Modeling Places for Interactive Media and Entertainment Applications

Rui Nóbrega  
CITI, Faculdade de Ciências e Tecnologia,  
Universidade Nova de Lisboa  
2829-516 Caparica, Portugal  
rui.nobrega@di.fct.unl.pt

Abstract  
Taking advantage of the multitude of cameras now available and capable of recording all aspects of our lives, this work explores the notion of virtualizing a physical place using cameras and sharing the resulting model with others. This social sharing would create new forms of relationship and common space discovery that would enhance video chats and virtual visiting of physical places. Furthermore, the research will consider the possible interactive applications, from games to augmented reality, which can take advantage of the created spatial and temporal models.

Keywords  
Computer Vision, Augmented Reality, Virtual Reality, Social Sharing

ACM Classification Keywords  
H.5.1 Multimedia Information Systems - Artificial, augmented, and virtual realities

General Terms  
Design, Experimentation, Algorithms

Introduction  
Cameras are everywhere; usually they serve a specific purpose, which is well defined by where we can find them. Camcorders are used to record memories, cell phone cameras are usually to take pictures and videos of spontaneous interesting moments, and webcams are essentially used for video chat. With the rise of the new interaction mechanisms in video consoles, cameras start to appear in the living room and other unexpected places. Would it be possible to use this visual acquisition power to integrate services in interactive applications where the user can share visual content to the world seamlessly and effortlessly? Can a camera record and acquire the room where we live and share it to the world? Far from the fears of a Big Brother, it would be possible for a person to show the new decoration of his/her room, alter a model of the same room, create augmented realities in virtual space or time, and run applications and games that mix reality with custom content.
What is proposed here is a framework to empower users to capture their worlds and spaces, and use them in a wide range of interactive applications. In the same way as applications such as Google Street View capture the outside world, here users can do the same indoors, controlling and manipulating what is shown and shared with others.

**Background and Related Work**
The use of cameras and the state-of-art of computer vision and Simultaneous Localization and Mapping (SLAM) algorithms are currently summarized in [4]. The use of a large collection of images to create a coherent model relating the overlapping of each picture was done in the Photo-tourism project [7]. Most of the techniques used depend on photo-stitching, most of them based on SIFT descriptor as in the panorama construction algorithm described in [2]. Taking advantage of these techniques several applications have been made to produce interactive results that can be manipulated by regular, untrained users. The project presented in [8] allows the real-time creation of a full panoramic scene from the cell phone. The PTAM project [5] goes further and interprets the space where the user is and creates augmented reality objects that interact with real world elements in 3D. Another important work is [6] where a panoramic image of someone or an object is created, altered and the alterations are reflected back in the original images and videos that created the panoramic. In terms of interface, [3] presents an application where the user can directly manipulate the frames, not following their natural sequence, but the desired direction of where the user wants the objects to go. In [1] the notion of rephotography is explored, where old photos can be compared with recent photos using the same techniques described above.

**Extending Vision as an Input Method**
The main research questions are related with the creation of spatial and temporal models of a certain physical place from images and videos and the ability to provide an intelligible and interactive form of manipulating and sharing this information. Specifically, the addressed research topics are:

- Data acquisition, image acquisition from webcams or mobile devices, offline for additional processing and in real-time for interaction [2,4,7].
- How to create a virtual model of a specific room? Using 3D reconstruction [7], panoramic scenes [2], and/or specific features detection [5]?
- How to add virtual content to the above model? One option is using markerless augmented reality [8].
- How to show [3] the model? How to share with other users or social network.
- What kind of interactions between users can arise from the sharing of a model of the space. Will others comment, modify or copy your space?

These are very diverse topics, that can be summed up into the problem: how to live and interact in someone else’s space without being physically there.

**Virtualizing and Sharing Your Space**
In order to address the proposed research statement, several principles and development strategies must be considered as presented next.
Space Acquisition
The data acquisition should be done using an available camera, such as a webcam. The user would go into the desired place and wave the camera around freely, capturing images from all desired spots of the room with some redundancy going back and forth. These images can also be taken at different moments of the day or even in different days in order to record a temporal history of the place. Free camera moving poses some problems and some restrictions [5] may need to be applied, but for now the more general problem will be considered. Here the main goal is to acquire a large set of images to work with.

Data Processing
In order to create better exploratory models the raw images should be processed to explore their relations and homographies. Using photo-stitching techniques [2] it is possible to pair each image with their similar counterparts. Knowing this, a graph can be created relating all the images spatially and temporally. Using this graph the goal is to construct three models: (1) image flow graph navigation, (2) partial 3D reconstruction and (3) panoramic view. The (1) flow model would allow browsing the space following the graph; this means that the order in which the user sees the images can be completely different from the order of how they were taken. Using the camera movement, namely translation, 3D characteristics (2) can be obtained such as, object depth, plane surfaces, detecting walls and floor and detecting the main lines that compose a scene. Finally, a large panoramic view (3) can be obtained by merging several images. These panoramic images can be realistic or distorted in order to accommodate weird points of view that the user may wish to add. All these models, could be browsed considering different temporal moments.

Interaction and Social Sharing
This is the main aspect that I will be concerned in my research. Having a rich visual model is only as important as the means to interact with it. With the model captured and analyzed the user should be able to undertake the following actions:

- Add/Remove content. Remove unwanted parts and add virtual objects or labels.
- Play games in that space, taking advantage of physical characteristics such as planes or surfaces.
- Create multiple alternative versions of the space.
- Share the content online for others to see.

