of a video enriched with haptic feedback. It was built around the **Presence** [18] and **Usability** [16] concepts.

Presence aims at measuring how much the user feels being *physically* situated in a virtual environment. Witmer and Singer [18] identified four factors to determine the presence: Control, Sensory, Realism and Distraction. "Control" determines how much the user can control and modify objects within the virtual environment. "Sensory" characterizes how each sensory modality is solicited during the interaction. "Realism" describes how much the environment is realistic and consistent with user's representation of the real world. "Distraction" identifies how much the user is disturbed by the apparatus used to create the virtual world. From this definition we focused on two factors: *Realism* and *Sensory*. As the user is passive with our system, Control factor was not relevant here. Moreover we did not measure Distraction in our QoE questionnaire but this aspect was interesting and was evaluated in a second questionnaire (post-test questionnaire).

Usability is defined by the norm ISO 9241-11 and aims at measuring how easy a system is to use. Three factors composed this concept: Efficiency, Effectiveness and Satisfaction. This latter measures how well the user enjoyed the system. "Effectiveness" means how well a user can perform a task while "Efficiency" indicates how much efforts are required. These two factors were not totally suitable for our system in the sense that it was not designed to perform a task. We preferred to use the term of *Comfort* to measure how well was the system to provide feedback. *Satisfaction* was however fully relevant in our situation.

Hence, the QoE of our system was evaluated by 4 items : Realism, Sensory, Comfort and Satisfaction (cf. Table 1). We defined only one question by item supposed to be rated on a five-point Likert-scale. The QoE is computed by the sum of these 4 items. This way the QoE questionnaire is easy to fill in and can be submitted for each condition.

A second questionnaire was designed in order to collect more information about user's feelings. It was composed by

Factor	Question
Realism	How much did your experiences in the virtual environment seem consistent with your real-world experiences?
Sensory	How much did the haptic feedback improve the interaction?
Comfort	How much was the system comfortable?
Satisfaction	How much was the system pleasant to use?

Table 1: QoE Qestionnaire. Each question is rated on a 5-point Likert-scale from 1 (Not at all) to 5 (Totally)

several open-ended questions supposed to be submitted at the end of the experiment. The participant was asked to label each haptic feedback, to indicate which one seemed the more realistic and why. To answer these questions he had the possibility to replay every sequences classified into 3 groups : Feedback 1, Feedback 2 and Feedback 3. He was also asked to tell how the system was tiring and comfortable to use, and finally to imagine several applications.

4.2.3 Procedure

15 participants have taken part to the experiment. They were aged from 21 to 59 (M=27.8 SD=9.7), 9 were Male, 1 participant was left-handed, 8 never used a Novint Falcon device. The whole experiment lasted from 30 to 40 minutes. The procedure for each user was organized as follows:

1. Fill in a consent form.
2. Fill in an information sheet.
3. Demonstration of the falcon and its capacities (forces it can provide). This step aims of reducing the "surprise effect" for novice users.
4. Presentation of the 12 conditions in a random order. For each one:
   - 4a. Experience the multimedia content (video sequence + haptic feedback),
   - 4b. Answer to the QoE questionnaire.
5. Answer the post-test questionnaire. Possibility to retry every sequences.

4.3 Results

The data collected were 4 notes (Realism, Sensory, Comfort and Satisfaction; from 1 to 5) for each condition per participant. The sum of these notes gives the QoE per conditions per participants (Figure 7). The QoE per feedback per participant was also computed. This metrics is termed QoE_{All}. The normality of the distributions was tested with the Shapiro-Wilk test and was rejected most of the time. Hence non-parametric tests were used to analyze the results presented in this section.

We first found that the QoE for each sequence follows the same pattern. The QoE_{All} of the "Random Feedback" condition ($M = 10.2\ SD = 1.6$) is higher than the QoE for the "No Feedback" condition ($M = 7.5\ SD = 2.1$), and the QoE of the "Realistic Feedback" ($M = 15.3\ SD = 2.6$) is higher than the random one. Figure 7 depicts this pattern. This result is significant according to the Friedman Anova ($\chi^2 = 24.71,\ df = 2,\ p < 0.05$).

