C-space

Simoes, Bruno; Aksenov, P.; Santos, Pedro; Arentze, T.A.; De Amicis, Raffaele

Published in:
2015 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2015, June 29 2015-July 3 2015, Turin, Italy

DOI:
10.1109/ICMEW.2015.7169833

Published: 28/07/2015

Document Version
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
ABSTRACT

A key theme in ubiquitous computing is to create smart environments in which there is seamless integration of people, information, and physical reality. In this manuscript, we describe a set of tools that facilitate the creation of such environments, e.g., a service to transform videos recorded with mobile devices into navigable 3D scenes, a service to compute and describe the emotional processes that occur during the user interaction with such content, a service that takes into account certain dynamic needs of users in personalizing solutions for allocating their leisure time and activities, a gamified crowdsourcing application, and a set of projection-based tools for creating and interacting with augmented environments. Ultimately, our objective is to have a framework that seamlessly integrates all these components, to foster creativity processes.

Index Terms— 4d, affective, smart routing, projection

1. INTRODUCTION

The idea of creating a digital replica of our planet that can be used to "go back in time" and then see how the planet looked in the past, or that can be used to experience the world as it looks now or as it might look like in the future, was initially introduced by Al Gore [1] in 1999. To some extent, the concretization of this idea can facilitate the creation of new forms of smart environments since it can function as a well-structured and intuitive spatial-temporal repository. However, market implementations felt short on his vision [2,4,7]. The lack of low cost technology to capture three-dimensional models and the high cost in terms of human resources were probably the main limitations.

In this manuscript we introduce framework that aims at easing the aforementioned limitations. On the one hand, our framework fills a gap that is not yet covered by any technology on the market: it reduces the cost and time needed to create 3D models of physical objects and events. Reduction in time is not just twofold but it is of more than one order of magnitude, potentially reducing a 3D production process that would take days, if not weeks to complete, to few hours. On the other hand, it delivers a set of tools that aim at boosting user engagement and at turning the real-world into a "virtual stage" for ubiquitous media-sharing. These tools combine strategies based on user emotional processes and needs, to create environments that seamlessly integrate people, information, and physical reality.

2. C-SPACE FRAMEWORK

In this section, we describe the objectives of the c-Space project and we provide an overview on its technical framework.

The first objective of the c-Space project is to facilitate the reconstruction of large and dynamic environments, by leveraging the trend of Internet-based sharing of "casual" videos recorded with mobile devices. With that in mind, we developed a web API and a mobile client that can be used to upload pictures and videos to a cloud service that creates, in nearly real-time, a reusable 4D model of the real scene.

![Fig. 1. The c-Space modelling pipeline.](image-url)
After the 3D reconstruction, the new spatial-temporal model is analyzed, optimized and then made available for user distribution.

The second objective of the c-Space project is to promote engagement as the strategy to decrease the time required to acquire a minimum number of videos (and pictures) for the reconstruction, and to increase user spatial coverage. As a solution we adopted a produsage model where users are both producers and consumers of the content generated [5]. Two strategies were adopted to create an interactive framework that capitalizes on user interest, which is essential for the success of the model [4]. First, we use gamification and serious games to engage users into video crowdsourcing activities. To make the framework flexible, we designed the framework to support user-customizable levels. Hence, we can promote or prioritize video crowdsourcing in specific areas. Last, we integrate somehow similar ideas into our domain-based applications. For example, in cultural heritage, we provide personalized smart routing and recommendation systems: we give tourists access to content and augmented experiences, otherwise not accessible, and in exchange they share some of the videos and pictures they took during the trip.

Our last objective is to deploy user-friendly tools for content access and creation. The strategy for supporting new ways of interaction with content is: to deploy an adaptive and affective content module that aims at reducing content searching times; to create interactions based on projected content (in addition to traditional augmented reality), so interactions with 3D content are similar to interactions with objects in the physical world.

In the next sections, we describe the five components of our framework: video-based 3D reconstruction, gamified crowdsourcing experiences, projected-based interactions, affective computing, and smart routing.

2.1. 3D reconstruction of dynamic scenes

Our 3D reconstruction framework aims at filling a gap not yet covered by any technology on the market. First, our framework handles the reconstruction of a 3D-scene from heterogeneous inputs, including high-quality video from professional video cameras, still images captured with digital photo cameras, and video sequences recorded from smart phones. Last, our 3D reconstruction framework serves as an input to new types of interaction that move from simply watching individual videos, as currently possible on YouTube, to enjoying complex events such as sports events and concerts in a dynamically reconstructed 4D way.

The scientific challenge that our framework has to sort out is that one of reconstructing dynamic objects from videos and pictures. The reconstruction of dynamic objects is much more difficult to tackle than the reconstruction of static objects. First, we have to split up video streams into single frames at finest granularity, such that there is a frame for every discrete time step at optimum quality. Second, we remove all irrelevant frames. Third, we sort the remaining ones along the time axis. Last, we identify scene’s dynamic objects and then we analyze the kinematics of their moving parts.

2.2. Gamified crowdsourcing strategy

In c-Space, we use crowdsourcing as a strategy to decrease the costs associated with the process of data collection. The current challenge is that existent literature does not provide a good understanding on how to best motivate massive crowds to participate. There are however two strategies that are widely used to influence user behavior towards doing a specific action.

