Dynamic Insertion of Virtual Objects in Photographs

Rui Nóbrega, CITI and Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal
Nuno Correia, CITI and Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

ABSTRACT

Introducing virtual objects in photographs or video sequences presents several challenges, such as the pose estimation and the visually correct interaction boundaries of such objects. In this article a framework for the introduction of virtual objects in user-captured photos is discussed. Furthermore, the introduced virtual objects should be interactive and respond to real physical environments. The proposed detection system is semi-automatic and thus depends on the user to obtain the elements it needs. This operation should be significantly simple to accommodate the needs of a non-expert user. The system analyses a photo taken by the user and detects high-level features such as vanishing points, floor and scene orientation. Using these features it will be possible to create virtual mixed and augmented reality applications where the user takes one or more photos of a certain place and interactively introduces virtual objects or elements that blend with the picture in real time. This article discusses the techniques required to acquire images and information about the scenario involving the user. To demonstrate the framework, a proof-of-concept implementation is presented. This implementation was used to conduct a user study regarding the evaluation of the reliability of the concept. The presented results show a high reliability in the scene detection and that users are able and motivated to use this type of systems.

Keywords: Augmented Reality, Computer Graphics, Computer Vision, Photos, Virtual Objects

INTRODUCTION

Most mixed and augmented reality applications use images and videos from the real world as input for the virtual scenario. These images must have certain properties: some have a fixed pre-built scenario, others include fiducial markers as reference. This article proposes an interaction model for mixed reality systems without a priori knowledge of the real world. This means that the applications should integrate with real world scenarios without using predefined markers or objects. Instead, the enhanced reality should be based in image recognition and automatic scene reconstruction. Virtual objects should recognize real world elements and react accordingly to them. It is exactly the combination of image analysis and recognition with the augmented reality concept that is explored in this article. The goal is to enable the user to choose the scenario where the AR application will take place. The starting point for this project is the motivation to build interactive applications that allow the user to recreate and change virtually
a personal space. Possible applications of this system would be important for decoration of rooms, rearranging house interiors, furniture manipulation, games, 3D modeling and animation systems, amongst others.

The main broad research questions addressed in this project can be divided in two parts. First, is it feasible to create augmented reality applications where virtual objects, characters and elements, live and interact on a user chosen (and not pre-defined) real world scenario? Secondly, are untrained users capable and motivated enough to perform the necessary steps to initialize the AR environment? Altogether, is there a future for interactive markerless AR applications? To answer these questions a computer vision system, which analyses pictures taken by the users, was developed (Nóbrega & Correia, 2012a). More than building a system for perfect photo-realistic 3D reconstruction, the intention is to construct a sufficiently robust model to be used by a non-expert user. The presented system was tested with several users in different environments to evaluate the reliability of the interaction.

To showcase the capabilities of the AR system, a game was implemented as a working proof-of-concept. The chosen game was a classic two-player snake game where players must avoid at all cost bumping into the boundaries of the game or any of the snakes itself. The main novelty here is that the game can be played in any real-world flat surface that the players can imagine. This can be a table, the living room floor, the backyard or the main street. Additionally, the users can built their own custom snake maze using sheets of paper or any other uniform material.

The initialization of the game requires taking a picture of an open uncluttered space. The user is then guided through a simple process where the game level is built. In a first stage, the user roughly selects the area of the game in the picture, then the picture is analyzed to build the corresponding internal 3D model, and finally the game is superimposed to the picture.

From now on, we will refer to the model that contains all the detected features including the 3D detected data from the image as just the model. This is just the information used to superimpose objects and should not be confused with a rendered model of the world.

In the next section, several related projects and technologies are listed and described. Following, the mixed and augmented reality system is explained with several algorithms that enable the virtual interface being detailed. Afterwards, the showcase prototype interface is presented along with a user study, which evaluates the system. In the end, several conclusions are drawn and perspectives for future work are presented.

**RELATED WORK**

Superimposing virtual information on top of real images and video is a concept that has been around for many years (Azuma, 1997). One of the main problems in AR applications is finding location, allowing to know what kind of information should be displayed on a given part of an image (Tillon & Marchal, 2011).

Most traditional AR applications depend, for location awareness, on fiducial markers, QR codes (Turcsanyi-Szabo & Simon, 2011) or GPS tracking. Fiducial markers (e.g., ARtoolkit (ARToolKit, 2003)) are algorithmically simple and fast to track and provide three-dimensional information of a certain spot. They are ideal for virtual insertion of 3D objects in a scene.

GPS tracking requires that the camera should be paired with a GPS device and most modern smartphones already offer this functionality. Several applications (LayAR, 2012) that currently use GPS use a “bubble” approach (Takeuchi & Perlin, 2012), where the information is presented in bubbles near the GPS spot.
Mobility, Liminality, and Digital Materiality


[www.lgi-global.com/chapter/mobility-liminality-digital-materiality/73630?camid=4v1a](www.lgi-global.com/chapter/mobility-liminality-digital-materiality/73630?camid=4v1a)