Immersive retrospection by video-photogrammetry

UX assessment tool of interactions in museums, a case study

Emmanuel Beaudry Marchand\textsuperscript{1}, Xueying Han\textsuperscript{2}, Tomás Dorta\textsuperscript{3}
\textsuperscript{1,3}University of Montreal, Design Research Laboratory Hybridlab \textsuperscript{2}Harbin Institute of Technology
\textsuperscript{1,3}\{emmanuel.beaudry.marchand|tomas.dorta\}@umontreal.ca\textsuperscript{2}hanxueyinghxy119@163.com

Studying interactions in museums often omits to consider the complexity of the space and the visitors’ behaviors. Visitors’ walking paths do not provide enough insight of their user experience (UX) since they are distant from the experiential realities. Videogrammetry can convey such dimensions of an environmental experience. Because of limitations of real-time playback, a twofold approach is suggested: “immersive videos” combined with “photogrammetric models”. A granular optimal experience assessment method using retrospection interviews is also applied providing a finer evaluation of the perceived experience through time. This method permits to characterize museum interactive installations, according to the perceived challenges and skills of the interaction’s task, based this time on immersive retrospection. This paper proposes the “Immersive retrospection” by “Immersive video-photogrammetry” as a UX assessment tool of interactions in museums. A hybrid virtual environment was used in this study, allowing social VR without the use of headsets, through a life-sized projection of interactive 3D content. The study showed that Immersive video-photogrammetry facilitates the recall of memories and allows a deepened self-observation analysis.

**Keywords:** immersive retrospection, photogrammetry, videogrammetry, UX assessment, museum environments

**INTRODUCTION**

Traditionally, museum visitors are engaging with various forms of content communication throughout exhibitions. Textual displays and audio-guides are some examples of what is used to situate the visitors in the exhibit and complement the displayed content. However, some studies (Samis, 2008) highlight that few people engage actively when experiencing those approaches. Part of the reason is the fact that the information disseminated this way focuses excessively on collection-centered rather than user-centered installations, leading to few carefully thought through interaction designs.

Tools commonly used to study visitors’ interactions in museums often omit to consider the complexity of the environment thus designed for inter-
active purposes and visitors’ behaviors. For example, only a top view map of the visitors’ positions and their walking paths (Nasir, Nahavandi & Creighton, 2012) does not provide any insight into their actual gaze or general awareness which could be affected differently through time and the tridimensional space, depending on multiple factors: other visitors’ changing positions, height of the exhibits, different heights of participants, etc.

Such aspects are hard to depict analysing two-dimensional images. These images do not account for subtler physical clues regarding a visitor’s focus of attention or apprehension of a situation in a specific moment: head orientation in contrast to the body’s general position and orientation, or simple gestures such as pointing a certain part of the exhibition’s content or talking to another visitor. Moreover, traditional tools like path analysis constitute a risk of perceptual representations being distant from certain experiential realities of the studied situations, such as the atmosphere (Van de Vreken & Safin, 2010).

In this study, we focus our attention on the interactions as lived by visitors rather than the aspects concerning the exhibit content transfer. We postulate that overcoming undesirable situation implies opting for an alternative approach to exhibition design and evaluation, in this case concerned more with the visitor’s actual experience of the perceived content. In this paper, we propose a new framework by working to relocate, in an immersive way, the participants in key parts of a previous museum visit to assess their experience.

From the original granular methodology of user experience based on retrospective interview (Safin et al. 2016), this study goes a step further by proposing immersive retrospection using immersive videophotogrammetry as strategy to better recall past visits and interactions of a given exhibit. This paper presents for the first time the use of this framework in the case of a given museum space. The preliminary results from only 3 visitors seem to point towards a better evaluation of the user experiences in museums’ interactions.

