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

Author 1, Author 2, Author 3, Author 4, Author 5

Address [REMOVED FOR REVIEWING]
{Author1, Author2, Author3, Author4, Author5}@xxx.com

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. We present the results of some preliminary user tests made at the laboratory level and they give satisfactory scores for the usability of the system and the individual's satisfaction. As well, we show the versatility of the system by means of different types of system setups and generated content that includes a first-person movie and others involving the technologies of virtual, augmented and mixed realities. The contents shown here are related to driving safety, though they could be adapted to other fields, such as learning, entertainment, love, sex, etc.

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 technology and environments [4, 5], usually they involve individual rather than collective experiences, and/or only refer 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 cinematic experience while creating a new and contemporary version of storytelling, which can be conceptualized as a "reboot cinema" [10].

adfa, p. 1, 2016.

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 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 in the Middle East 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 show the first outcomes regarding the usability of the system and the individual’s satisfaction. To the best of our knowledge, this is the first work reporting audience experiences in such a complete system.

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 *SEW Eurodrive* 2.2 Kw electric motors coupled with a 58.34 reduction-drive. The parallel 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.

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.

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

Fig. 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 second.

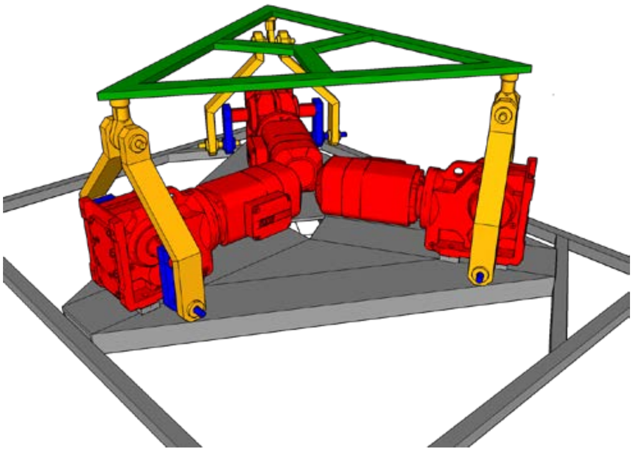

Fig. 2. 3-DOF parallel platform.

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.

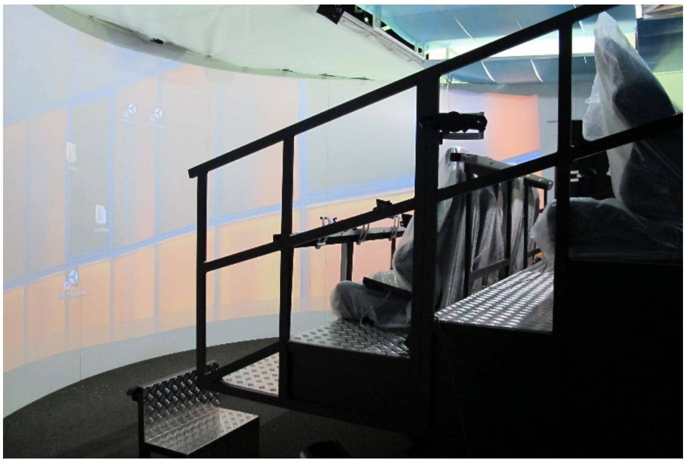

Fig. 3. Image of the curved screen and the house with seats.

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 set) (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.
- *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.

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.

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).

