A Markerless Augmented Reality Environment for On-Patient Medical Data Visualization

Márcio Macedo, Antonio Apolinário
Department of Computer Science
Federal University of Bahia
Salvador, BA, Brazil
Email: marciocfmacedo@gmail.com, apolinario@dcc.ufba.br

Abstract—On-patient medical data visualization is desirable in several situations such as surgical planning and training. Currently, it becomes possible with the augmented reality, a technology which enables the visualization of the patient's virtual anatomy at the location of the real one in a conventional display. In this work[1], we present a markerless augmented reality environment for on-patient medical data visualization, which supports: photorealism, advanced methods for real and virtual data composition and additional features such as occlusion. From an evaluation of the proposed environment, the results obtained highlight that it runs in real-time and provides good visual quality for the augmented scene.

Keywords-Augmented Reality; Markerless Registration; Volume Rendering; Image-Based Lighting; Focus + Context Visualization;

I. INTRODUCTION

Physicians see medical data, typically patient's anatomical structures, on a monitor and they must analyze and mentally compose what is shown on the screen to the patient. This mental model of the patient's anatomy will serve as a basis for health care in routine examinations or time-critical situations [2]. With the availability of the 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 his anatomy. On-patient or *in-situ* medical data visualization can be used for surgical planning, training, medical diagnosis and post-operative examination.

Augmented reality is a technology which augments the view of a real scene with additional virtual information. Medical AR is a sub-field of AR in which the virtual entity is a medical data. In an ideal scenario, a successful medical AR environment must support the following requirements:

- Real-Time Performance - for user interactivity;
- Accurate Tracking - for the proper alignment of the virtual object into the augmented scene;
- Volume Rendering - when the virtual entity consists of a 3D medical volume;
- High Visual Quality - to improve user's perception of the augmented scene;
- Photorealistic Rendering - to allow seamlessly integration of the virtual medical data into the augmented scene;

---

[1]M.Sc. Dissertation [1]

Due to the popularity of freely available AR libraries, the creation of AR applications which use fiducial markers for tracking of the virtual object became easier. However, in medical AR, fiducial markers can be intrusive in the scene and uncomfortable, mainly when the target model is the patient's body. Hence, markerless tracking is desirable in this scenario.

Another distinction between general AR and medical AR is related to the virtual data. Most of the AR applications use polygonal models for rendering. However, volumetric models are commonly used for rendering in medical AR due to the popularity of 3D scanners based on computed tomography (CT) or magnetic resonance imaging (MRI) data for medical procedures.

To improve the visual understanding of the scene by the physician, it is fundamental for a medical AR application to provide a high quality rendering of the virtual objects to be combined with the real scene. Instead of superimposing all the patient's virtual data onto the patient's color image, one solution is to show the patient's anatomic data in a focus region in the context of the patient's body. This process is formally known as focus + context (F+C) visualization and it is already known to improve the user's visual perception of the content being visualized [3]. To allow realistic rendering for the virtual data, local illumination components (i.e. diffuse and specular terms) must be computed from the real scene in real-time to lit the medical volume according to real-world lighting.

In this work, we show that from the recent advances in hardware, as Graphics Processing Units (GPUs) and 3D sensors, and the recent techniques proposed in the academy, it is possible to develop an integrated solution for on-patient medical data visualization which supports all those requirements listed before, still providing real-time performance and good visual quality of the AR scene [1]. The main motivation for our work relies on solving the problem of on-patient medical data visualization for patients with craniofacial traumas. Hence, we have assumed through all the work that the patient's region of interest (ROI) is the patient's head.

Contributions: The main contributions of our work are the proposition of:

- A medical markerless AR (MAR) environment for on-patient medical data visualization based on a new markerless tracking solution which is robust to rigid tracking failures;
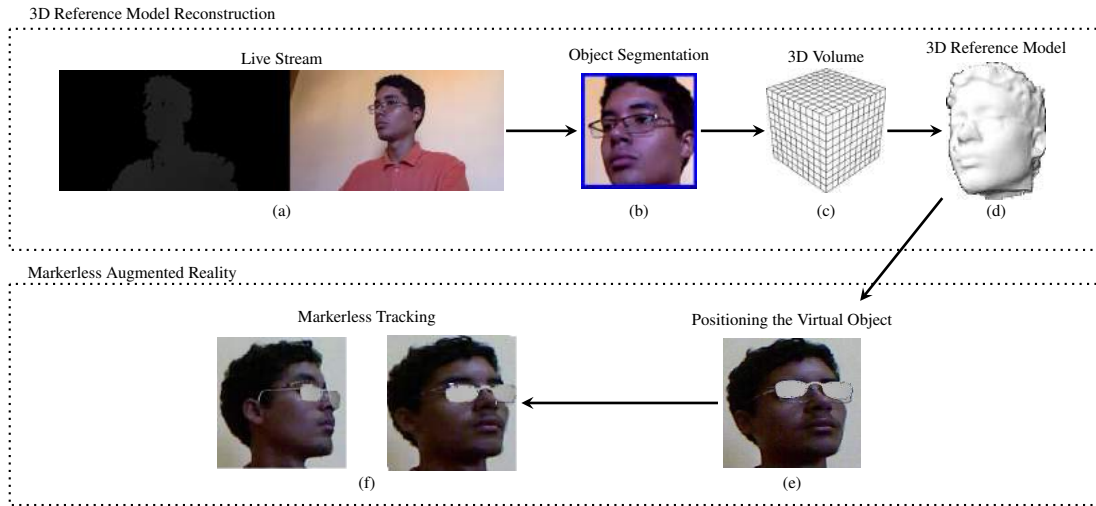
Fig. 1. Overview of the proposed markerless solution: From an RGB-D sensor live stream, the region of interest is detected and segmented in the scene. Different viewpoints of the same region of interest are captured, aligned and integrated in real-time into a single 3D reference model. Based on this model, the virtual data can be positioned into the augmented scene and markerless tracking can be performed in real-time.

- Three new F+C visualization techniques based on volume clipping for on-patient medical data visualization;
- An application for on-patient medical data visualization which supports photorealistic volume rendering based on local illumination components;

The rest of this paper is arranged as follows: Section II presents the markerless tracking solution developed. Section III introduces the MAR environment for on-patient medical data visualization. Section IV shows how photorealistic volume rendering can be done based on real local illumination components. Section V discusses the experimental results. In Section VI, a summary of the developed work is presented. Finally, in Section VII, we list the publications which resulted from the M.Sc. dissertation.

II. MARKERLESS TRACKING

The proposed approach for markerless tracking is based on an RGB-D sensor and a computer with a GPU. We use the Kinect as RGB-D sensor to capture patient's color and depth information [4] (Figure 1-(a)). To achieve real-time performance, all of the steps described in this Section (with exception of the head pose estimation algorithm) run on the GPU.

As we have assumed the patient's ROI to be the patient's own head, to locate and segment it from the scene, we use a face detector in the color image and transfer the segmented region to the depth map (Figure 1-(b)). The Kinect sensor is known to provide noisy depth data. We handle this issue by applying a bilateral filter over the depth map [5]. The depth of the background scene is segmented by applying a Z-axis threshold over the filtered depth map. Then, the filtered and segmented depth map is converted into a vertex map by using the process of intrinsic calibration. The normal map is computed based on a local covariance matrix computed for every vertex [6].

With the patient's ROI properly segmented, we reconstruct its 3D reference model in real-time by using the KinectFusion algorithm [7] and use this model to track patient's movements with a real-time variant of the Iterative Closest Point (ICP) algorithm [8] (Figure 1-(c)(d)).

3D reference model reconstruction is stopped by the user and the virtual data can be positioned into the scene (Figure 1-(e)). MAR live tracking of the real scene is done by aligning the current point cloud captured by the RGB-D sensor with the previous one represented by the 3D reference model in the previous camera pose (Figure 1-(f)). We also improve the robustness of the tracking algorithm by using a real-time head pose estimation algorithm to give an initial guess to the method when it fails [9], [10], [11].

III. ON-PATIENT VOLUMETRIC MEDICAL DATA VISUALIZATION

To enable on-patient volumetric medical data visualization, we must first render the medical volume (Figure 2-(a)(b)). To do so, we use several techniques from the field of volume rendering, such as: direct and context-preserving volume rendering, single-pass GPU ray casting, pre-integrated transfer functions, volume clipping; stochastic jittering, fast GPU-based tri-cubic filtering and Blinn-Phong shading for image quality improvement; empty-space leaping and early ray termination for performance optimization [13].