INTERACTIONS IN MUSEUMS AND UX
Nowadays museums have moved from basic maintenance of collections to public services communicating ideas and providing exchange between them and their visitors. It is obvious that the aim of museums has become to emphasize on human-centered design and holistic experience making. The current approach shifted to focus on visitor’s attention via interactive exhibitions in museums. Interactive exhibitions would encourage visitors to explore them more directly. However, exhibitions that simply include pressing buttons, using mobile guides, etc. are not truly interactive, but rather reactive (Mortensen, 2010). The user experience (UX) for interactive museums is strongly configured by users’ demands of identity-related visit motivations which create a primary direction for visit (Falk & Dierking, 2016). The explicit information of what users see and perceive is covered by the following factors (Falk & Dierking, 2016): User: (prior knowledge, experience, and interest); Physical museum: (exhibitions, programs, goals, interior design and signage); Information: (interactions between users, users and exhibitions, users and physical museum).

ASSESSING USER-EXPERIENCE IN MUSEUMS
There are several methods in UX research that have been used. Observations and interviews are the most suitable to gather UX from non-verbal behaviors. Users may be unaware of their experiences or unable to express them verbally (Hsi, 2003). Moreover, other methods are proposed (Hartson & Pyla, 2012) but require the visitors to hold digital devices or wear sensors to assess their UX while visiting the museum. This can hinder the UX itself. Besides that, surveys, diaries and storytelling have been regarded as an effective way to get information about UX (Arhippainen & Tähti, 2003). This way users can express some of their experiences in a written form, including their background (age, education), prior experiences, expectations and motivations. Another method is to interview users facing a video recording of themselves.
(Hsi, 2003), but eventually because of ethical reasons, videos of the kind are recorded from different perspectives that cannot capture all facial expressions. The last common method is the use of a storyline and prototyping. Storyline (Pujol et al. 2012) is the way to organize and remember experiences enabling users to communicate them. Prototyping experiences of different situations allows designers, clients or users to “experience it themselves” rather than just witness a demonstration of someone else’s experience (Hartson & Pyla, 2012).

**PERCEPTUAL EXPERIENCE OF ENVIRONMENTS AND VIDEOGRAMMETRIC REPRESENTATIONS**

Ecological perception appears as a fruitful frame for setting a canvas of the implications of our endeavor for a new approach to UX assessment, regarding the human-environment perspective, the experience of space and its interactions as formatted by museum exhibitions. Ecological perception informs us that, from a perceptual standpoint, a visitor’s exploration and apprehension of a museum exhibit (be it direct or mediated) will be directly supported by the perception of its affordances (Gibson, 1979). Perceiving events as changes in the organization of these affordances may provide visitors with additional indications on how to intervene on their direct environment (Chemero, Klein & Cordeiro, 2003), hence interact with the exhibition. Events are temporal by definition, resulting from a change in the state of a situation, demanding some form of dynamism to be observed. Immersive videogrammetry is a media in development with the potential of conveying efficiently such dimensions of an environmental experience, allowing visitors to better remember their perception of a previously recorded situation than when using traditional images and videos for UX analysis. This would bring more depth to the analysis of lived situations during “self-observation interviews” conducted accordingly to Boubée (2011).

Videogrammetry encompasses the process in which multiple synchronized video streams of a scene are used to reconstruct its spatial-visual properties as a sequence of 3D models (4D, including time), draped with photographic textures, through image processing. Whereas the process is analogous to photogrammetry, it differs in the way the images are captured: photogrammetry uses successive pictures from different angles with a single camera, while videogrammetry employs instead multiple devices to record videos simultaneously from different positions. While many aspects of this digital technology have been approached throughout literature in computer sciences (Kanade, Rander & Narayanan, 1997), cases of its practical use remain for now widely unexplored and under-documented, limiting our comprehension and acknowledgement of its contribution in practice.

**Immersive videogrammetry**

In light of the informational capacities of videogrammetry, we suggest that displaying the videogrammetric model at life-size in an immersive environment could contribute to a stronger sense of presence for the participant and support a deepened recall of events. A stronger state of presence, in its turn, would reinforce a greater perception of affordances, also aided by the recognition of other individuals that were acting in the space (Stoffregen et al., 1999). This would mean making intuitive the perceived possibilities of action within the perceptual and hybrid immersive representation (Riva et al., 2011) by altering how the users relate their actual bodies to the immersive space (Schubert, Friedmann & Regenbrecht, 2001). A visualization of the kind would allow us to assess visitors regarding their past experiences of an interactive installation without the need to interrupt their actions as they unfold.