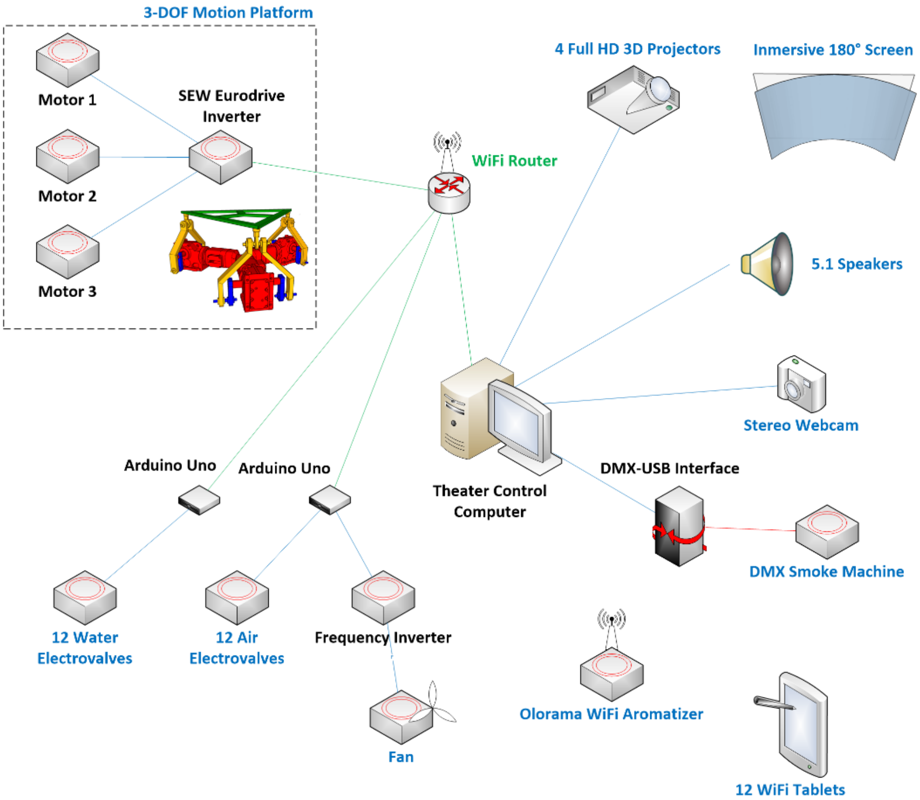

Fig. 5. Schema of the multimodal displays and other hardware involved in ROMOT.

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 driving-related 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 some audio consisting of ambient sounds and/or a locution that reinforces the images the user is looking at. 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

Situations such as accidents, outrages or serious violations of traffic safety norms cannot be easily recreated in the real world. On the other hand, by creating a virtual city, all of the aforementioned situations can be recreated without damaging anyone. 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 create 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.

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

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.

When each situation takes place, the audience can feel that they are in the car driving 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.

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).

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

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 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 User tests

In this section, we present the results of some preliminary user tests made at the laboratory level (Fig. 10). A total of 14 people tested the system and participated in its evaluation, which consisted on filling out a pair of questionnaires that were related to the usability of the system and the individual's satisfaction, where the System Usability Scale (SUS) [22] was chosen to measure usability.

The participants were some of the research staff of the [REMOVED FOR REVIEWING], where 7 of them were women and 7 were men. Only 3 of the participants contributed to the present research work at some of their stages. Though this might not be considered as a full objective audience, it can give us a good first notion for a user-related evaluation of the system.

Fig. 10. Image of some of the research staff testing ROMOT.

The results of the SUS questionnaire are listed in Table 2. In questions 1 to 10 the range 0–4 means: 0: strongly disagree, 4: strongly agree. The values of the SUS score, however, range from 0 to 100, meaning 100 the best imaginable result. In the case of the ROMOT evaluation, this score reaches 84.25 points, which can be considered excellent on the scale of scores provided by the questionnaire and taking into account the fact a minimum score of 68 would be deemed acceptable for a tool [23, 24].

Table 2. Results of the SUS questionnaire (mean, standard deviation, minimum and maximum).

Questions	Mean	S.d.	Min	Max
1. I think that I would like to use this system frequently	3.00	0.89	1	4
2. I found the system unnecessarily complex	0.70	0.78	0	2
3. I thought the system was easy to use	3.50	0.5	3	4
4. I think that I would need the support of a technical person to be able to use this system	1.30	1	0	3
5. I found the various functions in this system were well integrated	3.60	0.66	2	4
6. I thought there was too much inconsistency in this system	0.40	0.49	0	1
7. I would imagine that most people would learn to use this system very quickly	3.60	0.49	3	4
8. I found the system very cumbersome to use	0.60	0.8	0	2
9. I felt very confident using the system	3.30	0.64	2	4
10. I needed to learn a lot of things before I could get going with this system	0.30	0.46	0	1
SUS score	**84.25**			

The results of the individuals' satisfaction questionnaires are given in Table 2. The scores also range from 0 to 4, meaning: 0: strongly disagree, 4: strongly agree. As it can be seen, results are quite satisfactory as 8 out of the 12 mean scores are over 3 points and none are under 2.5 points. The lowest mean score belongs to question 10: "I didn't feel sick after using the ROMOT", which is an issue difficult to tackle, as simulator sickness highly depends on the individual. On the other hand, the highest mean score belongs to question 12: "I would like to recommend others to use ROMOT", meaning that overall the individuals are satisfied with the system.