In the same way, friends or other people that can access the users’ model online can do the following:

- Browse the users’ room. Check out the changes, how it was and how it will be.
- Virtual navigation of an interior space composed of several places.
- According to a permission system, it would be possible to comment, add or remove content.
- During an online video conversation, it would be possible to view the users’ room even if the camera was not pointing to it.
- Play games in the other user room.

In current social systems users can share photos, maps and comments. What is proposed here is the sharing of a physical location in a virtual model. Common video interaction is usually limited to where the camera is
pointing; these methods would allow additional interaction and discovery, adding playful elements.

**Expected Contributions and Conclusions**

The current ongoing research was essentially centered on the system requirements, and studying and developing the set of tools (as seen in Figure 1) that support the final interactive applications scenarios described.

The final goal of this project is to research the means so that users can engage in space sharing and interaction through social networks. The expected contributions of this work are the development of a framework that empowers the user to virtualize a certain place and share it online. Additionally it would be interesting to investigate the social implications of instantly sharing intimate places. How would it affect our privacy? Would it be just another level of socialization complementing images and micro-blogs?

In summary, the contributions will focus on the development of interactive methods for place manipulation and study the implications of these social interfaces.

---

**Acknowledgements**

I would like to thank CITI/FCT/UNL, especially the Interactive Multimedia Group for their support. This work was funded by FCT/MCTES through grant SFRH/BD/47511/2008.

**References**


Appendix 1 : Statement of Expected Benefits

From my participation in this doctoral consortium I hope to have an open discussion about my PhD research topic with other researchers, where I can get feedback on what is relevant to the HCI scientific community, and achieve some form of peer validation. I would very much enjoy meeting and talking with other people and experts that work in the same research field. I expect criticism and suggestions and look forward to contribute with my opinions and experiences to the other consortium participants’ work. Additionally I can volunteer to help in the organization of the doctoral consortium.

Rui Nóbrega
Appendix 2 : Supervisor's Letter of Support

I have supervised Rui Nóbrega for three years during the MSc and several research projects. Our last collaborative project was a multitouch table and software that was successfully used in a major art exhibition by more than 50,000 people at the Berardo Museum/CCB in Lisbon.

Rui Nóbrega is currently starting the second year of research (prior to that he did the Doctoral Program courses with very good marks). He has at least two years ahead before finishing the PhD. Currently, his work is in a stage where he has done related work and built tools that will lead to produce innovative scientific results.

In this way it is an interesting year to attend the Doctoral Consortium because he had done all the foundational work but there are still many options and directions that can be chosen.

Rui Nóbrega is quite resourceful and sociable, and will certainly interact with the other participants of the consortium. He has previous experience in organizing conferences and scientific events (e.g., Mobile HCI conference). He also has a wide variety of interests and experience in human computer interaction that would certainly benefit the other participants of the consortium.

Nuno Correia  
CITI, Faculdade de Ciências e Tecnologia,  
Universidade Nova de Lisboa  
2829-516 Caparica, Portugal  
cmp@dl.fct.unl.pt

Interactive Multimedia Group Coordinator  
Director of the Digital Media Program in collaboration with UT Austin.
Appendix 3 : Rui Nóbrega’s CV

Rui Pedro da Silva Nóbrega

Born: April 11th, 1983, Lisbon, Portugal
Emails: rui.nobrega@di.fct.unl.pt,
rui.nobrega@gmail.com
Webpage: http://img.di.fct.unl.pt/~rpn
Ph.D. Blog: http://doutorandoemfilosofia.blogspot.com

Currently I am a student researcher at the Interactive Multimedia Group (http://img.di.fct.unl.pt), which is the HCI part of the research centre CITI. (http://citi.di.fct.unl.pt). I have a B.Sc. (2006, grade 16 in 20) and a M.Sc. (2008, grade 18 in 20) in Computer Science Engineering from the Faculdade de Ciências e Tecnologia (Faculty of Science and Technology) of the Universidade Nova de Lisboa (University) (http://www.fct.unl.pt). Currently I am in the middle of a 4 year Computer Science Ph.D. (2 years finished, but the first year is a Doctoral Program where students experiment several topics and choose their research area). My topic is related with HCI. My Ph.D. is financed by the Portuguese Foundation for Science and Technology (http://www.fct.mctes.pt), and with some contact points with the Portugal-UT Austin program Digital Media. I was recently involved in the organization of Mobile HCI 2010, which was held in September 2010 in Lisbon. The next sections summarize the work done in the last three years of research.

Research Projects

Life-Saver - March 2007–September 2008 create a simulator to simulate emergency flood plans in river dams in conjunction with a visualization system with interaction through multi-touch tables. – http://img.di.fct.unl.pt/lifesaver

ArtTouch – March 2009–September 2010 - create a large scal multi-touch table and application suitable for a large scale exposition inside a museum. – http://img.di.fct.unl.pt/arttouch

Publications

Professional Experience
2006/2007 IT Consultant - 1 year
Novabase – Advanced Custom Development
http://www.novabase.pt
(This is one of the largest IT consulting companies in Portugal)

Teaching Assistant
Activity between March 2007 and July 2010, on FCT-UNL (http://www.fct.unl.pt)

Introduction to Computers and Programming
(Introductory programming course in MatLab to Engineers) – 3 Semesters

Geographic Information Technologies (Map related technologies, 1st year Computer Sc. M.Sc.) – 1 Semester