Video Stitching based on Optical Tracking

Jerônimo G. Grandi and Anderson Maciel
Instituto de Informática (INF)
Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre, Brazil
Email: {jggrandi,amaciel}@inf.ufrgs.br

***Abstract*—Medical imaging is essential for diagnosis. Some image modalities are 2D, others are 3D. A system that integrates both two and three-dimensional visualization modes would allow the physician to still interact with a technique they are familiar, and at the same time, the system would locate and show such structure in the three-dimensional volume automatically, providing other dimensionality of the problem. In this work in progress, we show limited results in two-dimensional space. We present a technique for video stitching, video images captured from different sources in real time are merged into one with increased resolution and view area.**

***Keywords*-video stitching; volume visualization; optical tracking system;**

I. Introduction

Medical images are essential for clinical practice as they support medical diagnosis, surgery planning and physiology studies, for example. These images are captured in several ways. Computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, x-ray and laparoscopic images are the most common. In CT and MRI, after the capture of a set of images, it is possible to generate three-dimensional volume models that are better abstractions of the data set, because they have the same dimensions of the world which we are used to. Doctors are usually better trained to make diagnosis based solely on two-dimensional images and get help from volumetric techniques once in a while, thus losing all the benefits that three-dimensional techniques can provide.

A system that integrates the two-dimensional and three-dimensional display modes is desired. In this way, the physician would still interact with the technique that they are familiar, but in close assistance of the technique that provides other dimensionality of the problem. As the information is correlated between the two techniques the transition among them would be seamless. For example, given a certain real-time image captured by an ultrasound device it is desired to locate such structure in three-dimensional MRI volume automatically.

Objectives: In this work in progress, we present the very first steps to unify the bidimensional and tridimensional techniques. Our initial objective is work in the two-dimensional space, merging two images captured in real time into one with increased resolution and view area, mostly known as video stitching. The main goal of video stitching is to unify in a single frame images captured by many cameras building a larger image with higher resolution. This technique has been used in several domains form security to entertainment. To do that we need to track the capture devices to blend the images into one. In our case we use common webcams for capture and an optical tracking system to track the position of the cameras.

II. Related Work

Research related to video stitching are well consolidated. In this section we will briefly present some previous studies.

In the paper ”An Effective Video Stitching Method” [2], the authors propose a novel method to create wideangle and high-resolution videos from video sequences by assembling the individual overlapping video frames. In the same venue the authors of ”Depth Adaptative Video Stitching” [3] propose an algorithm that utilizes the depth information to compensate current projection transforms for dynamic video content with depth variation. In ”Stitching Videos Streamed by Mobile Phones in Real-Time” [4], it is show an end-to-end system which receives video streams coming from different mobile phones, time synchronizes the streams and produces a single composite mosaic video in real-time. Finally, the ”Image Alignment and Stitching: A Tutorial ”[5] is a complete tutorial for image alignment and stitching that will guide the implementation steps of the system proposed.

None of them use any kind of environment tracking to acquire the position of the webcams in the real world. Using the tracking, the position and orientation of the webcams can be trivially achieved and this can facilitate the image merge process.

III. Optical System

The tracking system being used for development is the Bratrack[6]. Bratrack is an optical tracking system based on reflective markers. It gives as output the position and orientation of the tracked marker. It is possible to track multiple markers simultaneously. The system outputs the data (position and rotation) via network using UDP protocol. In the paper ”Reality Cues-Based Interaction Using Whole-Body Awareness” [7], the Bratrack is used as a technique for 3D navigation and selection exploiting the ”peephole” metaphor with tablet PC.

IV. Cameras

The cameras being used are the PS3eye. These cameras have a high frame rate with good image quality and sensitivity of

Fig. 1. Concept of video stitching. The image of multiple cameras capturing the same scene blend into one when the images are the same in some point. It generates a larger and higher resolution image. Image taken from ”A Winter Day- Video Stitching” [1]

the capture sensor. For this first attempt, we are assuming the cameras will not move in the Z axis. By locking this variable we will always have the same portion of the scene captured, since the cameras have the same lens parameters and hardware.