Table 3. Results of the individuals' satisfaction questionnaire (mean, standard deviation, minimum and maximum).

Questions	Mean	S.d.	Min	Max
1. Overall, I liked very much using ROMOT	3.14	0.74	2	4
2. I find it very easy to engage with the multimodal content	3.29	0.8	2	4
3. I enjoyed watching the 3D movies	3.29	0.8	1	4
4. The audio was very well integrated with the 3D movies	3.43	0.73	2	4
5. The smoke was very well integrated in the virtual reality interactive environment	2.71	0.88	1	4
6. The smell was very well integrated in the virtual reality interactive environment	2.93	1.1	0	4
7. The air and water were very well integrated in the virtual reality interactive environment	2.79	1.01	0	4
8. The movement of the platform was very well synchronized with the movies	3.36	0.72	2	4
9. The interaction with the tablet was very intuitive	3.07	0.8	2	4
10. I didn't feel sick after using ROMOT	2.64	1.23	0	4
11. I would like to use again ROMOT	3.36	0.61	2	4
12. I would like to recommend others to use ROMOT	3.64	0.48	3	4

6 Conclusion

In this paper we have presented the construction and first user evaluation of 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. As well, 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. Here we have shown different setups for content related to driving safety, though other filmic contents could be used, including some related to learning, entertainment, love, sex, etc. As the different setups, we have shown a first-person movie and others related to the technologies of virtual, augmented and mixed realities.

The outcomes regarding the usability of the system and the individual's satisfaction are very satisfactory, though we are aware that the system has only been evaluated at the laboratory level. It is also worth mentioning that, although there exist different commercial solutions (e.g. 4D/5D cinemas), we have not found complete research works dealing with the construction and audience evaluation of such systems.

As further work, we intend to evaluate the developed system in a real environment (the exhibition space) with the participation of hundreds of visitors. We also would like to create more filmic contents in order to make use of ROMOT for other purposes.

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.) Human Interface and the Management of Information. Information and Knowledge in Applications and Services: 16th International Conference, HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014. Proceedings, Part II, pp. 501-510. Springer International Publishing, Cham (2014)
5. Matsukura, H., Yoneda, T., Ishida, H.: Smelling Screen: Development and Evaluation of an Olfactory Display System for Presenting a Virtual Odor Source. IEEE Transactions on Visualization and Computer Graphics 19, 606-615 (2013)
6. http://www.cj4dx.com/about/about.asp
7. http://expressavenue.in/?q=store/pix-5d-cinema
8. http://www.5dcinema.hu/
9. Yecies, B.: Transnational collaboration of the multisensory kind: Exploiting Korean 4D cinema in China. Media International Australia 159, 22-31 (2016)
10. Tryon, C.: Reboot cinema. Convergence: The International Journal of Research into New Media Technologies 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. Journal of Vestibular Research 14, 375-385 (2004)
13. Stewart, D.: A Platform with six degrees of freedom. (Year)
14. Casas, S., Coma, I., Riera, J.V., Fernández, M.: Motion-Cuing Algorithms: Characterization of Users' Perception. Human Factors: The Journal of the Human Factors and Ergonomics Society 57, 144-162 (2015)
15. Nahon, M.A., Reid, L.D.: Simulator motion-drive algorithms - A designer's perspective. Journal of Guidance, Control, and Dynamics 13, 356-362 (1990)
16. Casas, S., Coma, I., Portalés, C., Fernández, M.: Towards a simulation-based tuning of motion cueing algorithms. Simulation Modelling Practice and 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. http://www.olorama.com/en/
20. 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. Brooke, J.: SUS-A quick and dirty usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L. (eds.) Usability Eval. Ind, pp. 189-194. Taylor & Francis (1996)
23. Bangor, A., Kortum, P., Miller, J.: Determining what individual SUS scores mean: adding an adjective rating scale. J. Usability Studies 4, 114-123 (2009)
24. Brooke, J.: SUS: a retrospective. J. Usability Studies 8, 29-40 (2013)