Marker Based vs. Natural Feature Tracking Augmented Reality Visualization of the 3D Foot Phantom

Saša Ćuković¹, Michele Gattullo², Frieder Pankratz³, Goran Devedžić¹, Ernesto Carrabba⁴, Khelifa Baizid⁵

¹ University of Kragujevac, Faculty of Engineering, Serbia, ² Politecnico di Bari, DMMM, Italy, ³ Technical University of Munich, CAMPAR Lab, Germany, ⁴ Politecnico di Milano, Dipartimento di Meccanica, KAEMaRT, Italy, ⁵ University of Cassino and Southern Lazio, Italy.

cukovic@kg.ac.rs, michele.gattullo@poliba.it, pankratz@in.tum.de, devedzic@kg.ac.rs, ernesto.carrabba@polimi.it, baizid@unicas.it

ABSTRACT

In this paper we present novel methods and algorithms for augmenting a foot 3D model in real-time. By overlaying reconstructed anatomical contents onto user’s field of view via an augmented reality visualization platform, the virtual Biomedical Computer Aid Design (BioCAD) objects, appear geo-referenced to the real environment, help physicians and/or students to improve the overall image of the bone condition and anatomical structures of patients. Moreover, we explain the principle of marker and markerless Augmented Reality (AR) and interactive integration of virtual objects in the actual scene by development of a PC and mobile AR applications. In particular, this should be a powerful tool to understand the complex 3D nature and structure of 3D foot bones, joints and assemblies.

KEYWORDS

Augmented Reality, Foot 3D Model, BioCAD, Android, CT Reconstruction, DICOM.

1 INTRODUCTION

Augmented Reality (AR) systems are used to enhance the perception of the real 3D world and real objects. A real scene that a person sees is augmented with computer-generated 3D objects. These virtual objects are registered in the scene in such a way that the computer-generated 3D information appears in the correct location with respect to the real objects in the scene [1]. AR can be classified along a virtuality continuum as illustrates Figure 1.

This paper presents novel algorithms and platforms of AR technology in domain of medical visualization. In the case of complex foot anatomy, we present two algorithms: the first is based on marker and the second is based on markerless principles (Natural Feature Tracking). The core concept of these interactive systems is video image processing techniques and interactive 3D model visualization.

1.1 Augmented Reality Concept

One of the most commonly used definitions of AR was given by Ron Azuma [1]. Independent
of specific technologies, an AR system has to meet the following requirements: combine real and virtual worlds, augmentations are interactive in real time and they are registered in 3D virtual to the real world.

In the last few years, medical AR applications experienced a rapid expansion, driven by advances in hardware (tracking, haptics, and displays), new concepts in user interface design, such as tangible user interface (TUI) and a set of new interface metaphors and display techniques, such as magic lens and virtual magic mirror [2] have been appeared.

In recent decade smart phones and tablets became an increasingly popular device for AR in medicine, industry and education [3], [4]. Those devices combine all needed components (camera, display and processing power) for video based AR in a compact form [4].

2 3D RECONSTRUCTION OF HUMAN STRUCTURES

Standardized procedure of anatomical 3D reconstruction is presented in [5]. To obtain 3D models of the bones we have used modern 3D imaging modalities, such as CT-DICOM, which enables extracting and generating 3D anatomical model of the human structures in detail. Those models are adequate for preoperative preparation and planning, as well as for anatomy learning. To this end, we made a master model – virtual foot phantom.

The Visible Foot Phantom

Phantom is another word for a life-size virtual replica of a human body or one of its parts. The visible foot phantom was generated using a 3D reconstruction method applied on patient’s DICOM images (SIEMENS/Sensation 64, 512x512pxl, Pixel size: 0.402mm, Slice increment: 0.5mm, FOV: 20.6cm). As illustrated in Figure 2, direct 3D reconstruction of each slice and rendering volume allows to achieve correct depth perception of inner organs and to perform segmentation of each bone. We made a full 3D model of the foot by combining two reconstruction packages [6], [7].

