ROMOT: A Robotic 3D-Movie Theater Allowing Interaction and Multimodal Experiences

Sergio Casas, Cristina Portalés(✉), María Vidal-González, Inma García-Pereira, and Marcos Fernández

Institute of Robotics and Information and Communication Technologies (IRTIC), Universitat de València, València, Spain
{sergio.casas,cristina.portales,maria.vidal, inmaculada.garcia-pereira,marcos.fernandez}@uv.es

Abstract. In this paper we introduce ROMOT, a RObotic 3D-MOvie Theater. ROMOT is built with a robotic motion platform, includes multimodal devices and supports audience-film interaction. Differently from other similar systems, ROMOT is highly versatile as it can support different setups, integrated hardware and contents. Regarding to the setups, here we present a first-person movie, a mixed reality environment, a virtual reality interactive environment and an augmented reality mirror-based scene. Regarding to integrated hardware, the system currently integrates a variety of devices and displays that allow audiences to see, hear, smell, touch and feel the movement, all synchronized with the filmic experience. Finally, regarding to contents, here we present some samples related to driving safety and, in the discussion section, we theorize about the expansion of ROMOT for love and sex-related interactive movies.

Keywords: Multimodal · Interaction · 3D-movies · Augmented reality · Virtual reality · Mixed reality · Driving safety

1 Introduction

In 1962 Morton L. Heilig patented the Sensorama [1], one of the earliest immersive, multisensory (or multimodal) machine. The technology integrated in the Sensorama allowed a single person to see a stereoscopic film enhanced with seat motion, vibration, stereo sound, wind and aromas, which were triggered during the film. It was also referred to as "the cinema of the future" [2, 3]. However, when referring to cinemas in a wider sense, i.e. movie theaters, the filmic experience is collective rather than individual. Although there exist many research works dealing with multimodal technologies and environments [4, 5], they usually involve individual rather than collective experiences, and/or refer only to the involved technology.

On the other hand, the rapid technological advancements of the last years have allowed the development of commercial solutions that integrate a variety of multimodal displays in movie theaters, such as in [6–8], where these systems are usually referred to as 4D or 5D cinemas or theatres. Some claim that this technology shifts the cinema experience from "watching the movie to almost living it" [9], also enhancing the

A.D. Cheok et al. (Eds.): LSR 2016, LNAI 10237, pp. 50–63, 2017.
DOI: 10.1007/978-3-319-57738-8_5

cinematic experience while creating a new and contemporary version of storytelling, which can be conceptualized as a “reboot cinema” [10].

However, the criterion followed to establish the number of dimensions is not unified. In fact, the naming seems to follow commercial purposes rather than referring to physical-based dimensions. For instance, according to some, 4D cinema expands the 3D cinema by allowing a range of real-time sensory effects including seat movements, leg and back pulsation, projected wind and mist blasts, fog, lightning, scent perfume discharge, etc., all synchronized with the narrative of the film [9]. According to others, the fourth dimension corresponds to the movement and/or vibration of the seat, whereas the fifth dimension integrates the rest of sensory effects [11]. In [11] further distinctions are made that include some kind of interaction and wearing virtual reality glasses, justifying more than 7D, though the referred systems seem more close to single-person virtual reality simulators than to movie theaters. Having said this, we prefer to use the term “multimodal/multisensory 3D-movie theater” when referring to rooms exhibiting stereoscopic films enhanced with sensorial stimuli that can be experienced by a group of persons simultaneously. A step beyond would be to add to the film feedback of the users, leading to interactive multimodal/multisensory 3D-movie theaters.

In this paper we present ROMOT, a RObotized 3D-MOvie Theater. ROMOT follows the concept of 3D-movie theater with a robotized motion platform and integrated multimodal devices. Additionally, it supports audience-film interaction. Based on this, the audience gets some kind of reward by the system. Furthermore, in this sense, the whole system can be perceived as being alive, a kind-of huge robot around the audience. Additionally, ROMOT is highly versatile as it is prepared to support different types of setups, hardware and content, including films/animations that could be related to learning, entertainment, love, sex, etc. In this paper we present different kind of stereoscopic content related to driving safety, as ROMOT was initially built for an exhibition with this end. The following setups are integrated in ROMOT and shown in the paper: a first-person movie, a Mixed Reality environment, a Virtual Reality interactive environment and an Augmented Reality mirror-based scene. The contents of all of the different setups are based on a storytelling and are seen stereoscopically, so they can be broadly referred to as 3D-movies.