The first strategy is to design a serious game. Serious games are games that are designed to motivate the user to do specific actions. Consequently, serious games do not have entertainment as their initial design purpose. The second strategy is gamification. The goal of gamification it to exploit the user context in order to provide motivation that is specific to the situation. The difference between serious games and gamification is still up to debate, and many serious games use the modes of gamification and vice-versa.

To maximize the adoption and the results of our framework, we implement both strategies. With serious games we target users that are more likely to enjoy games. In our game, users have to take pictures in order to collect digital creatures that are hidden at physical locations, or to complete specific levels to claim reward points. Our gamified application is tailed specifically for tourists. The application suggests, when we believe that is timely appropriated, activities that are relevant to user’s location and context.

2.3. Affective computing capabilities

The affective computing module consents to detect and convey emotions to users, among other features. This module is not intended to be used as a standalone system. It is designed to be integrated into creative applications in order to guide users through the creative process, by providing a better understanding of their cognitive process through the analysis of their emotional reactions. For
example, this is very important in order to capture and maintain their attention. Moreover, it facilitates also the creative process by reducing content search times and by providing feedback on the quality of the content.

2.4. Projection-based interactions

The use of projective technology has been studied for many years. Therefore, its use is somehow not new. The novelty of our approach consists of two main aspects: how we adapt the image that is projected to the physical properties of the environment; and how we can interact with content projected by devices, such as smart phones with an embedded projector, camcorders that can project their videos, digital cameras that can project their photos, and stand-alone mobile projectors with and without built-in memory.

![Projected augmented reality concept.](image)

Although projection devices can afford interesting applications, the level of interactivity with such technology is, at the moment, somewhat pedestrian. The way most applications work is similar to the idea of an ordinary projector connected to a laptop. In many applications, the image that is projected is actually a clone of the image of the device controlling it, and the user interaction is basically limited to the interaction with the device itself.

This implies that the display space and the user interaction remain separate; information does not adapt to the shape of physical objects and users cannot interact with the projection itself. Typically, the interaction with the device (e.g., proximity, using tilt gestures, etc.) or on the device (e.g., tap, touch, etc.) has direct influence on the display space and not on the projection space.

Most of the cases, the main advantage of using projections is that the projected image can easily be shared among a multiple of users. Hence, the user interaction with the projection itself is somehow limited to translation, rotation and scale, as the situation requires.

But, there are more exciting ways of using content projections. Our objective is to use computer vision to decouple the display of the mobile device, for instance, from the projection, and to facilitate the creation of smart environments where users can interact with physical objects that are interconnected with 3D content.

2.5. Personalized smart routing

In planning a trip, often there is an assumption that tourists’ preferences for activities depend on static attributes of the person and choice alternatives. This stands in contrast to a central notion of need-based theories of choice behaviour [6] which state that individuals’ motivations to engage in certain activities are at least partly dependent on needs that change depending on previous activities.

![A user’s preference for an activity depends on dynamic needs that change depending on previous activities.](image)

The added value of “personalised smart routing” is in that it aims to take into account tourists’ experiences at a destination in real time and to personalize each suggested cultural program (tour) in a “smart” way, from the perspective of both the cultural and the overall experience of taking it (Fig. 4).

In order to take into account the dynamics, we introduce a new model, named LATUS (Leisure Activity Travel Utility Simulation) that deals with leisure activity and travel choice specifically. The model’s central notion is that process utility derived from leisure activities emerges from the satisfaction of the needs people have and it generates positive and negative substitution relationships between activities contributing significantly to the value of an overall tourist experience.

The development and evaluation of the model are supported by a number of stated choice experiments specifically designed to measure the relative impact of dynamic needs and variety seeking on tourists’ activity choice. The analysis of the results provides the key parameters of the LATUS model.
Real-time affect recognition can be used to adapt personalized recommendations of attractions based on the affects they cause.

Besides, one of the objectives of the c-Space framework is to leverage on user-generated content, such as videos and photos taken during the journey, and to offer 3D-reconstruction of scenes of cultural heritage and tourist attractions. From the perspective of the tourist experience, this feature offers an additional source of information about a tourist’s dynamics, which this time takes the form of affective feedback collected during the tourist’s interaction with the aforementioned 3D-content. And one of the directions to include this affective feedback into personalising cultural program recommendations to tourists is to create affective maps of places (Fig. 5), where each attraction and its collocated activities correspond to a (set of) particular affective response(s), so that the recommendation process takes into account the current and/or the desired affective state of a tourist.

3. CONCLUSIONS

In this manuscript, we describe a framework that facilitates the creation of smart environments. It includes a service to transform videos recorded with mobile devices into navigable 3D scenes; an affective computing module that consents to detect user emotions that occur during the user interaction with such content; a service that takes into account certain dynamic needs of users in personalizing solutions for allocating their leisure time and activities; and a set of projection-based tools to augment the surrounding environment as well as to facilitate creativity processes. The sustainability of our framework is briefly described.

4. ACKNOWLEDGEMENTS

This research has been supported by the European Commission (EC) under the project c-Space (Grant Agreement N. 611040). The authors are solely responsible for the content of the paper. It does not represent the opinion of the European Community. The European Community is not responsible for any use that might be made of information contained herein.

5. REFERENCES