Figure 2. Creating 3D models of the foot - point cloud and contour extraction from DICOM slices

First phases of the reconstruction process are realized in Materialize Mimics (a biomedical software specially developed for medical image
processing), segmentation and highly accurate 3D models of patient’s anatomy calculation (from Computer Tomography - CT, Magnetic Resonance - MRI, Micro CT - μCT, Ultrasound, Confocal Microscopy) [5]. In this case we used a trial version of Mimics for:
- Importing DICOM data of the patient obtained from CT device,
- Segmenting 29 foot bones, slice by slice according to the level of gray color and Hounsfield units,
- Generating polylines and point clouds of the bone models,
- Exporting 3D models in STL and IGS formats for further processing.

With reverse engineering (RE) software Geomagic Studio [7], we have made further point clouds processing through few phases (noise reduction, filtering outliers and wrapping). In the phase of polygonal meshes we have done decimating, spikes removing, mesh relaxation, holes filling and defeaturing (Figure 3). Final stage of mesh processing is contour detection and patches generation. This is a preliminary step for the surface phase. Once we have exported models with NURBS patches, we processed it in PLM system CATIA [8], by joining patches into unique surface model. By adding volumetric features we have made a solid 3D model for each bone, and then by defining joints and constraints we have created a final foot assembly. Final assembly contains 28 bones including distal parts of the fibula and tibia. Models of the foot assembly exported in STL format are given in Table 1.

![Figure 3. Creating 3D models of vertebrae - Polygonal and surface model generation](image)

<table>
<thead>
<tr>
<th>Table 1. Foot bones</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Os metatarsale I.stl,</td>
<td>14 Os cuneiforme intermedium.stl,</td>
<td></td>
</tr>
<tr>
<td>2 Os metatarsale II.stl,</td>
<td>15 Os cuneiforme laterale.stl,</td>
<td></td>
</tr>
<tr>
<td>3 Calcaneus.stl,</td>
<td>16 Os cuneiforme mediale.stl,</td>
<td></td>
</tr>
<tr>
<td>4 Os cuboideum.stl,</td>
<td>17 Os metatarsale IV.stl,</td>
<td></td>
</tr>
<tr>
<td>5 Os naviculare.stl,</td>
<td>18 Os metatarsale V.stl,</td>
<td></td>
</tr>
<tr>
<td>6 Phalanx distalis V.stl,</td>
<td>19 Os metatarsale stl,</td>
<td></td>
</tr>
<tr>
<td>7 Phalanx distalis I.stl,</td>
<td>20 Phalanx distalis II.stl,</td>
<td></td>
</tr>
<tr>
<td>8 Phalanx distalis III.stl,</td>
<td>21 Phalanx distalis IV.stl,</td>
<td></td>
</tr>
<tr>
<td>9 Phalanx media II.stl,</td>
<td>22 Phalanx media III.stl,</td>
<td></td>
</tr>
<tr>
<td>10 Phalanx media V.stl,</td>
<td>23 Phalanx media IV.stl,</td>
<td></td>
</tr>
<tr>
<td>11 Phalanx proximalis I.stl,</td>
<td>24 Phalanx proximalis II.stl,</td>
<td></td>
</tr>
<tr>
<td>12 Phalanx proximalis III.stl,</td>
<td>25 Phalanx proximalis IV.stl,</td>
<td></td>
</tr>
<tr>
<td>13 Phalanx proximalis V.stl,</td>
<td>26 Talus.stl.</td>
<td></td>
</tr>
</tbody>
</table>
In the Table 2 are details of reconstruction process of all parts from DICOM to solid 3D models.

### Table 2. 3D reconstruction of anatomical models

<table>
<thead>
<tr>
<th>Phase</th>
<th>Detail</th>
<th>Processing data</th>
<th>File format</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Tresholding, Segmentation, Region growing, Polylines detection, Exporting.</td>
<td>Number of slices: 250</td>
<td>*.dcm, *.mcs</td>
</tr>
<tr>
<td>II</td>
<td>Point cloud filtering, Warping, Decimating, Hole filling, Exporting.</td>
<td>Initial number of points: 156686 Reduced number of points: 144296</td>
<td>*.igs, *.wrp, *.stl</td>
</tr>
<tr>
<td>III</td>
<td>Surface extraction, Solid modeling, Exporting.</td>
<td>Number of triangles: 284019</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>Surface extraction, Solid modeling, Exporting.</td>
<td>Number of features/models: 28</td>
<td>*.CATPart, *.vrml</td>
</tr>
</tbody>
</table>