This paper is organized in the following way. First, we show the main technical aspects behind the construction of ROMOT and the integrated multimodal devices and interaction capabilities. It is worth mentioning that, differently from other existing commercial solutions, we have used a 180º curved screen to enhance user immersion. Then, we show the different kind of setups and content that were created for the exhibition. Finally, we discuss the expansion of ROMOT for love and sex-related movies.

2 Construction of the Robotized House (Audience)

The house (audience) was robotized by means of a 3-DOF motion platform with capacity for 12 people (Fig. 1). The seats are distributed in two rows, where the first row has 5 seats and the second one, 7 seats. This motion platform is equipped with three 2.2 kW *SEW Eurodrive* electric motors coupled with a 58.34 reduction-drive. The parallel

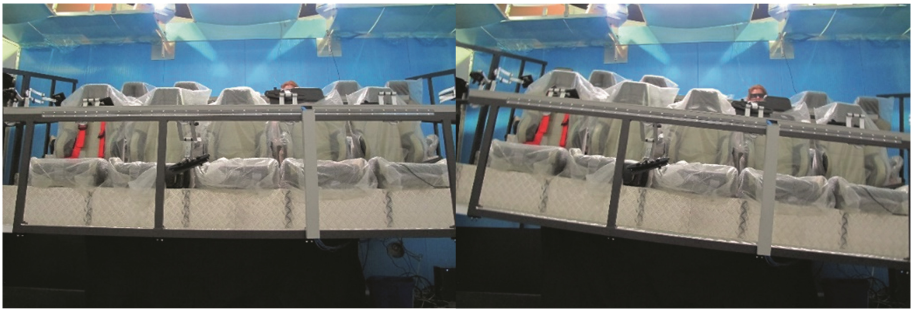

Fig. 1. Images of the robotized house. In the left image, the motion platform is at rest. In the right image, the platform is tilted.

design of the robotic manipulator alongside with the powerful 880 N·m motor-reduction set, provide a total payload of 1500 kg, enough to withstand and move the 12 people and their seats.

The design of the robotized motion platform allows for two rotational movements (pitch and roll tilt) and one translational displacement along the vertical axis (heave displacement). The motion platform is capable of featuring two pure rotational DOF, one pure translational DOF (the vertical displacement) plus two "simulated" translational DOF by making use of the tilt-coordination technique [12] (using pitch and roll tilt to simulate low-frequency forward and lateral accelerations). Thus, it is capable of working with five DOF, being the yaw rotation the only one completely missing. It is, therefore, a good compromise between performance and cost, since it is considerably cheaper to build than a 6-DOF Stewart motion platform [13], but its performance could be similar for some applications [14]. The motion platform is controlled by self-written software using the MODBUS/TCP protocol. The software includes not only the

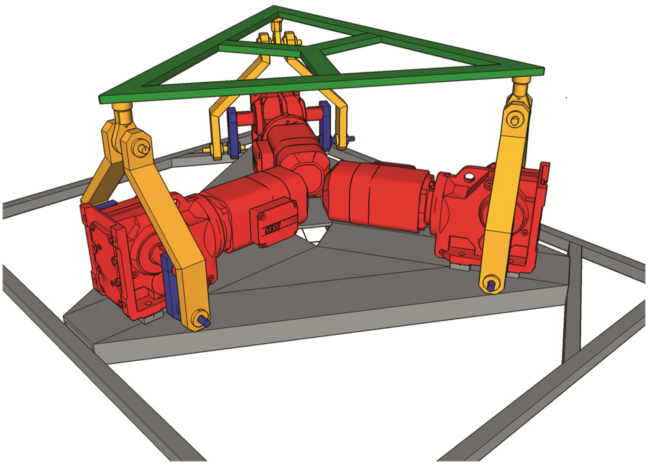

Fig. 2. 3-DOF parallel platform. (Color figure online)

actuators' control but also the classical washout algorithm [15], tuned with the method described in [16].

