EXPLORING 3D VIRTUAL ENVIRONMENTS THROUGH OPTIMISED SPHERICAL PANORAMA NAVIGATION

Andrea Ferracani, Lea Landucci, Pietro Pala

Università degli Studi di Firenze - MICC
name.surname@unifi.it

ABSTRACT

Commonly, immersive and virtual reality systems simulate real environments exploiting 3D computer graphics. This entails a considerable work to be done in models development and objects textures mapping in order to obtain a good degree of realism. Furthermore, the excessive complexity and the high rendering quality of the models can compromise system performance, especially in a web environment. This paper describes a vision based approach which allows a user to immersively navigate a real cultural environment through a lightweight web based system for the interactive walk-through and browsing of an ordered sequence of spherical panoramas.

Index Terms— Virtual reality, Sub-scene matching, Spherical images, Interaction design, Web application.

1. INTRODUCTION

In the last few years web browsers have been providing an increasing support to third dimension technologies, although 3D rendering engines are not fully ubiquitous and there are still some issues regarding the performance and realism of these systems. Many researchers have then focused on panoramic images due to their attractive and immersive display usage, first of all focusing on how to build a panorama starting from image sequences (two or more) [2], and then addressing the panorama navigation issues. Several works exist which use different mapping representations for outdoor panorama navigation, for example Google Street View [1]. However the navigation metaphor is often more intended to give users a good experience of exploring one panorama (panning, zooming and titling), rather than to optimise the switching from one panorama to another. Our goal is to enable smooth transitions across panoramas so as to reduce their perceived discontinuity and to propose an innovative interaction metaphor for a better multimedia fruition including 3D models, PDFs, galleries of images/videos and indoor panoramas related to the Point Of Interests (POIs) distributed along a cultural walkthrough.
The paper is organised as follows. We discuss the image-based method for panoramas transition in Section 2. Section 3 presents the web interface and its interaction design metaphor. Section 4 outlines conclusions and future work.

2. OPTIMISED SPHERICAL PANORAMA NAVIGATION

A spherical panoramic image is created by warping the radially undistorted perspective images onto a unit sphere assuming one virtual optical center in which the user is supposed to watch at the panorama. Navigation within a single panorama is provided by two main actions: dragging and zooming; navigation among different panoramas is performed through a smooth transition with the replacement of the panoramic texture. In order to support the navigation between panoramas the interaction design model uses a metaphor based on the zooming action of the user: the transition takes place only when the current zoom level exceeds a particular threshold (see Sec. 2.1). In order to minimise the gap in the transition between the two scenes, we studied and developed an algorithm that performs sub-scene matching.

2.1. Sub-scene matching algorithm

The sub-scene matching algorithm is exploited in order to find the best zoom factor threshold for triggering the transition to the next (or previous) panorama. Given two images I_k and I_{k+1} we want to find the rectangular sub-region of I_k whose content is most similar to I_{k+1}. The sub-region and I_{k+1} must have the same aspect ratio, although their size is expected to be different. Let us introduce the following notation:

- w_k and h_k are the width and height in pixels of the image I_k;
- the function $Crop(I, x, \Delta x, y, \Delta y)$ returns the rectangular sub-region of the image I identified by the bottom-left coordinates (x, y) and the top-right coordinates $(x + \Delta x, y + \Delta y)$;
- the function $D(I_n, I_m)$ returns the dissimilarity between the content of two images I_n and I_m : zero in the case that the content of the two images is the same and a number greater than zero otherwise. The measure of dissimilarity is invariant to the size of the two images.