All vertebral models are exported in an appropriate exchange format such as *.vrml file (Virtual Reality Modeling Language) and then as a *.obj (Object) models that are adopted, later, for AR environments. These patient-specific models or phantoms can be used for a variety of engineering applications in external software like statistical, CAD/CAM, or FEA packages.

### 3 SYSTEM ARCHITECTURE

In related articles, some promising methods have been proposed for improving visualization during interventional therapy via AR using head-mounted displays, external cameras or intra-operative projector systems.

To fulfill all requirements for a given AR system, the foundation is to estimate the position and orientation of the camera in respect to the world frame or vice versa. The combination of a position and an orientation is called a pose. To do this we employ two image-based tracking techniques: marker-based tracking [4], and Natural Feature Tracking [9].

#### 3.1 Marker Based AR System

Marker tracking makes use of the camera image to find optical square markers and estimate their relative pose to the camera. A square marker consists of a black square with a white border with a predefined size. Within the square

![1. Acquirer image](image1)

![2. Calculate exact edges](image2)

![3. Calculate position and orientation of marker](image3)

![4. Registration of the specific model to the specific marker](image4)

**Figure 4. Marker based tracker pipeline**
the ID of the marker is encoded. Different techniques can be used to encode the ID like template matching or encoding as a binary number, as in our case. The marker tracking pipeline [4] is illustrated in Figure 4.

After getting the image from the camera is converted to a gray scale image to speed up the image processing algorithm, edge detection and registration. By using the camera image as the background in our display (real world) and using the pose of the marker we, now, can superimpose the camera image and virtual object (virtual world) as seen in Figure 4. When either the marker or the camera is moving, the augmentation stays on the marker (registered in 3D space). The marker tracking pipeline is computationally inexpensive, so we can keep all interactions with virtual objects in real time. Students and/or doctors observe the scene through a video camera, where the rendered image of the virtual phantom is composed with video stream, directed at the patient’s foot, as it is shown in detail in Figure 6.

Our marker-based system is composed of a tracking framework to provide the necessary tracking data and a game engine for rendering the virtual models and augmentations interactions. As the tracking framework we implemented UbiTrack [10]. UbiTrack is an open source, general purpose tracking framework for AR developed by the “Fachgebiet Augmented Reality” group of the Technische Universität München, published under the LGPL license.

For the game engine we used Unity3D [11] for the ease of use and its independency. This allows us to deploy our application to all desktop systems and mobile devices without any need of changing the source code of the application. By employing a desktop version or a Smartphone equipped with a camera of our developed AR platform, students and physicians are able to use it easily. Focusing the camera on the markers retrieve the virtual 3D object from database. Information and graphics are then overlaid onto a screen in order to improve perception of the nature of foot structures.

By recognizing the ID of the square-sized marker, the application determines which model and information to display. The interaction with the CAD model is handled by single touch rotation gestures on the touch-screen of the device or by mouse inputs in case of using desktop application.

### 3.2 Natural Feature Tracking

Natural Feature Tracking is an image-based tracking technique that detects and tracks the features that are naturally found in the image itself. These could be corners, edges, blobs, etc., without using specifically designed ID markers.

There are several Natural Feature Tracking approaches such SIFT, SURF, Ferns [9], [12]. They differ mostly by image features that are associated between the video image and a model of the object or environment to be tracked [9]. However, the tracking pipeline can be sketched as shown in Figure 5. For instant, “Keypoint detection” is usually accomplished with a FAST corner detector to calculate feature points in the camera image.