**OPTIMAL USER EXPERIENCE AND GRANULAR ASSESSMENT**

To assess UX of museums’ interactions we selected the model of the optimal experience of Csikszentmihalyi (1997). The optimal experience or state of Flow, is characterized by high “challenge” of the task to be
performed and high “skills” of the users. The Flow is an autotelic experience, considered as a memorable, gratifying experience where users lose the notion of time (Csikszentmihalyi, 1997). Other psychological states can be experienced according to the changes between the challenges and skills regarding the task: Stress (high challenge and low skills), Control (low challenges and high skills) and Boredom (low challenges and low skills) (Massimini & Carli, 1988); (Safin et al. 2016). In the case of visitors’ museum, it would be possible to measure the perceived challenges and skills when they must engage with a given interactive installation. Visitors expectations and involvement with the interaction can be unfolded this way, but the remaining issue would be to observe it during time to better understand the reasons of the UX changes. In the context of evaluating designers’ experiences, Safin et al. (2016) developed a methodology of assessing UX in a granular way by using retrospection of self-observation interviews and two sliders (one measuring the challenges and other the skills) in a specific device (Korg NanoKONTROL2(TM)) and software (Max Run Time). This way the whole UX, including its different states (Stress, Flow, Control and Boredom) can be assessed for each second, allowing to combine this data with other kinds of information, like in this study, analysing different kinds of interactive installations. Moreover, the limits of recalling a past activity (or museum visit) using traditional footage during the retrospective interviews could be improved using immersive videos and photogrammetry.

**TOWARDS VIDEOGRAMMETRY: VIDEO-PHOTOGRAMMETRY**

Considering the current technological limitations associated to videogrammetry - namely its real-time interactive playback and the uncertainty regarding the quality of the models it offers - we propose a more reliable twofold alternative approach to relocate users in parts of their previous museum visit (Fig. 1). To proceed, the events are first recorded as they unfold in the chosen exhibition room using both the videogrammetric setup and a 360° camera (Nikon KeyMission 360(TM)) positioned at the center of the space. Afterwards, immersive videos are presented to the participants allowing them to observe their own behaviours. This immersive spherical video was edited from the 360° footage (Fig. 2)
and adapted to be used in Hyve-3D(TM) (Hybrid Virtual Environment 3D) (Dorta, Kinayoglu & Hoffmann, 2016). Hyve-3D offers a Social virtual reality (VR) experience, without the use of headsets, through an anamorphic life-sized projection of interactive 3D content. The Social VR experience allowed the recognition of other visitors’ behaviours that were participating, asynchronously in the immersive video and synchronously via collective interviews. Through this first re-visit and after the analysis of the UX states, moments of interest are identified to indicate which frames to extract from the videogrammetric data for the production of 3D models.

**METHODOLOGY**

Using the granular experience assessment method developed by Safin et al. (2016) as described above, key moments of the subjective experience were identified through the visitors’ evaluations as they were spectating their recorded visit in the immersive video (life-sized scale) from a nodal perspective (from 360° camera point of view). This method calls for participants to indicate how they perceived the “challenges” of the previously lived interactions and their felt “skills” level, all through an interface with corresponding sliders (Fig. 2). This way, we aimed to obtain a finer evaluation and evolution of the lived experience through time, allowing us to pinpoint specific interactive settings based on the UX state they induced. The information thus collected was used as markers indicating which frames to extract from the videogrammetric data to present as immersive 3D models (photogrammetric). A further spatial exploration and evaluation then took place inside the social VR environment of Hyve-3D (Fig. 3) through collective interviews where participants explained together their UX for each interactive setting selected.