Afterwards, the medical volume is coarsely registered into the augmented scene by using information about the size and position of the 3D reference model together with the pose estimated from a head pose estimation algorithm [14] (Figure 2-(c)). Then, medical volume position and orientation can be finely adjusted by the user and finally rendered into the augmented scene [15], [16].

To enable high quality on-patient medical data visualization, we use the F+C paradigm to visualize the virtual anatomy as a
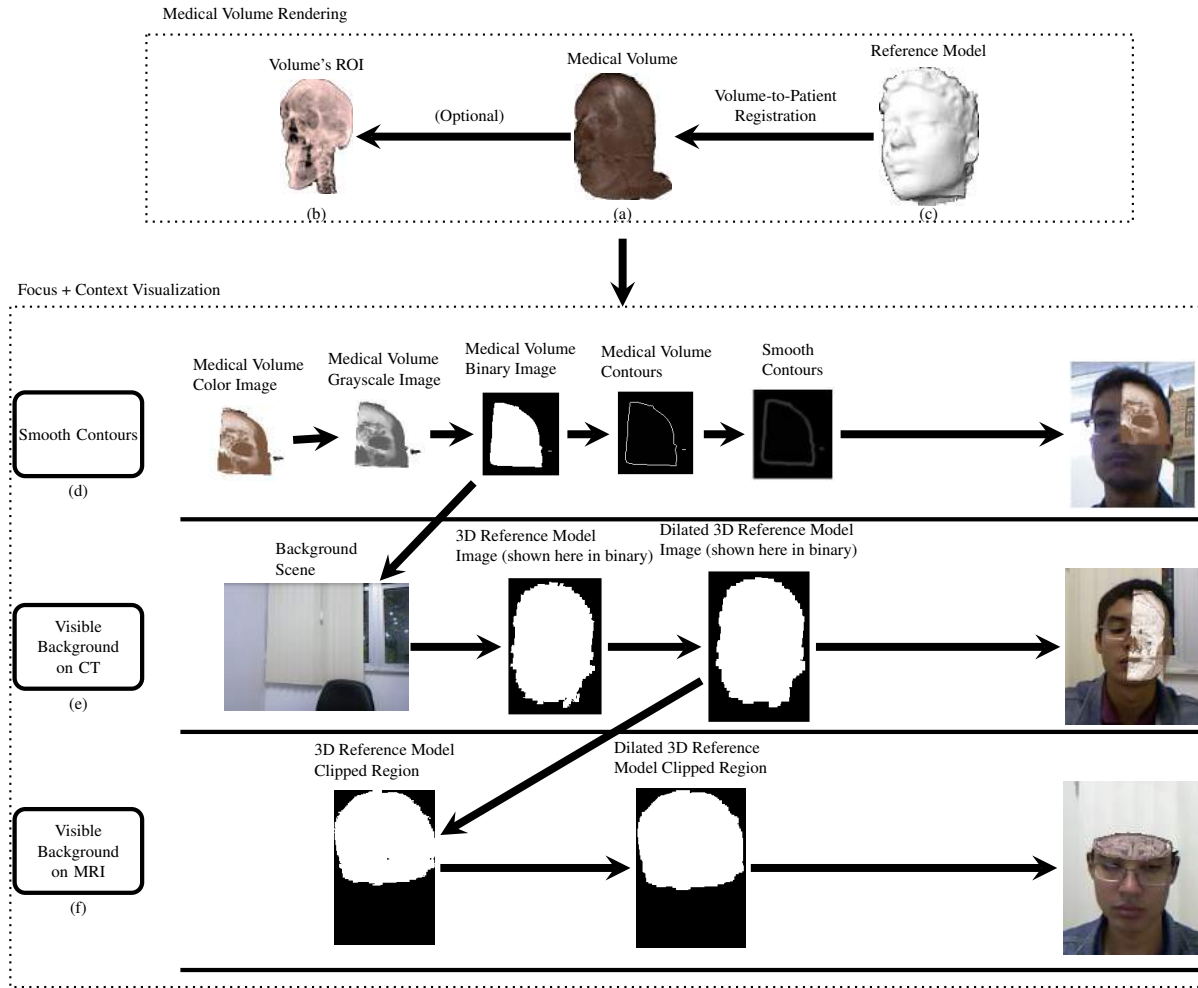
Fig. 2. Overview of the proposed on-patient visualization solution: Medical volume is rendered and, based on the 3D reference model reconstructed previously, the volume data can be positioned into the augmented scene. To improve the visual perception of the scene, we have proposed three new focus + context visualization techniques [12].

focus region in the context of the patient's body. In our work, we have designed three new F+C techniques to improve the visual perception of the augmented scene: smooth contours, visible background on CT data and visible background on MRI data (Figure 2-(d)(e)(f)) [12].