![Figure 5. Natural Feature Tracking Pipeline](image-url)
The “Descriptor creation and Matching phase” depends on the choice of a descriptor; for example, with the SIFT pipeline [9], the algorithm estimates the dominant keypoint orientation using gradients, compensates for detected orientation and finally describes the keypoints in terms of the gradients surrounding it. In an offline step, a database file with all descriptors and their position on the original image is built. Then, during real-time tracking the algorithm matches the descriptors from the live video against those in the database. The “Outlier Removal” consists of a cascade of keypoints removal techniques, starting with cheapest (simple geometric tests such as line tests), finishing by most expensive (homography-based tests). Pose from homography is used as starting point for “Pose estimation”. Then, a refinement step is accomplished based on Gauss-Newton iteration: minimization of the re-projection error of the keypoints. Typically 2-4 iterations are enough.

As tracking framework, we used Vuforia SDK image targets [13]. We created our own image target using the page of the Atlas book [14]. Figure 6 shows the image used as target and the features extracted from it. The augmentable rating of such image was equal to 5 in a range of 0 to 5, this means that the image is easily tracked by the proposed AR system. In fact, it contains a lot of zones with good local contrast especially on bones edges.

We, then, developed our AR application using Unity3D software [11], importing the tracking database generated in the Vuforia developer portal [15]. We took also the advantage of the “extended tracking” option, which enables our application to have a continuous experience whether or not the target remains visible in the camera field of view. As the target goes out of view, Vuforia uses other information from the environment to infer the
target position by visually tracking the environment. Vuforia builds a map around the target specifically for this purpose and assumes that both the environment and target are largely static. Results are shown in figure 7.

4 CONCLUSIONS AND FUTURE WORK

AR technology can be used as a new tool to support teaching activities and diagnosis aided tool to enhance the traditional learning experience and medical diagnosis. The greatest advantage of the developed UbiTrack Framework is the use of so called “Spatial Relationship Patterns”. This leads to a component based design that allows an easy replacement of specific hardware drivers and tracking methods, as well as it represents a less error prone way to develop and setup more complex AR systems. UbiTrack has been successfully ported to Microsoft Windows, Linux, Mac OS and Android. However, the placing of a marker on all the book pages is not a viable solutions: it requires some extra space on the page, it is not scalable to existing books, this could annoy some users who do not want to use it in AR mode. As presented earlier, solutions to these problems do not exist using the Natural Feature Tracking, where the page itself is used for tracking, without the need of a customized marker. Furthermore, with the newest Vuforia SDK releases, the quality of the tracking is improved a lot, resulting better in some cases than that with ID markers. Since the foot joints are very complex, our conceptual systems can emphasize the perception of students and medical staff and offers information that is not perceived directly by the use of their own senses or standard x-rays. Besides understanding the structure and joins AR platform aims to comprehend geometric relationships and measurements of the key internal parameters. Such models can be further enhanced by integrating types of information other than just simple 3D representations. These types can include audio, text annotations, 2D images, and diagrams. As the development of the UbiTrack framework continues, we plan to extend the use of future ports of UbiTrack framework to other operating systems and all types of desktop and mobile devices. A video presenting marker based application is available on: https://www.youtube.com/watch?v=B3_FpP768Rg&list=UUd0rnbGWo-NL0yncoQ1ICH-Q.

Acknowledgment

This work is supported by national project “Application of Biomedical Engineering in Preclinical and Clinical Practice”, supported by the Serbian Ministry of Education, Science and Technological Development (III-41007) and TEMPUS Project, “BioEMIS” (530423 - TEMPUS), funded by European Commission.

REFERENCES

[7] Geomagic Studio V12 Trial Version -
Last access: 13/10/2013.

[8] Software Dassault System CATIA - Student version
V5R20: http://www.3ds.com/products-
services/catia/welcome/. Last access: 09/10/2013.

[9] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond,
D. Schmalstieg, “Pose tracking from natural features
on mobile phones”, In Proceedings of the 7th
IEEE/ACM International Symposium on Mixed and

access: 13/10/2014.


[12] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond,
D. Schmalstieg, “Real-time detection and tracking
for augmented reality on mobile phones”, IEEE

[13] https://developer.vuforia.com/resources/dev-

Anatomy, Based on the International Nomenclature”
Stuttgart, Germany, Thieme, ISBN 0-86577-928-7,
2000.

access: 17/11/2014.