Figure 2 shows the kinematic design of the motion platform. The 12 seats and people are placed on the motion base (green), which is moved by three powerful rotational motors (red) that actuate over the robot legs (blue). The elements in yellow transmit the rotational motion of the motors to the motion base while ensuring that the robot does not turn around the vertical axis (yaw).

The motion envelope of parallel manipulators is always a complex hyper-volume. Therefore, only the maximum linear/angular displacements for each individual DOF can be shown (see Table 1). Combining different DOF results in a reduction of the amount of reachable linear/angular displacement of each DOF. Nevertheless, this parallel design allows for big payloads, which was one of the key needs for this project, and fast motion [17]. In fact, the robotized motion platform is capable of performing a whole excursion in less than 1 s.

Table 1. Motion platform excursions for each individual DOF.

	Heave [m]	Pitch [º]	Roll [º]
Minimum	−0.125	−12.89	−10.83
Maximum	+0.125	+12.89	+10.83

In front of the motion platform, a curved 180º screen is placed (Fig. 3), with 3 m height (and a 1.4 m high extension to display additional content) and with a radius of 3.4 m. Four projectors display a continuous scene on the screen, generated from two different camera positions to allow stereoscopy. Therefore, to properly see the 3D content, users need to wear 3D glasses.

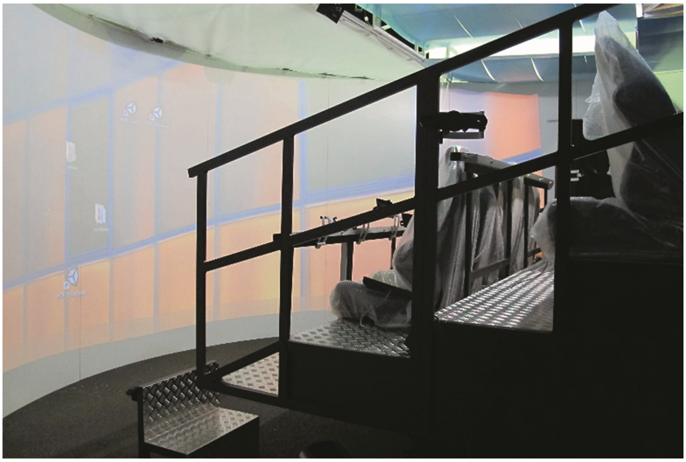

Fig. 3. Image of the curved screen and the house with seats.

Although some smaller setups introduce the display infrastructure on the motion-platform (so that they move together and inertial cues are correctly correlated with visual cues), the dimensions of the screen strongly recommend that the display infrastructure is kept fixed on the ground. Therefore, the visual parallax produced when the motion platform tilts or is displaced with respect to the screen needs to be corrected by reshaping the virtual camera properties so that the inertial and visual cues match. This introduces an additional complexity to the system, but allows the motion platform to be lighter and produce higher accelerations, increasing the motion fidelity [18].

3 Adding Multimodal Devices to ROMOT

In order to enrich the experience of the users and make the filmic scenes more realistic, a set of multimodal displays was added to the robotized platform:

- *An olfactory display*. We used the *Olorama* wireless aromatizer [19]. It features12 scents arranged in 12 pre-charged channels, that can be chosen and triggered by means of a UDP packet. The device is equipped with a programmable fan that spreads the scent around. Both the intensity of the chosen scent (amount of time the scent valve is open) and the amount of fan time can be programmed.
- *A smoke generator*. We used a Quarkpro QF-1200. It is equipped with a DMX interface, so it is possible to control and synchronize the amount of smoke from a computer, by using a DMX-USB interface such as the *Enttec Open DMX USB* [20].
- *Air and water dispensers*. A total of 12 air and water dispensers (one for each seat) (Fig. 4). The water and air system was built using an air compressor, a water recipient, 12 air electro-valves, 12 water electro-valves, 24 electric relays and two *Arduino Uno* to be able to control the relays from the PC and open the electro-valves to spray water or produce air.

Fig. 4. An image showing some of the air and water dispensers to the back of the first-row of seats, facing the audience located in the second row of seats.