This case study, presented here as a proof of concept of the proposed framework, develops an assessment of three users (3 researchers from the lab) engaging for the first time different interactions within a recently renewed museum dealing with different engineering and design aspects from the development of a company’s flagship vehicles (the name of the company is withheld). In the context of this pilot project and considering its limited sample size, our contribution relates itself mainly to a new framework that can be used for guiding the design and evaluation of interactive installations in museums.

**RESULTS AND ANALYSIS**

The granular UX assessment data was visualized as graphs to facilitate the identification of key moments of UX that should be explored in depth. The graphs (Fig. 4) were composed of opposing curves corresponding to the perceived challenge and skill levels, to which was overlaid a four levels line graph illustrating the evolution of the person’s UX, the bottom indicating boredom. A brief review of the original video footage was conducted in parallel with an exhaustive numbering of the exhibition room’s installations. This made possible to annotate each participant’s graph with the installation concerned at every moment of the visit. A revision lead us to isolate what the data showed to be the most engaging in-
interaction moments (flow state) as well as controversial installations rendered by contrasted experiences oscillating mainly between stress and flow. To lighten the 3D photogrammetric model building workflow, a further refinement of the selection was based on the plurality of experiences - i.e. the number of affected visitors - encompassed by the precise key frame and its neighboring moments. Scenes extracted accordingly are presented in Table 1.

The hypothesis behind the selection criteria is that memories from the key moments of these UX state changes should give us an insight not only into the environmental variables supporting the optimal experience, but also on the reasons leading to their interruption, shedding light on eventual recommendations. The scene occurring at $t = 270$ sec. (see Fig. 4) was discarded due to the camera angles and the participants’ positions, making it impossible to obtain

<table>
<thead>
<tr>
<th>Installation</th>
<th>6</th>
<th>9</th>
<th>11</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>70</td>
<td>100</td>
<td>155</td>
<td>165</td>
<td>185</td>
</tr>
<tr>
<td>Visitor 1</td>
<td>Flow</td>
<td>Stress</td>
<td>Stress</td>
<td>Stress</td>
<td>Stress</td>
</tr>
<tr>
<td>Visitor 2</td>
<td>Flow</td>
<td>Flow; Stress; Control</td>
<td>Stress</td>
<td>Flow</td>
<td>Flow; Stress</td>
</tr>
<tr>
<td>Visitor 3</td>
<td>Stress</td>
<td>Flow; Control</td>
<td>Stress</td>
<td>Stress</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 2
Brief descriptions of the selected installations.

<table>
<thead>
<tr>
<th>Installation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>A fully functional vehicle engine (modified to be powered by electricity) displayed in a transparent acrylic casing. It is equipped with a fixed replica of the vehicle’s handlebars, at the counter’s height, as an interface to run the engine using a single push-button. The handlebar has multiple fictitious buttons for decoration.</td>
</tr>
<tr>
<td>9</td>
<td>Complete suspension mechanism used for the traction unit, encased in a similar transparent cabinet as installation 6. The components are connected into a single longitudinal piece that is secured on two articulated pistons situated at both ends. Two vertical handles are installed side-by-side on the counter, in front of the casing, allowing visitors to observe the component’s various behaviours by pushing and pulling horizontally each handles independently.</td>
</tr>
<tr>
<td>11</td>
<td>Three different versions, from different years, of a part of the vehicle used that attaches itself to its bottom for ground support and vehicle motion. Parts differ greatly in terms of proportions and materials while having a similar shaping and recurrents details.</td>
</tr>
<tr>
<td>14</td>
<td>A complete and real vehicle mounted above head height and further back on an angled surface that extends across the periphery of the exhibit room.</td>
</tr>
<tr>
<td>15</td>
<td>A coarsely textured traction unit fixed [no moving parts] independently on the counter.</td>
</tr>
</tbody>
</table>

a properly reconstructed 3D model, even if it corresponded to the criteria (as illustrated in the graphs in striped lines). Although, few spatial properties were differentiating it from the selected instant occurring a few seconds later at $t = 275$. As participants were confronted with an extended temporal span of their behaviors while watching the video, they had a chance to reflect openly on the omitted moment. After reviewing the recordings of the interviews, we extracted the main ideas underlying the accounts from each participant.