Identification of the rectangular sub-region of I_k that best matches the content of I_{k+1} in correspondence to the inward direction is accomplished by minimisation of the following cost function with respect to the three variables (x, y, η):

$$\begin{cases} (x_0, y_0, \eta_0) = arg[\min_{x,y,\eta} D(f(x, y, \eta), I_{k+1})] \\ f(x, y, \eta) = Crop(I_k, x, x + \eta w_{k+1}, y, y + \eta h_{k+1}) \end{cases} \quad (1)$$

The result of Eq.(1) is the triplet (x_o, y_o, η_o) of the variables that minimise the cost function; they are the coordinates of the bottom-left vertex of the sub-region $P_o = (x_o, y_o)$ and the zoom factor $Z_o = 1/\eta_o$ to be applied to the sub-region of I_k in order to match the size of I_{k+1}. Experimentally, we observed that computation of the dissimilarity function $D(I_n, I_m)$ through the distance of the image histograms provides higher effectiveness compared to solutions based on scale invariant local keypoint descriptors such as SIFT. This is mainly caused by the fact that in the general case, the scenes represented in two consecutive panoramic images can differ significantly in some parts, due to severe occlusions that can take place depending on the 3D structure of the scene captured by the panoramic images. To reduce the computation time associated with the minimisation of the cost function and speed up the computation of the histograms we adopted the technique of Integral Histograms [3].

3. APPLICATION INTERFACE

The system has been developed as a web application based on HTML5, CSS3, Javascript and WebGL and has been implemented using the 3D 'open source' library *three.js*[1]. The navigation among spherical images has been carried out through the dynamic replacement of the texture mapped into the sphere, according to user interaction. The update of the texture is done taking into account the optimal zoom factor threshold for the transition (see Sec. 2.1). This can occur when the user zooms in the outward direction and exceeds the threshold pre-computed by the sub-scene matching module. Otherwise, when the direction of zooming is not aligned to the outward direction the rendered image is progressively magnified—and the field of view is reduced accordingly—until the maximum zoom level is reached. The application can be configured in order to show interactive hotspots in the virtual walkthrough. Each hotspot can be activated by users and it provides several multimedia additional material about specific POIs. Hotspots are constituted by four graphic icons arranged in a circular menu. Each item represents one of the following multimedia categories: 1) gallery of images/videos (yellow icon), 2) PDF (red icon), 3) 3D object (green icon), 4) Indoor panorama (blue icon), cfr. Fig. 1. All these multimedia artefacts are shown inside a floating panel oriented contextually to the panoramic environment and are arranged

Fig. 1. Examples of interaction with an hot-spot: Indoor panorama

in real-time by the system with the best spatial position according to the user point-of-view. In this way, the application overcomes the limitation of standard panorama-based interfaces in which additional content (3D or 2D) is shown in two-dimensional *lightboxes* covering the main navigation area. Furthermore, hotspots can be dragged using an handle at the centre in order to let users better organise the content in the interface view. Context awareness is provided through a mini-map at the bottom-left angle of the screen. Position and orientation are shown on the map and a little circular handle on the route path can be dragged in order to move to different panoramas. POIs, hotspots and associated multimedia materials are searchable using an autosuggest input field.

4. CONCLUSION

In this paper we present a web tool for an immersive interactive walk-through in an urban cultural scenario and we propose a new interface and interaction metaphor for accessing cultural content. An optimisation method for the transition between spherical images based on a Sub-scene matching paradigm is proposed.

5. ACKNOWLEDGMENTS

The research was conducted as part of a project developed in collaboration with F.A.S.T S.R.L. (Fine Art Science and Technology), which provided all the multimedia material, and funded by Regione Toscana.

6. REFERENCES

[1] D. A. at Al. Google street view: Capturing the world at street level. *Computer - IEEE*, june 2010.

[2] O. Cogal, V. Popovic, and Y. Leblebici. Spherical panorama construction using multi sensor registration priors and its real-time hardware. In *Multimedia (ISM), 2013 IEEE International Symposium on*, pages 171–178, Dec 2013.

[3] F. Porikli. Integral histogram: a fast way to extract histograms in cartesian spaces. *In Proceedings of IEEE CVPR, 2005*, 2005.

[1]http://threejs.org/