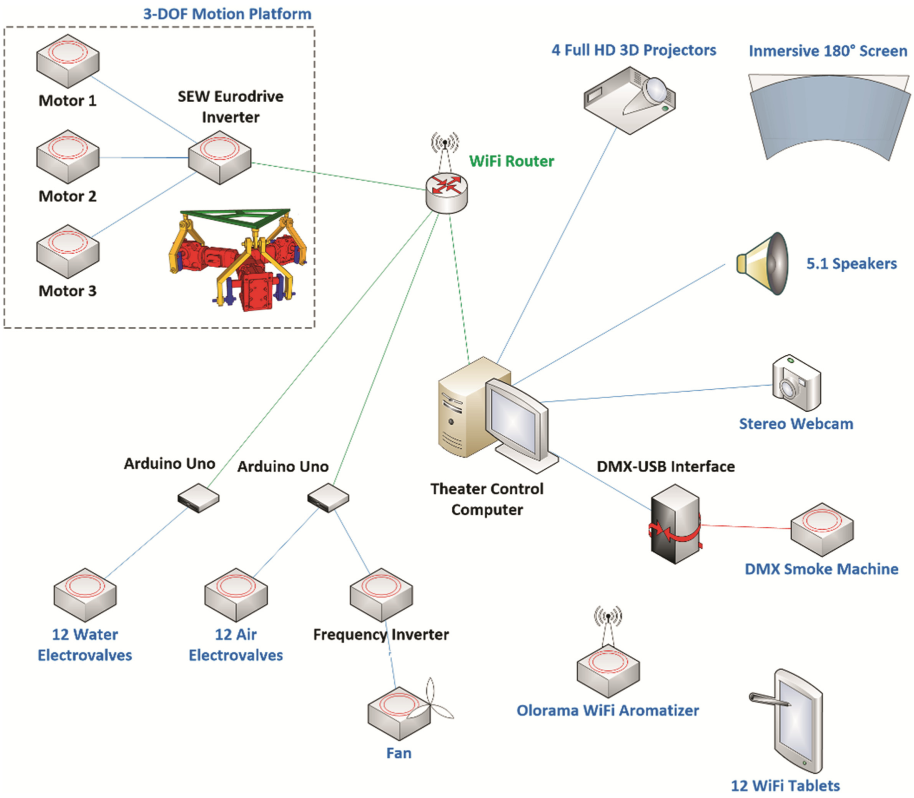

Fig. 5. Schema of the multimodal displays and other hardware involved in ROMOT.

- *An electric fan*. This fan is controllable by means of a frequency inverter connected to one of the previous *Arduino Uno* devices.
- *Projectors*. A total of 4 Full HD 3D projectors.
- *Glasses*. A total of 12 3D glasses (one for each person).
- *Loudspeaker*. A 5.0 loudspeaker system to produce binaural sound.
- *Tablets*. A total of 12 individual tablets (one for each person).
- *Webcam*. A stereoscopic webcam to be able to construct an augmented reality mirror-based environment.

It is important that all the multimodal actuators can be controlled from a computer, so that they can be synchronized with the displayed content and with the motion platform (Fig. 5).

Within this set of multimodal displays, users are able to feel the system's response through five of their senses:

- *Sight*: they can see a 3D representation of the scenes at the curved screen and through the 3D glasses; they can see additional interactive content at the tablets; they can see the smoke.
- *Hearing*: they can hear the sound synchronized with the 3D content.

- *Smell*: they can smell essences. For instance, when a car crashes, they can smell the smoke. In fact, they can even feel the smoke around them.
- *Touch*: they can feel the touch of air and water on their bodies; they can touch the tablets.
- *Kinesthetic*: they can feel the movement of the 3-DOF platform.

Apart from that, the audience can provide inputs to ROMOT through the provided tablets (one tablet per person). This interaction is integrated in the setup of the "3D-virtual reality interactive environment", which is explained as part of the following section.

4 Developed Stereoscopic Content

Four different stereoscopic content were elaborated for different system setups, which are described in the following sub-sections.

4.1 First-Person Movie

A set of videos were recorded using two GoPro cameras to create a 3D movie set in the streets of a city. Most of the videos were filmed attaching the GoPro cameras to a car's hood in order to create a journey with a first-person view and increase the audience immersive experience by locating them at the center of the view, as if they were the protagonists of the journey.