**Installation 6**
Of the selected scenes, installation 6 came out as the only one eliciting the flow state for all three visitors: The most important reason is the multisensorial aspect of the setup (distinguished sound when interacting with). This factor was remembered as soliciting a heightened engagement: “When others use this interaction, the sound of it creates interest for me” (Visitor 3); “What attracted me to that interface was the sound [Visitor 2] was doing, [...] I was looking and I wanted to do it myself to feel it” (Visitor 1). For another participant, it was the possibility of visually exploring the moving parts from different angles and discovering “complexities inside the moving engine” (Visitor 2) that was felt as an engaging feature, thanks to the object being three-dimensional, close and fully visible through a transparent casing. Yet, a stress state was induced due to shaking upon use, interrupting the flow state. Perceiving this instability, a participant claimed to feel a certain risk that lead him to self-constrain his interactions: “I got scared at some point when it started spinning fast, and I didn’t want- the box didn’t look that solid” (Visitor 2). The same visitor noted the sound he perceived as inauthentic created a disinterest tipping him further out of his engagement. For Visitor 1, inauthenticity was also brought up as a trait hindering the interactive experience. The participant expressed, why it didn’t replicate the handlebars he remembered from using an actual vehicle. For visitor 3, an initial stress state was linked to a poor contextualization of the isolated mechanical component. According to her, the disposition of this installation made it difficult to understand how the engine should integrate itself into an actual vehicle (Visitor 3).

**Installation 9**
For all three participants, the initial apprehension of this installation correlated with a stress state. The main reason was a lack of information discernable at first glance and clarifying the relation of the component and its operation with the broader context of the exhibition. For one visitor, this notable deficiency extended itself beyond the counter-level installation,
to the signage design of the room. He reported that after discovering specific marks spread across the room’s floor, the fact they had an obviously intentional placement and yet seemed irrelevant could have contributed to create confusion: “Only now I realize why this is there. [...] You don’t know when you are there.” (Visitor 1). The perceived affordances were a mitigated component of the installation. One participant detailed how the interaction wasn’t self-explanatory for her to figure out how to use it: “I was stressed because only after I see others using it I can know how to use it. [...] I don't know what’s the function of this device- it's for what” (Visitor 3). However, for another the design of the installation including two large upright handles “really sticking out” (Visitor 2) turned out highly inviting.

**Installation 11**
Here again, all three participants experienced a stress state during a part of their interaction. For Visitor 2, a lack of information properly designed within the actual installation lead to an interruption of his flow state. Indeed, because the elements were simply laid out, some more specialized mechanical details were less prone to being interpreted with certainty: “They show the [specific parts] but they don't detail on what I suppose to understand. I see a [specific parts], okay, but now tell me what is behind the it, what it's actually made of...” (Visitor 2). For another participant, the stress state was induced by a poor contextualisation of the displayed element. Although, by taking the time to turn around and browse the rest of the exhibition, she better understood how the component integrated itself to the vehicles, briefly prompting a state of flow.

**Installation 14**
Interviews suggest that, for two participants who experienced stress or control, this installation made them realise the exhibition’s general setup fell short of clarifying the link between the timeline’s evolution and the occurrence of the displayed technological developments: “I just want to find the connection between two [vehicles]... and I didn’t find it. I know the ‘mock-ups’ are better and better, but I can’t find the link... I want to know which part of the design makes which one better (while pointing at a displayed vehicle)” (Visitor 3). It seems that bad information design failed to efficiently associate the ascending years printed on the counter with the scattered vehicle components: “I realize now that I didn’t attach each object to a specific year, it was just something part of the vehicles” (Visitor 2). Otherwise, the single participant who experienced a state of flow here related mainly to the authenticity of the products that were shown and their general configuration (while gesturing an englobing circular motion with both arms) in the space with “very good lighting” (Visitor 1) as enjoyable.