Figure 8 shows the mean score for each item of the QoE_{All} for the three feedback conditions. The more realistic the feedback is the higher are the Realism, Sensory and Satisfaction items: Realism $M_{None} = 1.3\ M_{Random} = 1.9\ M_{Realistic} = 3.8$, Sensory $M_{None} = 1.1\ M_{Random} = 2.4\ M_{Realistic} = 3.9$ and Satisfaction $M_{None} = 2.1\ M_{Random} = $
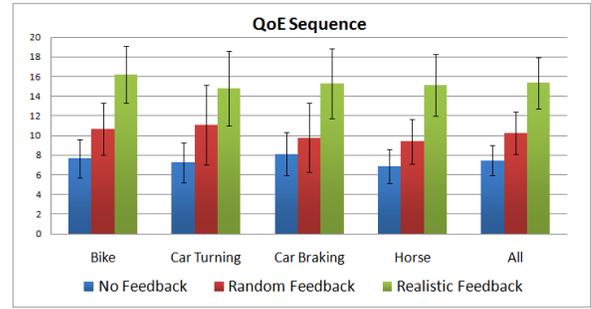

Figure 7: QoE of each sequence and haptic feedback. For each sequences, participants found that a realistic haptic feedback improves the experience. Interestingly a random feedback was more appreciated that no feedback.

$2.8\ M_{Realistic} = 4.1$. These results are also significant (Friedman Anova: $\chi^2 = 24.0.3,\ df = 2,\ p_{Realism} < 0.05$; $\chi^2 = 26.68,\ df = 2,\ p_{Sensory} < 0.05$; $\chi^2 = 23.05,\ df = 2,\ p_{Satisfaction} < 0.05$). However Comfort appears to be relatively stable all along the experiment ($M_{None} = 2.9\ M_{Random} = 3.2\ M_{Realistic} = 3.6$). But this result is less significant (Friedman Anova, $\chi^2 = 8.79,\ df = 2,\ p = 0.012$).
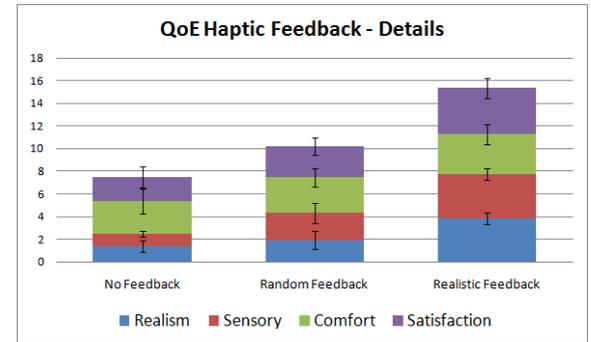

Figure 8: QoE of each haptic feedback and details of the components. The Comfort component of the QoE remains the same whatever the feedback perceived. However the three others increase with no feedback, random feedback and realistic feedback respectively.

We also observed that the QoE_{All} for the realistic haptic feedback remains the same for those who never used a Novint Falcon ($M = 15.5\ SD = 2.9$) and for those who did ($M = 15.25\ SD = 2.5$). The expertise of the participant do not affect the result significantly (Wilcoxon Mann-Whitney test, $p < 0.05$).

4.4 Discussion

The main result of this study is that QoE increases with haptic feedback and more particularly with haptic feedback consistent with audiovisual content. Moreover participants' expertise with the Novint Falcon do not affect the QoE. This observation let us think that our main result is not due to a "surprise effect". However the low score obtained by sequences with no feedback can be in part due to our experimental protocol. Whatever the condition, participants were asked to hold the Novint Falcon device in their dominant hand. Thus they might have been frustrated when they felt no haptic feedback. Obviously if there is a haptic device people are expecting haptic feedback.

We also observed that haptic feedback may change user's

perception of the audiovisual content, especially if the meaning of the video is ambiguous. For instance one cannot see a bike in the bike sequence although a head of a horse is visible in the horse sequence as well as a part of a car in the two car sequences. During the experiment a participant thought that the bike sequence represented a buggy riding video because he felt that the haptic feedback (realistic feedback condition) was closed to his own buggy driving experience. Thus it appears that users build a mental representation of the multimedia content consistent with their own experience, and it is interesting to see how haptic feedback can influence this representation when audiovisual content is ambiguous.

Another interesting behavior was observed while participant experienced video enriched with random feedback. Most of them tried to find a meaning for this haptic feedback, consistent with their own personal experience. This observation may explain higher QoE for random feedback than for no feedback. The phenomenon was particularly highlighted in the Car Turning and Car Braking conditions. Several participants supposed that the haptic feedback was mapped to the gear shift of the car. This can also explain why QoE for Random Feedback in these two conditions is better than in Bike and Horse conditions.

Finally participants reported in the post-test questionnaire to feel comfortable all along the experiment although the position of the arm and the hand-grip were reported as quite uncomfortable. This setup is obviously not suitable for watching a 2-hours movie. Of course, the limitations mainly due to the haptic rendering system and to the screen size, should be improved to reduce the user's fatigue in a out-of-the-lab context.