At every moment there's audio consisting of ambient sounds and/or a locution that reinforces the images the user is watching. In some cases, synchronized soft platform movements or effects like a nice smell or a gentle breeze help to create the perfect ambient at each part of the movie, and make the experience more enjoyable for the audience.

4.2 Mixed Reality Environment

3D video and 3D virtual content can be mixed creating a Mixed Reality movie that helps the audience to perceive the virtual content as if it were real, making the transition from a real movie to a virtual situation easier.

In this setup, the created 3D virtual content – a 3D virtual character – interacts with parts of the video by creating the virtual animation in such a way that it is synchronized with the contents of the recorded real scene. Virtual shadows are also considered to make the whole scene more real (Fig. 6).

Fig. 6. Example of the mixed reality setup.

4.3 Virtual Reality Interactive Environment

In this case, different buildings were created and merged to a map of the streets of a city. Street furniture, traffic signs, traffic lights, etc. were added too in order to recreate the city as detailed as possible. Vehicles and pedestrian were further animated to create every situation as real as possible (Fig. 7).

Each situation was created using a storyboard that contains all the contents, camera movements, special effects, locutions, etc. So at the end, a set of situations were derived that could be part of a movie.

In this case, we want the audience not to just look at the screen and enjoy a movie but to make them feel each situation, to be part of it and to react to it. That is why platform movements and all the other multimodal displays are so important.

Fig. 7. The created 3D city with vehicles and pedestrians.

When each situation takes place, the audience can feel that they are driving inside the car, thanks to the platform movements that simulate the movements of a real car (accelerations, decelerations, turns…). In some of the scenes, the virtual situation pauses and asks the audience for their collaboration (Fig. 8). At that moment, the different tablets vibrate and a question appears, giving the individual users some time to answer it by selecting one of the possible answers. When the time is up, they are prompted to report whether the answer was correct or not, and the virtual situation resumes, showing the consequences of choosing a right or a wrong decision. Crashes, outrages, rollovers… the audience can feel in first person the consequences of having an accident thanks to the platform movements and the rest of multimodal feedback, such as smoke, smell, etc.

Fig. 8. Tablet pause. Users have to look to their tablets and choose one of the options.

Each correct answer increases the individual score at each of the tablets. When the deployed situation finishes, the audience can see the final score on the big curved screen. The people having the greatest score are the winners who are somehow rewarded by the system by receiving a special visit, a 3D virtual character that congratulates them for their safe driving (see next sub-section).

4.4 Augmented Reality Mirror-Based Scene

This setup consists of a video-based Augmented Reality Mirror (ARM) [21] scene, which is also seen stereoscopically. ARMs can bring a further step in user immersion, as the audience can actually see a real-time image of themselves and feel part of the created environment.

This ARM environment is used in the final scenes of the aforementioned virtual reality interactive environment (previous sub-section), where the user(s) with the highest score get(s) rewarded by a virtual 3D character that walks towards him/her/them. Together with this action, virtual confetti and colored stars appear on the environment, accompanied with winning music that includes applauses (Fig. 9).

Fig. 9. Audience immersed in the Augmented Reality Mirror-based scenario. One person receives the visit of the virtual character that congratulates him for being the winner.

5 Discussion: Expansion of ROMOT for Romantic and Sex Content

ROMOT is a laboratory system (hardware and software) that has been built from scratch. Because of this, and differently from overall commercial systems, ROMOT is highly versatile, being easily adapted to different kinds of public, purposes, contents, setups, etc., as both the hardware and the software can be modified with relatively little effort. In this section we theorize about the possible expansion of ROMOT for love and sex-related interactive content and/or experiences.

Romantic content can be easily adapted to ROMOT. For instance, users can experience films with alternative endings. At a certain point of the movie, we can ask the audience “will she marry him?” and it will be the public who decides. User studies could be performed on this, on whether public prefers a happy or an unhappy ending. Gender analysis could be furthermore performed. This might be already done in traditional movies, but with ROMOT this will acquire a richer experience, as the number of multimodal stimuli and the different kind of scenarios can expand the possibilities and enhance user immersion. For instance: audience could smell flowers when he gives them to her as a present; we can simulate the motion of a car, when they go on a romantic trip crossing the country, etc.