V. Video Stitching Technique

The idea is to blend the image of PS3eye camera when they are showing the same scene. This concept is shown in the Figure 1. Our approach is to create a point cloud where each point is a pixel grabbed from the captured video frame. As the PS3eye moves, the tracking sends the updated position and orientation to the application and the point cloud moves as well. When the two point clouds (that represent each camera image) overlap, the image is blended, stitching the video frame.

VI. Current results

In this paper we have shown the first steps to build a system that seamlessly allow the physician to work with two visualization modalities, the two-dimensional, direct analyzing of the medical images and three-dimensional with the volumetric view assistance.

Our first step was to work only with 2D images, performing video stitching among them. None of the references searched use environment tracking to know the position of the capture devices. We believe that using this spatial information we can overcome the problems of depth difference between cameras positioned at different distances from the focused object.

What we have done so far, was a search to know what is being published in video stitching field in order to better understand the problem and direct solutions based on the strengths and weaknesses of other early studies.

The implementation part is divided in three main modules: tracking system, video frame grabber and the main application. The first one is the tracking system Bratrack. We have already done the setup of the tracking, the markers for the cameras and it is possible to receive information about positioning and rotation of PS3eye camera via network. The video frame grabber is responsible to grab the video of each camera and store it in a data structure for fast read and generate the point cloud. The main application unifies the other modules above and is based on OpenGL/GLUT API. It receives the position and the image of the tracked cameras and generates two point clouds, one for each camera. The point clouds, that represent the video frame in real time, moves as the cameras move.

VII. Future Work

For future work, we intend to expand the idea of video stitching of two-dimensional images with three-dimensional volumes. For future implementations we will need a volume rendering system that is fast to render the volume in real-time, change the optical tracking system to image processing algorithms that is reliable to find the patterns of a given image and show the region in the volume with the same characteristics.

References

[1] G. Volcovich, “A winter day- video stitching,” 2011. [Online]. Available: http://vimeo.com/25483660

[2] L. Chen, X. Wang, and X. Liang, “An effective video stitching method,” in *Computer Design and Applications (ICCDA), 2010 International Conference on*, vol. 1, june 2010, pp. V1–297 –V1–301.

[3] W. Zeng and H. Zhang, “Depth adaptive video stitching,” in *Proceedings of the 2009 Eigth IEEE/ACIS International Conference on Computer and Information Science*, ser. ICIS ’09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 1100–1105. [Online]. Available: http://dx.doi.org/10.1109/ICIS.2009.129

[4] M. A. El-Saban, M. Refaat, A. Kaheel, and A. Abdul-Hamid, “Stitching videos streamed by mobile phones in real-time,” in *Proceedings of the 17th ACM international conference on Multimedia*, ser. MM ’09. New York, NY, USA: ACM, 2009, pp. 1009–1010. [Online]. Available: http://doi.acm.org/10.1145/1631272.1631493

[5] R. Szeliski, “Image alignment and stitching: a tutorial,” *Found. Trends. Comput. Graph. Vis.*, vol. 2, no. 1, pp. 1–104, Jan. 2006. [Online]. Available: http://dx.doi.org/10.1561/0600000009

[6] F. Pinto, A. Buaes, D. Francio, A. Binotto, and P. Santos, “Bratrack: a low-cost marker-based optical stereo tracking system,” in *ACM SIGGRAPH 2008 posters*, ser. SIGGRAPH ’08. New York, NY, USA: ACM, 2008, pp. 131:1–131:1. [Online]. Available: http://doi.acm.org/10.1145/1400885.1401026

[7] A. Maciel, L. P. Nedel, V. A. M. Jorge, J. M. T. Ibiapina, and L. F. M. S. Silva, “Reality cues-based interaction using whole-body awareness,” in *Proceedings of the 2010 ACM Symposium on Applied Computing*, ser. SAC ’10. New York, NY, USA: ACM, 2010, pp. 1224–1228. [Online]. Available: http://doi.acm.org/10.1145/1774088.1774345