Based on the statement that depth perception is improved by smoothing the transition between the volume in the focus region and the rest of AR scene [17], the smooth contours technique adds a smooth transition between the volume rendered and the patient's image by binarizing the medical volume image with a pre-defined threshold, extracting its contours with the Sobel filter and blurring them with one iteration of a two-pass Gaussian blur (kernel size: 3×3 pixels) (Figure 2-(d)).

In CT data, we can visualize the bones apart from the soft tissue of the volume. Visible background on CT data enables focus for the visualization of bone structures while deemphasizing the visualization of soft tissues. To deemphasize the soft tissue visualization, it is blended with the real background scene captured previously. Medical volume grayscale image is used to separate bone and soft tissue. Medical volume binary

image and patient's 3D reference model are used to delimit the region on the scene where this technique may operate (Figure 2-(e)).

In an AR environment, the best way to visualize organs in MRI data is by clipping not only the medical volume, but the patient's ROI. Visible background on MRI data enables this kind of visualization by allowing the user to cut away volume and patient's ROI in real-time (Figure 2-(f)).

In the visible background-based techniques, dilation is performed over the 3D reference model images if they do not overlap perfectly the patient's ROI captured by the color image of the RGB-D sensor (Figure 2-(e)(f)).

Finally, our MAR environment supports the standard F+C technique proposed in the field of on-patient medical data visualization: the Contextual Anatomic Mimesis (CAM). This technique only uses a focus point and radius to define focus and context regions. In Section V, we discuss the performance and visual quality of all these methods in our MAR environment.

IV. Photorealistic Local Illumination

Techniques for solving the problem of real world lighting estimation have been proposed in the field of image-based lighting. The most common approach relies on the computation of a high dynamic range (HDR) image captured previously to lit the virtual objects [18]. HDR images cover true radiance values for a given scene, however, as they were previously computed based on a static scene, they do not give support for real-time lighting estimation in scenes where lighting condition changes in the environment. We solve this issue by capturing a low dynamic range (LDR) image (i.e. the kind of data captured by consumer cameras) from the light probe (i.e. mirrored sphere which reflects light from the environment) by using an additional webcam directed to it.

After light probe capturing, spherical harmonics basis is used to compactly represent the lighting and its local illumination components [19]. By using nine spherical harmonics coefficients, we can recover diffuse and specular lighting terms to lit the medical volume according to the real-world environmental lighting [20], [21].

V. Results and Discussion

In this section we describe the experimental setup and analyze the experimental results achieved with our approach with respect to performance, accuracy and visual quality.

A. Experimental Setup

For all tests we used an Intel Core i7, 8GB RAM, NVIDIA GeForce GTX 660 as a computer equipped with a GPU. Kinect was used as RGB-D sensor [4]. Light probe is a common mirrored sphere and Philips SPC530NC was used as webcam to capture the light probe image.

Medical datasets are an MRI volume of a head of resolution 256^3 available in Volume Library [22] and a CT volume of a head of the Visible Human Project of resolution $128 \times 256 \times 256$ [23]. The user's head was reconstructed with the KinectFusion using a grid with the volume size of $70cm \times 70cm \times 140cm$ and resolution of 512^3. Resolution of 256^3 was only used to evaluate the performance of the MAR environment. The light probe image was captured and cropped to a fixed resolution of 128^2.

We evaluate performance and visual quality of the MAR environment in a scenario where the patient's head is augmented with a generic medical volume of a head. The use of a generic medical volume does not affect our performance evaluation since the dataset is of typical resolution. Moreover, it does not affect our visual quality evaluation since the volume is scaled and positioned semi-automatically into the MAR scene [14].