In order to have a first evaluation of the intention of the general public to see romantic or sex scenes within ROMOT, we have done a survey where people had to answer a short questionnaire. To that end, a romantic scene and a sex scene of know movies were selected. In the following lines, these case studies are introduced and then the results of the questionnaire and outlined.

5.1 Description of the Selected Scenes

As a first case study, the known as the most romantic scene of Titanic is here reported [22]: *Rose gasps. There is nothing in her field of vision but water. It's like there is no ship under them at all, just the two of them soaring. The Atlantic unrolls toward her, a hammered copper shield under a dusk sky. There is only the wind, and the hiss of the water 50 feel below. Rose: "I'm flying!" She leans forward, arching her back. He puts his hands on her waist to steady her. Jack: "Come Josephine in my flying machine…" Rose closes her eyes, feeling herself floating weightless far above the sea. She smiles dreamily, then leans back, gently pressing her back against his chest. He pushes forward slightly against her.* How could this scene be reproduced in ROMOT? The audience would feel fresh air on their faces, smell of sea, humidity, their seats moving accompanied with the waves. We believe that they could reach a greater immersion in the movie and thus feel like they are Rose; they are flying, too (perceive the scene in first person).

In the same way, sex scenes could be reproduced and enhanced in ROMOT. As the second case study, the sex scene in *The Twilight Saga: Breaking Dawn (Part 1)* is here reported. In that movie, Bella marries Eduard, the vampire. They are in their honeymoon. They decide to have sex for the first time (partially taken from [23]): *Bella walks into another area of the house. Bedroom. Stares at the bed, touches the curtains around the bed. Can see ocean in background […] Edward walks outside. Takes off shirt to get into water. […] Bella walks up to Edward in the moonlight. Bella is naked and walking towards the water. Edward already in. They kiss each other. Then, back to the bedroom. They have sex. The next day, the bed appears completely broken and Bella covered with pillow feathers*. In ROMOT audience would feel the ocean breeze. Then, at the bedroom, bed movements could be directly transferred to smooth vibrations and platforms movements. The scene would end with very abrupt movements, and the audience would feel like something has broken, too. Some pillow feathers could also be projected onto the audience.

5.2 Results of the Questionnaire

A total of 22 persons participated in the survey, 50% woman and 50% men with ages ranging from 20 to 50 years old. Those people were related to the IRTIC lab, as we wanted to have the feedback of people that had already interacted with ROMOT. We have asked them the following question for four different situations: "I would feel comfortable watching romantic/sex scenes in ROMOT in the following situations: (a) If I don't know the rest of the audience; (b) If I go with my fiancée/husband/wife, but we don't know the rest of the audience; (c) If I go with friends/colleagues; (d) If I go with my family/relatives (adults)". For each of the four situations and for each case study (romantic/sex) they had to give a score from 1 (completely disagree) to 5 (completely agree).

The results of the given answers are depicted in Fig. 10 (romantic case study) and Fig. 11 (sex case study). As it can be seen, when comparing the results of the different case studies, we can say that people would feel more comfortable having romantic experiences in ROMOT. When looking to the different situations for both scenarios, most people would feel uncomfortable when going with their family or relatives, especially in the sex case scenario. This happens even if the selected scene does not involve explicit sex content. However, most people would feel comfortable with the sex case scenario if they go with friends or colleagues.