**RECOMMENDATIONS**
Our findings are in line with the idea that authentic stimuli are perceived as a desirable factor from the visitors’ viewpoint (Levent & Pascual-Leone, 2014) and highlight the strength of a multisensorial UX design of these stimuli. On different occasions, participants recalled how parts of their visit evoked memories compared to which the installations’ per-
ceived authenticity played a definite role in the states reached. At one stage when talking about poor contextualization, a participant pointed out that, “after all, to [him], the [vehicle] is a tool, something humans use, so [the museum] should try to figure out a way to show that to people” (Visitor 2). By taking this into consideration, exhibition designers could help future visitors understand more directly the actual uses of different components, and in some cases, strengthen the link between displayed content and everyday life. In Table 3 we synthesized the analysis as punctual design recommendations for the installations to give a brief example of how the framework put forth could percolate in practice.

**DISCUSSION**

The pilot study hints at the fact that, even with a methodology using 3D models as freezeframes for exploration instead of live videogrammetry, participants’ grasp of memories appeared to overflow the represented moment. Collective interviews inside the social VR system show that participants gained the ability to reflect on what was experienced in the moments both preceding and following the reconstructed instant. We suggest that, by combining the spatial immersive exploration of static settings while switching back and forth with the immersive video including the recorded sound of the surrounding moments, retrospections reached a greater level of detail than they could have prior to this stage. In fact, self-observation interviews showed that participants’ memories of their visit cleared up even though almost three months had passed since it was conducted. Some even openly expressed being surprised at how much they could remember about otherwise forgotten facets of their visit. According to the aspects that were referred to during the interviews, a clearer link traces itself between the reflections shared and the mediatic characteristics of both the immersive spherical video and photogrammetric models. While the sound of various elements was brought up as a constituent part of the exhibition visit, the positions and gestures of other visitors in the recorded images also appeared to have triggered deepened retrospection. This observation points toward the particular potential of social VR environments and fully tridimensional event depiction as a productive interview context. The anchoring of the collective interviews on specific instants, as they were identified by the UX granular assessment, also helped to structure the workflow. Not only did they give participants a starting point to help them verbally express their experiences, they also offered researchers a tool to organize the interviews and analysis.

**CONCLUSION**

The primary aim of the study was to experiment a new framework for assessing UX in interactive spaces. By conducting a case study in a museum context, we managed to provide a more rigorous picture detailing the impact of using a methodology based on video-photogrammetry and immersive interviews. The study showed that immersing interviewees in interactive 3D photogrammetric models depicting frozen moments of lived events facilitates the recall of memories and allows a deepened self-observation analysis. Future work should try to clarify the link between the quality of the spatial and embodied experience provided by the immersive environment and the details and aspects brought up by participants. More tests could also be conducted to include 3D sound ambiances during the stage when participants navigate through the 3D models. This work also served as an actual eval-

<table>
<thead>
<tr>
<th>Installation</th>
<th>6</th>
<th>9</th>
<th>11</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design recom-</td>
<td>Be</td>
<td>Clear</td>
<td>Better</td>
<td>Clear</td>
<td>Reachability</td>
</tr>
<tr>
<td>mendations</td>
<td>authentic</td>
<td>signage</td>
<td>information</td>
<td>configuration</td>
<td></td>
</tr>
<tr>
<td>Avoid Shaking</td>
<td>Better</td>
<td>information</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contextualize</td>
<td>Better</td>
<td>affordances</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
evaluation of a designed exhibition, an empirical terrain through which some of the findings corroborate certain propositions from museum design literature. Although the design recommendations here were brief and haven’t been thoroughly delved into by an actual group of designers in a professional context, future explorations could attempt to focus on the use of 3D sketches also provided by Hyve-3D as a communication and ideation tool in steps following a similarly structured UX assessment protocol.

REFERENCES
Gibson, JJ 2014, Approche écologique de la perception visuelle, Dehors, Bellevaux