GPU-accelerated Rendering for Medical Augmented Reality in Minimally-invasive Procedures

Matthias Wieczorek[1], André Aichert[1], Oliver Kutter[1], Christoph Bichlmeier[1],
Jürgen Landes[2], Sandro Michael Heining[2], Ekkehard Euler[2], Nassir Navab[1]

[1]Chair for Computer Aided Medical Procedures (CAMP)
Technische Universität München, Germany
[2]Ludwig-Maximilians-Universität München, Chirugische Klinik und Poliklinik
Innenstadt AG CAS, Germany
`wieczore@cs.tum.edu`

Abstract. Recent advances in GPU programmability and performance have enabled development of real-time high quality volume visualization algorithms. Medical augmented reality systems can benefit from these developments. Task-specific visualization aids physicians in better understanding the patient's anatomy and supports navigation of medical instruments in absence of a direct line of sight in minimally-invasive procedures. In this paper we present our results of integration of a hardware accelerated volume renderer into a medical augmented reality framework using a video see-through head mounted display (HMD). The performance of the system is evaluated in an experiment for two human CT datasets. Compared to the literature, our approach allows direct real-time stereo visualization of volumetric medical data on a HMD without prior time consuming pre-processing or segmentation. To further improve the visual perception and interaction of real and virtual objects, the renderer implements a virtual mirror and occlusion handling with the physicians hands and tracked medical instruments.

1 Introduction

Augmented reality (AR) was introduced as an alternative to monitor based visualization for the presentation of medical image data during surgery. The key task of any augmented reality system is to provide its user with good perception of the scene, especially the correct visual cues at the right time. Existing solutions for medical augmented reality often avoid direct volume rendering (DVR) and fall back to simple rendering techniques, e.g. wireframe or single slice display, or reduced rendering quality in order to maintain real-time performance. In addition many methods require time consuming pre-processing of the volume data. More advanced approaches enhance 3D perception within their stereo systems i.e. viewing windows [1] into the body. In the recent years, GPU-accelerated ray-casting has emerged as the de-facto standard algorithm for DVR [2]. ClearView by Krüger et al. improves understanding and insight of the 3D data data by extracting focus and context (F+C) layers on the fly and clever compositing [3]. In [4] the potential of F+C rendering for medical AR is presented. F+C rendering is employed to intelligently embed the virtual scene parts in the video image within the AR scene. In [5] virtual mirror (VM) rendering for in-situ AR is proposed to overcome perspective limitations in navigated surgery settings.

This work summarizes efficient implementations for GPU-accelerated direct volume rendering (DVR) and proposes smart extensions to get the best visual perception while ensuring the required real-time update rate. Important visual cues, such as shape from shading, depth from occlusion or motion parallax can greatly improve user perception of the AR scene.

2 Methods and Materials

In this work we extend and improve the volume rendering and F+C techniques within the AR scene [6], by additional techniques in order to support several new in-situ visualization features. The AR system for optical tracking and the video see-through HMD for in-situ visualization was originally developed by Sauer et al. [7]. Tracking of the objects in the scene is accomplished by two separate optical tracking systems. An outside-in optical tracking system consisting of four infrared ARTtrack2 (A.R.T. GmbH, http://www.ar-tracking.de/) cameras mounted to the ceiling of the room and an inside-out optical tracking system using an IR camera mounted directly on the HMD. Position and orientation of the users head are accurately determined by the inside-out tracking system, while the outside in tracking covers a much greater tracking volume and is used for all tracked objects in the scene. A reference system composed of a known arrangement of optical tracking markers, visible by both tracking systems serves as the common reference frame of the scene. For more details on the used system setup, and calibration of tracking targets (e.g., phantom, instrument), see [6].