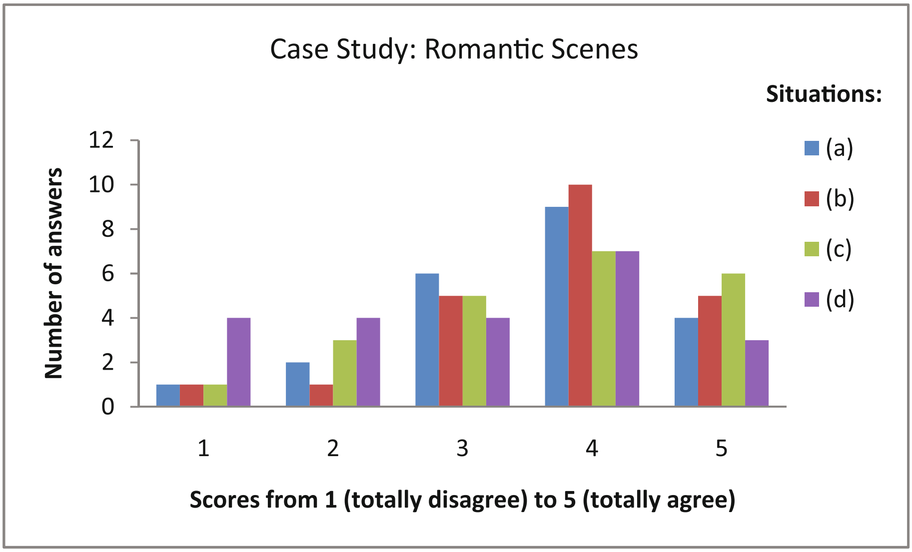

Fig. 10. Answers given for the case study of romantic scenes.

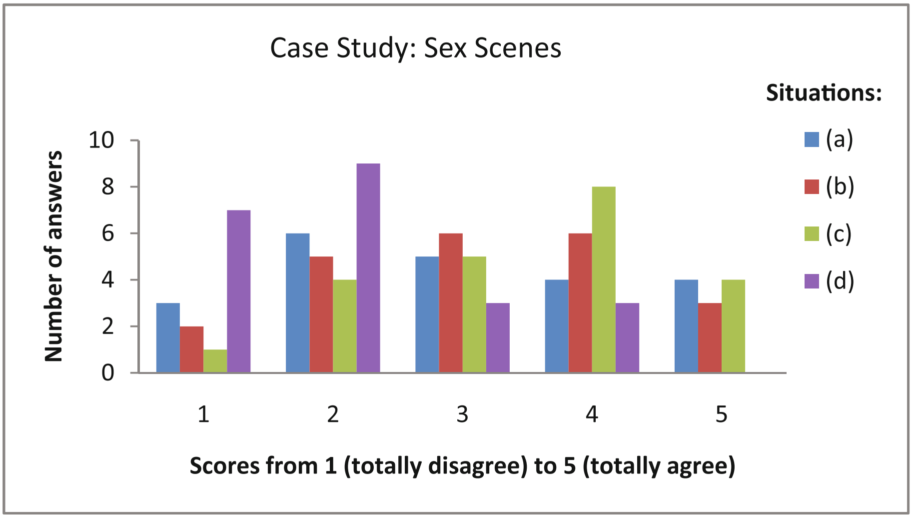

Fig. 11. Answers given for the case study of sex scenes.

5.3 Integrating Other Display Technologies

Also, other kind of display technologies could be included in ROMOT to enhance movies with other stimuli such as taste or touch. For instance, the *Kissinger* technology [24] could provide the audience with a kiss previously recorded by the actors themselves, and individuals could choose between being kissed by her or by him. Could you image being kissed by Jack or by Rose in the Titanic scene? In a similar way, hugs could be also transferred to the audience with a technology like the *Huggy Pajama* [25], so the audience could feel the hugs of actors in romantic and sex scenes. This technology could open new avenues in the related creative industries, where actors could collaborate to create a new kind of perceptions to attract huge audiences.

Finally, we would like to highlight that, as ROMOT involves the whole audience and thus it can be perceived as a collective experience, it could bring ethical issues when related to explicit sex content, and thus this could be an interesting field of research. For instance, would the audience immersed in ROMOT feel like they are experiencing a kind-of orgy and perhaps making them feel uncomfortable? Or would they feel the scene as isolated from the rest of the audience?

6 Conclusion

In this paper we have presented ROMOT, a robotized 3D-movie theatre. The work shown in this paper relates to the enhancement of audience experiences when integrating multimodal stimuli and making it interactive. In addition, we show the versatility of the system by means of the different kind of generated content.

Both the setups and the filmic contents of ROMOT can be changed for different kind of user experiences. As the different setups, we have shown a first-person movie and others related to the technologies of virtual, augmented and mixed realities. Regarding to the content, we have briefly shown samples related to driving safety and added a discussion for the use of ROMOT with romantic and sex content. To support our statements, a first user evaluation has been given that reveals that people would feel more comfortable when experiencing romantic content with ROMOT, but also they would feel comfortable experiencing sex content if they go with friends or colleagues.