5 Conclusion and Perspectives

In this paper we presented a comprehensive framework to add realistic haptic effects of motion to a video. A novel approach to capture both audiovisual content and motion of the camera was first detailed. Then an original way to render this multimedia content was described. Our playback system relies on a force-feedback device to make the user feel the captured motion. Finally we presented a method to evaluate users' Quality of Experience with our system. Results show that the user experience increases with a realistic haptic feedback.

The work presented here focuses on the production and on the rendering of haptic effects. A third issue important to consider would be the distribution of the audiovisual content enriched with haptic effects. Cha et al. [3] introduced the notion of haptic broadcasting which correspond to the techniques to transmit haptic effects synchronized to the other components of the media. In a similar way, Walt [17] described a data format to formalize haptic effects and to synchronize them to an audiovisual content. These techniques would be an interesting extension for this framework, especially in a consumer context where users used to watch videos through streaming platforms.

The first user study yielded promising results. Therefore, it would be first interesting to conduct a large-scale study in order to characterize more precisely the user's feelings and attitude with the system as well as the design of a more ecological setup. Moreover research efforts are necessary to determine when the user perceives a haptic feedback as consistent or not with an audiovisual content. This will help to finely design haptic effects necessary to trigger an immersion feeling. Second, sequences representing a diversity movements have also to be captured. More fundamentally,

it is important to understand how users feel a sensation of motion through a force-feedback device. Finally it would also be interesting to evaluate the perception of effects with third-person point of view records.

To conclude, the framework we proposed in this paper allows to create more immersive audiovisual content by involving our sense of touch. This brings a new way to experience multimedia content and can enhance many viewing contexts such as movies, extreme sports videos or video games.

References

[1] A. Brady, B. MacDonald, I. Oakley, S. Hughes, and S. O'Modhrain. Relay: a futuristic interface for remote driving. In *Eurohaptics*, 2002.

[2] J. Cha, M. Eid, and A. E. Saddik. Touchable 3D video system. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 5(4):1–25, Oct. 2009.

[3] J. Cha, Y.-S. Ho, Y. Kim, J. Ryu, and I. Oakley. A Framework for Haptic Broadcasting. *IEEE Multimedia*, 16(3):16–27, July 2009.

[4] B. Dasgupta. The Stewart platform manipulator: a review. *Mechanism and Machine Theory*, 35(1):15–40, Jan. 2000.

[5] D. Gaw, D. Morris, and K. Salisbury. Haptically Annotated Movies: Reaching Out and Touching the Silver Screen. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 287–288. IEEE, 2006.

[6] W. Hu, T. Tan, L. Wang, and S. Maybank. A Survey on Visual Surveillance of Object Motion and Behaviors. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, 34(3):334–352, Aug. 2004.

[7] K. Kilkki. Quality of experience in communications ecosystem. *Journal of universal computer science*, 14(5):615–624, 2008.

[8] Y. Kim, J. Cha, J. Ryu, and I. Oakley. A Tactile Glove Design and Authoring System for Immersive Multimedia. *IEEE Multimedia*, pages 2–12, 2010.

[9] S. O'Modhrain and I. Oakley. Touch TV: Adding feeling to broadcast media. In *European Conference on Interactive Television: from Viewers to Actors*, 2003.

[10] N. Ouarti, A. Lécuyer, and A. Berthoz. Method for simulating specific movements by haptic feedback, and device implementing the method, September 2009. French Patent $N°$ 09 56406.

[11] M. A. Rahman, A. Alkhaldi, and J. Cha. Adding haptic feature to YouTube. In *ACM International Conference on Multimedia*, pages 1643–1646, 2010.

[12] M. Reiner. The Role of Haptics in Immersive Telecommunication Environments. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(3):392–401, Mar. 2004.

[13] A. M. Sabatini. Quaternion-Based Extended Kalman Filter for Determining Orientation by Inertial and Magnetic Sensing. *IEEE Transactions on Biomedical Engineering*, 53(7):1346–1356, 2006.

[14] L. Sigal and M. J. Black. Humaneva: Synchronized video and motion capture dataset for evaluation of articulated human motion. Technical Report CS-06-08, Brown Univertsity TR, 2006.

[15] R. Slyper and J. K. Hodgins. Action capture with accelerometers. In *Eurographics/ ACM SIGGRAPH Symposium on Computer Animation*, 2008.

[16] T. Tullis and W. Albert. *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Morgan Kaufmann, 2008.

[17] M. Waltl. *Enriching Multimedia with Sensory Effects: Annotation and Simulation Tools for the Representation of Sensory Effects*. VDM Verlag Saarbrücken, Germany, 2010.

[18] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments : A presence questionnaire. *Presence*, pages 225–240, 1998.