In contrast to [6] we rearranged the rendering pipeline. In the first stage a first hit iso surface ray-casting extracts a certain surface (e.g. skin or bone). This hit texture should be extracted before all other rendering passes, because for several rendering techniques this hit texture has to be fully defined and thus not be effected by occlusion handling. In the second stage of the pipeline the hand occlusion is implemented. In the refined version the size of the focus region and hit texture positions are incorporated to improve detection stability. The following stages implement volume rendering and additional effects, e.g. virtual mirror, MPRs (multi-planar reformations). As proposed by Fischer et al. [8] the hit texture is used for occlusion handling for medical instruments. This occlusion handling is again implemented using the OpenGL shading language (GLSL). While the instrument is rendered, the parts of the instrument which are located above the skin, are replaced by the corresponding values read from the video camera image, while the inner parts are rendered traditionally as wireframe models. The individual results of the different rendering stages are illustrated in Fig. 1(b-d).

In order to integrate the paradigm of the virtual mirror [5] into our setup, we combined the virtual mirror and the presented DVR techniques. The mirrored image for the virtual mirror is computed by an additional raycast pass. The computational costs for this additional render pass are kept to a minimum by using the mirror as a stencil for this pass. The usability of the mirror is improved by several features like attachment to an instrument, automatic alignment to the instruments tip, different focus point for the mirror in order to provide the best visual information and linear interpolation based zooming.

Because of the popularity of MPRs amongst medical staff, we integrated additional MPR planes in the visualization of our AR system. We decided to render only the parts of the MPRs which are below the skin, thus the MPR perfectly fits into the F+C visualization. Like the virtual mirror, MPRs can be attached to an instrument as well in order to move and position the MPR within the body.

3 Experiments and Results

We evaluated the performance in an experiment for two human CT datasets, head (296x320x420) and thorax (256x368x522), extracted from the original Visible Korean Human CT data. Bichlmeier et al. [9] created a phantom based on the VKH data to facilitate repeatable realistic experiments. The life-size phantom ranges from the head to the lower hips. Exemplary result images for the head

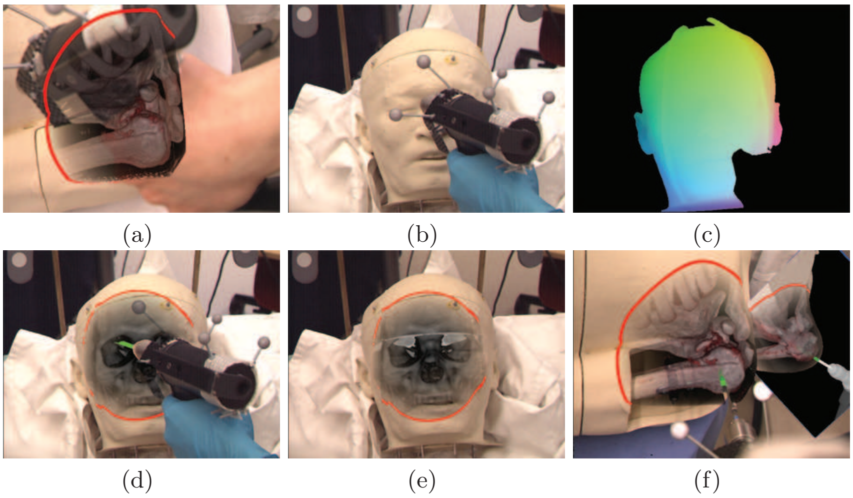

Fig. 1. (a) Illustration of the occlusion problem. (b,c,d) Render pipeline for correct occlusion handling, (b) video texture, (c) hit texture for the skin, (d) final composition of (b) and (c). Video image is used as context layer, Focus layer is rendered with volume rendering. Occlusion handling is shown for instruments and hands. (e) like (d) with in-body MPR. (f) Focus and Context rendering with shaded volume rendering for the focus layer (bone), virtual mirror and instrument.

Table 1. Average rendering performance (fps) results for visible korean head and thorax datasets over 2 minutes. (Pre-Int stands for preintegrated transferfunction) (a) Conventional DVR (b,c) F+C rendering. A video isosurface context layer is always rendered and paired with the same DVR modes as in the conventional DVR evaluation. (c) One virtual mirror and one surgical drill model added to (b).