We also have discussed the further integration of other display technologies within ROMOT that include the possibility to transmit kisses and hugs to the audience. In this way, people could feel like being touched by the favorite actors, and thus this can open new business avenues for the creative industries.

References

1. Heilig, M.L.: Sensorama simulator. Google Patents (1962)
2. Heilig, M.L.: El cine del futuro: the cinema of the future. Presence Teleoper. Virtual Environ. **1**, 279–294 (1992)

3. Robinett, W.: Interactivity and individual viewpoint in shared virtual worlds: the big screen vs. networked personal displays. SIGGRAPH. Comput. Graph. **28**, 127–130 (1994)
4. Ikei, Y., Okuya, Y., Shimabukuro, S., Abe, K., Amemiya, T., Hirota, K.: To relive a valuable experience of the world at the digital museum. In: Yamamoto, S. (ed.) HCI 2014. LNCS, vol. 8522, pp. 501–510. Springer, Cham (2014). doi:10.1007/978-3-319-07863-2_48
5. Matsukura, H., Yoneda, T., Ishida, H.: Smelling screen: development and evaluation of an olfactory display system for presenting a virtual odor source. IEEE Trans. Visual Comput. Graphics **19**, 606–615 (2013)
6. 4DX. http://www.cj4dx.com/about/about.asp
7. Pix 5D Cinema. http://expressavenue.in/?q=store/pix-5d-cinema
8. 5D cinema extreme. http://www.5dcinema.hu/
9. Yecies, B.: Transnational collaboration of the multisensory kind: exploiting Korean 4D cinema in China. Media Int. Australia **159**, 22–31 (2016)
10. Tryon, C.: Reboot cinema. Convergence: Int. J. Res. New Media Technol. **19**, 432–437 (2013)
11. http://www.xd-cinema.com/the-difference-between-4d-5d-6d-7d-8d-9d-xd-cinema/
12. Groen, E.L., Bles, W.: How to use body tilt for the simulation of linear self motion. J. Vestib. Res. **14**, 375–385 (2004)
13. Stewart, D.: A platform with six degrees of freedom (1965)
14. Casas, S., Coma, I., Riera, J.V., Fernández, M.: Motion-cuing algorithms: characterization of users' perception. Hum. Factors: J. Hum. Factors Ergon. Soc. **57**, 144–162 (2015)
15. Nahon, M.A., Reid, L.D.: Simulator motion-drive algorithms - a designer's perspective. J. Guid. Control Dyn. **13**, 356–362 (1990)
16. Casas, S., Coma, I., Portalés, C., Fernández, M.: Towards a simulation-based tuning of motion cueing algorithms. Simul. Model. Pract. Theory **67**, 137–154 (2016)
17. Küçük, S.: Serial and Parallel Robot Manipulators - Kinematics, Dynamics, Control and Optimization. InTech (2012)
18. Sinacori, J.B.: The Determination of Some Requirements for a Helicopter Flight Research Simulation Facility. Moffet Field (1977)
19. Olorama Technology. http://www.olorama.com/en/
20. Enttec. http://www.enttec.com/
21. Portalés, C., Gimeno, J., Casas, S., Olanda, R., Giner, F.: Interacting with augmented reality mirrors. In: Rodrigues, J., Cardoso, P., Monteiro, J., Figueiredo, M. (eds.) Handbook of Research on Human-Computer Interfaces, Developments and Applications, pp. 216–244. IGI-Global (2016)
22. IMSDb. http://www.imsdb.com/scripts/Titanic.html
23. Hypable. http://www.hypable.com/detailed-description-of-breaking-dawn-honeymoon-scene
24. Lovotics. http://kissenger.lovotics.com/
25. Teh, J.K.S., Cheok, A.D., Peiris, R.L., Choi, Y., Thuong, V., Lai, S.: Huggy Pajama: a mobile parent and child hugging communication system. In: Proceedings of the 7th International Conference on Interaction Design and Children, Chicago, Illinois, USA, 11–13 June 2008, Proceedings, pp. 250–257. ACM (2008)