B. Performance

The performance of the on-patient medical data visualization application is shown in Figure 3 for two different sizes of 3D grid used to reconstruct the 3D reference model. By changing the resolution of the 3D grid from 512^3 to 256^3, one will decrease tracking and 3D reference model reconstruction

accuracy, as well as the quality of the occlusion and the final composition for several F+C techniques. Moreover, in Figure 3, we evaluate the impact on the use of photorealism in the application's processing time.

3D reference model reconstruction runs between 35 and 40 frames per second (FPS). Photorealism does not influence its processing time because 3D reconstruction is not an AR step of the application (i.e. there is no virtual data to be rendered photorealistically). Three of the F+C rendering techniques for on-patient medical data visualization run above 30 FPS even by using the resolution of 512^3 and enabling photorealism. The exception for this situation occurs in the method of visible background on MRI data, which is slightly slower than the others, but it still enables user interactivity with the application, achieving performance above 25 FPS in all the tests conducted.

In terms of medical volume-to-patient registration, we have observed that the user takes less than 10 seconds to position and adjust the volume in the scene. The algorithm for coarse medical volume-to-patient registration is used only once (i.e. at the transition between 3D reference model reconstruction and on-patient medical data visualization, Figure 2) and it takes $\approx 60ms$ using the hardware described in the previous subsection.
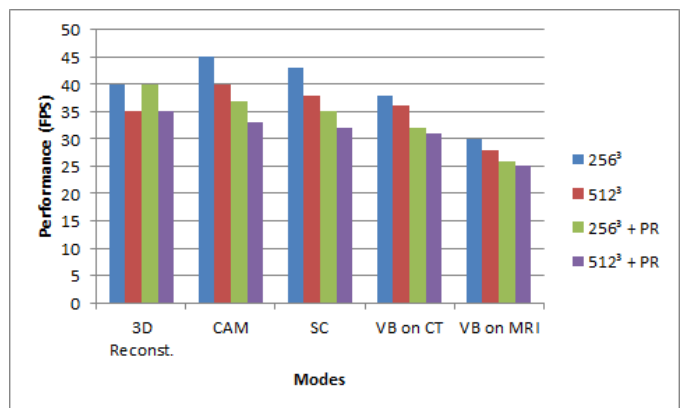

Fig. 3. Performance results measured in frames per second (FPS) for each one of main steps/modes of visualization discussed in this paper. PR - Photorealism, 3D Reconst. - 3D reference model reconstruction, CAM - Contextual Anatomic Mimesis, SC - Smooth Contours and VB - Visible Background.

C. Accuracy

3D reference model reconstruction has accuracy of $\approx 10mm$, as evaluated in related work [24]. Live tracking has an accuracy of $\approx 2mm$, which is not incremental (i.e. the error does not accumulate between frames). Accuracy of the registration between the virtual data and the reference model depends on the user's adjustments due to the use of a generic volume instead of the patient's own to validate the proposed environment.

Given these accuracy measurements, we observe that, visually, the registration is accurate. In this sense, the full MAR solution may be suitable for training purposes. However, for

scenarios in which high accuracy is required, such as the ones for surgery procedures, our solution is not recommended.

To improve tracking accuracy, one can increase the number of ICP iterations, decreasing application's performance. An alternative method for real-time 3D reconstruction may be used to improve the accuracy of the 3D reference model.

D. Visual Quality

Our entire solution handles occlusion by comparing depth data from current and previous frames (Figure 4).

Fig. 4.   Occlusion support.

We can see an integration between two F+C techniques, smooth contours and CAM, in Figure 5. In Figure 5-left image, a circular mask is defined over the window to select which parts of the medical volume must be rendered into the augmented scene. With the CAM method, there is no clear handling of the contours which result from the volume clipping (zoomed in Figure 5-left image). By using the smooth contours technique (Figure 5-right image), we can solve this problem by smoothing the contours inside the focus window.

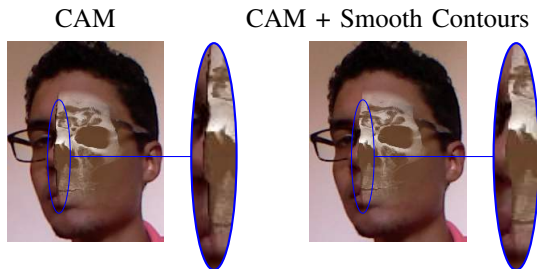CAM   CAM + Smooth Contours

Fig. 5.   F+C visualization based on the CAM algorithm (left image) and its extension with smooth contours (right image).

