Computers & Graphics 53 (2015) 196–209

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Special Section on SIBGRAPI 2014

Focus plus context visualization based on volume clipping for markerless on-patient medical data visualization

Márcio C.F. Macedo *, Antônio L. Apolinário Jr.

Department of Computer Science, Federal University of Bahia (UFBA), Brazil

ARTICLE INFO

Article history:
Received 27 November 2014
Received in revised form
30 August 2015
Accepted 10 September 2015
Available online 26 September 2015

Keywords:
Volume clipping
Focus plus context visualization
Augmented reality
Volume rendering

ABSTRACT

Focus plus context visualization can be used in augmented reality to improve the visual perception of the augmented scene. In the scope of in situ or on-patient medical data visualization, the focus plus context paradigm is used to improve depth perception for physicians showing the patient's anatomy as a focus region in the context of the patient's body. Volume clipping is one technique to realize focus plus context visualization. However, some of the existing methods for focus plus context visualization based on volume clipping do not run in full real time or are prone to artifacts. In this paper, we present an extension for two of these techniques to improve performance and image quality of the original approaches. We validate all the techniques in a markerless augmented reality environment. A 3D reference model is tracked by the application, and volumetric medical data are shown to the user at the position of the patient's anatomy. Our technique is able to handle multiple anatomic regions, although the main region of interest used in this paper is the face. Moreover, tracking accuracy is improved by the use of a hierarchical approach. From an evaluation of the proposed techniques, the results obtained highlight that all of them are free of artifacts, optimized for real-time performance, and improve the visual quality of the augmented scene.

1. Introduction

Physicians see medical data, typically images of a patient's anatomic structures, on a monitor and they must analyze and mentally compose what is shown on the screen. This mental model of the patient's anatomy will help the physician provide health care in time-critical situations. Therefore, the physician must have sufficient knowledge of the patient's and general human anatomy to proceed appropriately during any medical procedure (e.g., diagnosis, surgery). With the availability of augmented reality (AR) technology, one can take over this task of mental mapping by transferring it to a computer. Therefore, the physician will be able to visualize, at the same time, the patient and a part of the patient's anatomy. On-patient or in situ medical data visualization can be used to improve surgical planning, training, medical diagnosis, and post-operative examination. This kind of application is desirable in fields such as those involving craniofacial data, in which the visualization of 3D examinations on the patient may help the physician understand the trauma.

AR is a technology which augments the view of a real scene with additional virtual information. Accurate tracking of the real scene, realistic rendering of the virtual data, and real-time user interactivity are the most important technical challenges of AR applications [1]. The face is a part of the body in which depth- or texture-based tracking is easier because of the availability of face detection algorithms and the presence of distinguishable geometric structures. We take advantage of this to focus on the problem of on-patient medical data visualization for patients with craniofacial traumas. The decision to use a markerless AR (MAR) environment for tracking resulted from observations of the current limitations of the techniques proposed in the field of on-patient medical data visualization. Here, we are mainly interested in investigating the possibility of developing an MAR environment for on-patient medical data visualization which supports high-quality on-patient visualization and depth-based tracking (invariant to illumination conditions). Taking advantage of our main motivation to improve the physician's knowledge of the patient with craniofacial trauma, in this work we focused our tests on the patient's head as the region of interest (ROI). Although we have developed a solution for the scenario of craniofacial data visualization, in this paper we show how the MAR environment can be adapted for other patient ROI (i.e., torso and pelvis; Section 6). The generality of the proposed work is discussed in this paper.

* Corresponding author.
E-mail addresses: marciocfmacedo@gmail.com (M.C.F. Macedo), apolinario@dcc.ufba.br (A.L. Apolinário Jr.).

http://dx.doi.org/10.1016/j.cag.2015.09.007
0097-8493/

Traditionally, on-patient medical data visualization applications superimpose virtual medical data on the patient. However, in such applications, the virtual content seems to be floating in front of the patient. As stated in previous work [2–5], a better solution is to show the patient's anatomy as a focus region in the context of the patient's own body. This process is known as focus plus context (F+C) visualization paradigm [6], and it is known to improve the visual perception of the content being visualized. In the field of volume rendering, one way to improve the understanding and extend the exploration of the medical volume is by use of volume clipping. Therefore, the effect of volume clipping added in an F+C visualization technique is a new tool for the user to explore and understand the augmented scene.

The existing techniques for F+C visualization based on volume clipping are prone to artifacts or do not run in full real time [5]. Such issues decrease the application's visual quality and performance. One way to solve both problems is by use of an adaptive strategy to mitigate artifacts and shaders to execute the technique in parallel.