	(a) DVR		(b) F+C		(c) F+C, VM	
Rendermode	Head	Thorax	Head	Thorax	Head	Thorax
DVR	>30	18	> 30	26	> 30	25
DVR Shaded	>30	8	> 30	12	> 30	12
Pre-Int	27	18	> 30	25	> 30	25
Pre-Int Shaded	13	7	28	16	28	15

and thorax datasets of the visible korean human phantom are shown in Fig. 1(d-f). For the performance measurements the internal render target resolution was set to the camera resolution (640×480), the viewport resolution was set to the HMD resolution (1024×768). Empty space leaping and early ray termination optimizations were activated. A PC workstation with a Nvidia$^{\mathrm{TM}}$Geforce GTX 275 is used to run the renderer. For the evaluation a volunteer wore the HMD and inspected the phantom by continuous motion around it for two minutes. The average framerate was measured using Fraps (http://www.fraps.com/), a freely available benchmarking tool. Note that the maximum framerate of the system is limited by the update rate of the slowest component in the system. Thus, the average framerate can never be more than 30 frames per second, the analog video camera refresh rate, in our experiments.

The measured average framerate for various DVR rendering modes is depicted in Tab. 1(a), for F+C rendering in Table 1(b) and for F+C rendering with virtual mirror and instrument in Tab. 1(c). The better performance of the F+C rendering modes compared to the DVR rendering modes can be explained by the reduced number of pixels for which rays are cast through the dataset when the focus region optimization is enabled.

4 Discussion

This paper discusses the integration and possibilities of GPU-accelerated DVR in a medical augmented reality environment combined with support for virtual mirror, occlusion handling for physician hands and medical instruments and in-body augmentation of medical instruments. The demonstrated techniques are an important step towards the integration of in-situ AR systems to medical navigation applications. F+C volume visualization not only provides improved visual perception, but also aids in maintaining real-time performance, compared to conventional DVR, as demonstrated by the experiments.

Ongoing work is the evaluation of the system in simulated surgical navigated procedures together with our clinical partners. Future work will focus on (i) technical improvement of the system, e.g. real-time depth map reconstruction using the stereo camera setup, and (ii) automatic medical workflow optimized visualization mode selection. Therefore, the specific requirements have to be analyzed for each procedure over the complete workflow, from pre-operative imaging till the end of the intervention. For these dataset the visualization modes will be iteratively refined based on feedback from our medical partners.

References

1. Bajura M, Fuchs H, Ohbuchi R. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. In: Proc SIGGRAPH; 1992. p. 203–10.
2. Engel K, Hadwiger M, Kniss JM, et al. Real-Time volume graphics. AK Peters, Ltd.; 2006.
3. Krüger J, Schneider J, Westermann R. ClearView: an interactive context preserving hotspot visualization technique. IEEE Trans Vis Comput Graph. 2006;12(5).
4. Bichlmeier C, Wimmer F, Heining SM, et al. Contextual anatomic mimesis: hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In: Proc ISMAR; 2007. p. 129–38.
5. Bichlmeier C, Sielhorst T, Navab N. The tangible virtual mirror: new visualization paradigm for navigated surgery. In: Proc AMI-ARCS Workshop; 2006.
6. Kutter O, Aichert A, Bichlmeier C, et al. Real-time volume rendering for high quality visualization in augmented reality. In: Proc AMI-ARCS Workshop; 2008.
7. Sauer F, Khamene A, Bascle B, et al. A head-mounted display system for augmented reality image guidance: Towards clinical evaluation for iMRI-guided neurosurgery. In: Proc MICCAI; 2001. p. 707–16.
8. Fischer J, Bartz D, Straßer W. Occlusion handling for medical augmented reality using a volumetric phantom model. In: Proc VRST; 2004. p. 174–7.
9. Bichlmeier C, Ockert B, Kutter O, et al. The visible Korean human phantom: realistic test & development environments for medical augmented reality. In: Proc AMI-ARCS Workshop; 2008.