For the F+C visualization based on visible background on CT data, bone and soft tissue structures can be separated according to a user-defined parameter. By using this technique, the volume can be rendered almost completely invisible, with the soft tissue linearly interpolated with the background scene or with the volume rendered without the background scene (Figure 6).

In Figure 7, we can see more examples of interactions with the F+C visualization based on visible background on MRI data. The user controls both clip planes used to cut away patient's ROI and medical volume. By moving his ROI in

Fig. 6.   F+C visualization based on visible background on CT data. By using an user-defined parameter, the medical volume can be rendered in different ways [12].

front of the sensor, the user can see different parts of the MRI data.

Fig. 7.   Examples of interactions with the F+C visualization based on visible background on MRI data.

In terms of photorealistic on-patient medical data visualization, our approach can compute diffuse and dominant specular terms for several challenging illumination conditions, as depicted in Figure 8. Thereby, allowing a consistent, seamlessly integration of the virtual medical volume into the real scene.

Light
Probe

Real
Scene

Fig. 8.   Photorealistic on-patient medical data visualization based on local illumination components.

VI. Conclusion

In this work we have proposed the first MAR environment for on-patient medical data visualization based on photo-realistic local illumination and F+C visualization. From an evaluation of the proposed solution, we conclude that it runs in full real-time on typical medical datasets, provides high visual

quality of the final augmented scene through the use of F+C visualization paradigm and photorealistic rendering. Moreover, it provides accuracy enough for applications that need good "visual" accuracy for the registration (i.e. good composition and tracking of the virtual object into the augmented scene). In terms of generality, we believe that with some additional changes mainly for the segmentation and tracking algorithms, one can use this MAR environment for other parts of the patient's body. A more in-depth discussion about all the steps composing the final solution as well as the results achieved can be found in [1].

VII. PUBLICATIONS

The development of this work (M.Sc. dissertation) [1] resulted in the following scientific publications:

- [9], work presented at the XV Symposium on Virtual and Augmented Reality (oral presentation), selected as one of the best works of the conference, and invited for publication - extended version - in the Journal on Interactive Systems (conference with *qualis B4*);
- [10], work presented at the XVI Conference on Graphics, Patterns and Images (poster presentation) (conference with *qualis B1*);
- [11] and [14], works published in the Journal on Interactive Systems (journal with *qualis B4*);
- [15], work presented at the XVI Symposium on Virtual and Augmented Reality (oral presentation), selected as one of the best works of the conference, and invited for publication - extended version - in the Journal on Interactive Systems (conference with *qualis B4*);
- [16], work presented at the XIV *Workshop de Informática Médica* (oral presentation), winner of the Best Full Paper Award (conference with *qualis B4*).
- [12], work presented at the XVII Conference on Graphics, Patterns and Images (oral presentation), selected as one of the best works of the conference, and invited for publication - extended version - in the journal Computers and Graphics (conference with *qualis B1*);
- [25], work presented at the XIII ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry (oral presentation) (conference with *qualis B2*);
- [26], work presented at the IEEE Virtual Reality (poster presentation) (conference with *qualis A2*);