In this paper, we present improvements in terms of performance and visual quality over the F+C visualization techniques based on volume clipping proposed in [5]. We expand the evaluation of the MAR environment for different ROI in the patient and improve tracking accuracy through the use of a hierarchical algorithm. A more detailed description of the algorithms used in the entire solution (i.e., MAR environment and F+C visualization) and an in-depth analysis of the results obtained and the limitations of the proposed approach are presented as well.

The remainder of this paper is organized as follows. Section 2 reviews recent related work on medical AR and F+C visualization applied in AR. Section 3 introduces the MAR environment used in this paper for validation of the F+C techniques. Section 4 presents the F+C techniques based on volume clipping for on-patient medical data visualization. Section 5 presents the tests conducted and the experimental results obtained. Section 6 discusses the results obtained and the limitations of this work. In Section 7, a summary of the article and recommendations for future work are presented.

2. Related work

Medical AR systems for on-patient medical data visualization have been driven by different approaches in recent years. In this section, we classify the approaches on the basis of their tracking technology: marker based or markerless.

Over the past decades, many relevant approaches have been proposed for marker-based medical AR, such as those in [3,4,7]. Artificial fiducial markers provide fast and accurate tracking because of their shape; however, they are commonly associated with some issues which make this technology unsuitable for on-patient medical data visualization applications:

- They are intrusive, because they are not part of the original scene.
- When the traditional fiducial marker, such as the one used in popular applications such as ARToolKit [8], is not used, the optical tracking system hardware may be too expensive.
- In general, this kind of tracking must operate only on the image space, according to features computed from the pixels. The main drawbacks for this color- or texture-based tracking are the susceptibility to illumination conditions and marker occlusion, which may affect the accuracy of the tracking algorithm.

Recently, systems have been proposed in the field of markerless medical AR. Some of them do not run in real-time (more than 15 frames per second, FPS) [9,10] and others rely on specific prior knowledge about the ROI to be tracked (see [11–13] for the body and [14] for the face). To the best of our knowledge, there is only one exception which can be used for general-purpose markerless on-patient medical data visualization: the semiautomatic approach proposed in [15–17].

The semiautomatic MAR environment uses an RGB-D sensor to reconstruct and track a 3D reference model of the patient's ROI through the AR live stream. Then, after the virtual medical data positioning, it can be displayed for a physician at the location of the patient's real anatomy. Real-time performance is achieved by exploitation of the parallelism provided by the graphics processing unit (GPU).

To validate the F+C visualization techniques, we use a marker-free tracking algorithm because it requires a low processing time and can operate on customer hardware with good accuracy. A first necessary step is to evaluate the performance and visual quality of the proposed approach. In this sense, the semiautomatic MAR environment proposed in [15–17] is used because it runs in real time and, with some adaptations, its tracking solution can be applied for several ROI in the patient, in contrast to other state-of-the-art solutions. Such adaptations are discussed in Section 6.

An application for on-patient medical data visualization requires special attention to be paid to the composition of the virtual and real entities of the AR environment. Recently, many approaches have been proposed in the field of F+C visualization to dynamically define how this composition should be done. These, also known as ghosting or X-ray vision techniques, share the concept of an importance map, a mask (similar to an alpha mask) which controls how real and virtual entities should be blended.

Sandor et al. [18] designed a method for importance map computation based on the feature regions of both real and virtual objects inspired by three features: luminosity—to preserve regions with high illumination; hue—to preserve strong colors; motion—to preserve moving structures in the final rendering. As stated by Sandor et al. [18], this work was an extension of the work of [19], which is based on edge overlay to improve spatial perception.

Mendez et al. [20] proposed an F+C technique in which the lightness and color contrast of a given image are modified according to the importance map computed from a live color video. By adding subtle changes in the image, they guarantee temporal and spatial coherence between frames. The problem with this approach is its performance, which does not achieve the full 30 FPS even when it is implemented on the GPU.

An adaptive F+C visualization technique was recently introduced by Kalkofen et al. [21]. In their approach, an importance map is computed for the occluder [20] and the occludee is inserted into the scene. Then, another importance map is computed from this resulting image and is then compared against the first map computed. Regions on the first importance map that are not present in the final rendering are then emphasized to be visible. This approach improves the visual quality of the augmented scene and it runs in real time. However, it is not suitable for MAR environments, as it alone requires a processing time of 33 ms. Therefore, this additional time would decrease severely the performance of an MAR application.

F+C rendering was also proposed for visualization of underground structures in street scenes [22–24]. In these approaches, a method is used to dynamically compute when the underground structures must be rendered in relation to moving objects present in the scene. Although the final visual quality is good, the performance of the existing techniques is not full real time.

Traditional methods which compute the importance maps from live color video of the real scene are prone to errors because they are dependent on the illumination and material properties of the real environment. To overcome these problems, Mendez and Schmalstieg [25] proposed a method to compute an importance