ACKNOWLEDGMENT

REFERENCES

[1] M. Macedo, "A markerless augmented reality environment for on-patient medical data visualization," Master's thesis, Federal University of Bahia, Nov 2014.
[2] C. Bichlmeier, "Immersive, interactive and contextual in-situ visualization for medical applications," Dissertation, Technische Universität München, München, 2010.
[3] S. K. Card, J. D. Mackinlay, and B. Shneiderman, Eds., *Readings in Information Visualization: Using Vision to Think*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1999.
[4] L. Cruz, D. Lucio, and L. Velho, "Kinect and rgbd images: Challenges and applications," in *Graphics, Patterns and Images Tutorials (SIBGRAPI-T), 2012 25th SIBGRAPI Conference on*, 2012, pp. 36–49.
[5] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *ICCV*, jan 1998, pp. 839 –846.
[6] S. Holzer, R. Rusu, M. Dixon, S. Gedikli, and N. Navab, "Adaptive neighborhood selection for real-time surface normal estimation from organized point cloud data using integral images," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, Oct 2012, pp. 2684–2689.
[7] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon, "Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera," ser. UIST '11. USA: ACM, 2011, pp. 559–568.
[8] S. Rusinkiewicz and M. Levoy, "Efficient variants of the icp algorithm," in *3DIM*, 2001, pp. 145–152.
[9] M. Macedo, A. Apolinario, and A. Souza, "A robust real-time face tracking using head pose estimation for a markerless ar system," in *Virtual and Augmented Reality (SVR), 2013 XV Symposium on*, May 2013, pp. 224–227.
[10] M. Macedo, A. L. Apolinario, and A. C. Souza, "A Markerless Augmented Reality Approach Based on Real-Time 3D Reconstruction using Kinect," in *Workshop of Works in Progress (WIP) in SIBGRAPI*, Arequipa, Peru, august 2013.
[11] ——, "KinectFusion for Faces: Real-Time 3D Face Tracking and Modeling Using a Kinect Camera for a Markerless AR System," *Journal on 3D Interactive Systems*, vol. 4, pp. 2–7, 2013.
[12] M. C. d. F. Macedo and A. L. Apolinario, "Improving on-patient medical data visualization in a markerless augmented reality environment by volume clipping," in *Graphics, Patterns and Images (SIBGRAPI), 2014 27th SIBGRAPI Conference on*, Aug 2014, pp. 149–156.
[13] M. Hadwiger, J. M. Kniss, C. Rezk-salama, D. Weiskopf, and K. Engel, *Real-time Volume Graphics*. USA: A. K. Peters, Ltd., 2006.
[14] M. Macedo, A. Apolinario, A. C. Souza, and G. A. Giraldi, "High-Quality On-Patient Medical Data Visualization in a Markerless Augmented Reality Environment," *Journal on 3D Interactive Systems*, vol. 5, pp. 41–52, 2014.
[15] M. C. Macedo, A. L. Apolinario, A. C. Souza, and G. A. Giraldi, "A semi-automatic markerless augmented reality approach for on-patient volumetric medical data visualization," in *SVR*, May 2014, pp. 63–70.
[16] M. Macedo, C. Almeida, A. Souza, J. Silva, A. Apolinario, and G. Giraldi, "A Markerless Augmented Reality Environment for Medical Data Visualization," in *WIM*, Brazil, 2014, pp. 1682–1691.
[17] C. Bichlmeier, F. Wimmer, S. M. Heining, and N. Navab, "Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality," ser. ISMAR '07. Washington, DC, USA: IEEE, 2007, pp. 1–10.
[18] P. Debevec, "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography," ser. SIGGRAPH '98. New York, NY, USA: ACM, 1998, pp. 189–198.
[19] P.-P. Sloan, "Stupid spherical harmonics (sh) tricks," in *GDC*, 2009.
[20] R. Ramamoorthi and P. Hanrahan, "An efficient representation for irradiance environment maps," ser. SIGGRAPH '01. New York, NY, USA: ACM, 2001, pp. 497–500.
[21] D. Nowrouzezahrai, S. Geiger, K. Mitchell, R. Sumner, W. Jarosz, and M. Gross, "Light factorization for mixed-frequency shadows in augmented reality," ser. ISMAR '11. Washington, DC, USA: IEEE Computer Society, 2011, pp. 173–179.
[22] Volume Library, "Volume library," http://www9.informatik.uni-erlangen. de/External/vollib/, 2014, accessed 22 September 2014.
[23] National Library of Medicine, "Visible human data," http://www.nlm. nih.gov/research/visible/, 2014, accessed 22 September 2014.
[24] S. Meister, S. Izadi, P. Kohli, M. Hämmerle, C. Rother, and D. Kondermann, "When can we use kinectfusion for ground truth acquisition?" in *IROS*. IEEE Computer Society, 2012.
[25] A. C. S. Souza, M. C. F. Macedo, and A. L. Apolinario, Jr., "Multi-frame adaptive non-rigid registration for markerless augmented reality," ser. VRCAI '14. New York, NY, USA: ACM, 2014, pp. 7–16.
[26] ——, "A GPU-Based Adaptive Algorithm for Non-Rigid Surface Registration," in *Virtual Reality (VR), 2015 iEEE*, March 2015.