Flexible Architecture for Multimodal Augmented Reality Engineering Applications

M. Dellisanti, M. Fiorentino, G. Monno, A.E. Uva
DIMeG
Politecnico di Bari, Bari, Italy
{m.dellisantifabiano, m.fiorentino, gmonno, a.uva}@poliba.it

Abstract

We present a flexible framework for multimodal engineering applications using augmented reality. Our goal is to simplify the configuration procedures and to provide a higher grade of flexibility in multimodal interfaces. The system architecture is based on an extensible plug-in approach. A specific component has been designed to manage and synchronize the different multimodal inputs. A configurable XML based layer manages user preferences in a hierarchical way. We tested two engineering applications: a structural component re-design case and an industrial plant visualization. Industrial experts were positively impressed by the augmented visualization and by the usability of the interface. Most of them agree that the multimodal interface surpasses the desktop based interface.

Keywords: *Augmented Reality, Multimodal Interaction, Computer Aided Engineering.*

1. INTRODUCTION

Global competition led to a continuously increasing complexity in product engineering (geometry, integration, maintenance and support, etc.). The market presents several software packages to aid and to manage the integration of engineering processes, through Computer Aided Engineering (CAE). Unfortunately CAE human computer interfaces did not evolve from the two dimensional approach of the WIMP metaphor (Window, Icon, Menu, Pointing device).

Two emerging technologies can provide a natural and integrated support to CAE tools: Augmented Reality (AR) and Multimodal Interfaces (MI).

Augmented Reality is a promising technology able to integrate real products and scenarios with engineering data. It also allows the use of common hardware (pc/laptop, webcam, projectors, etc.) instead of expensive and complex VR devices. An important survey of present and future AR systems is provided by Azuma [1].

Differently, multimodal Interfaces put on practice the so called pervasive computing, which simulates human-like sensory perception by interpreting continuous simultaneous input from different sources such as human speech, pen input, gesture, gaze, and other natural behaviours. Multimodal interface have been proved to get better input accuracy and reduce task completion time [2].

Unfortunately, both AR and MI technologies require a large number of parameters to be set for optimizing the interface and for adapting the system to subjective preferences in specific engineering tasks.

We addressed system configuration complexity presenting a novel framework designed to integrate the multimodal interface metaphor in an AR visualization system for engineering purposes. The system architecture has been developed using a plug-in approach to decouple each service from the core functionalities.

2. RELATED WORKS

Multimodal Interfaces (MI) and Augmented Reality (AR) technologies are candidate to be the ideal interface for engineering software tools (CAE).

Several AR industrial applications have been described in ARVIKA project [3]. Shin et al. [4] demonstrated that the perception of 3D design in an augmented scene is improved compared with monitor display. Dunston et al. [5] proved that AR improves user geometry comprehension of complex 3D models. One of the first contributions to address multimodal input was proposed by Vo and Weibel [6]. They developed a Java based framework and a grammar structure for a broad class of multimodal applications. Milota in [7] studied the general strategies employed in pen interactions when combining speech and gesture. Gutierrez et al. [8] presented a system for real-time configuration of multimodal interfaces to Virtual Environments (VE). Their framework uses portable XML descriptors to define the I/O channels of a variety of interaction devices.

Although several approaches have been presented, none of the previous, as far as the authors know, fully addresses engineering applications using flexible multimodal interface in augmented reality.

3. THE FRAMEWORK ARCHITECTURE

Our system has been designed mixing different technologies to maximize flexibility and expandability. To reach this goal we implemented some of the functionalities into plug-ins (Figure 1). This solution allows to change dynamically the behaviour of the system.

The main application consists of a base GUI which provides a graphic window controlled by the rendering module. Each plug-in, when loaded and started, receives access to a specific node of the scene-graph managed by the renderer. The scene-graph is the in-memory database where all 3D graphical entities are stored. Therefore each plug-in is able to control its own entities in the virtual scene. Broadly the system is structured as depicted in Figure 1.

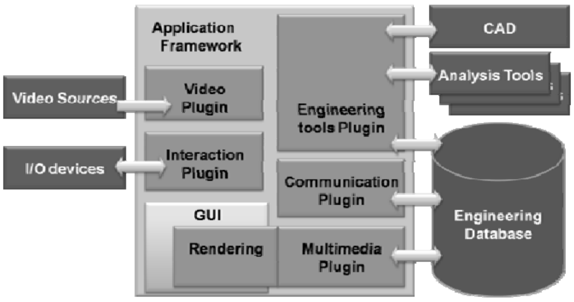

Figure 1. The framework architecture.

The different functional blocks of the application can be configured separately using XML configuration files: (i) Application, (ii) Interaction, (iii) Video, (iv) Engineering and (v) Multimedia. The use of XML files delivers the flexibility to test different setups. For example, it is possible to switch from a desktop environment to a mobile (laptop, tablet, PDA, etc) one at runtime leaving the application unchanged. This approach is very useful for the final user and for the developer which can test different configurations without re-compiling the application.

3.1. Rendering Module

The rendering module (Figure 2) manages the scene-graph where all 3D elements, which describe the visual environment, are stored. It includes two rendering submodules: one targeted to 3D visualization and another targeting 2D elements.

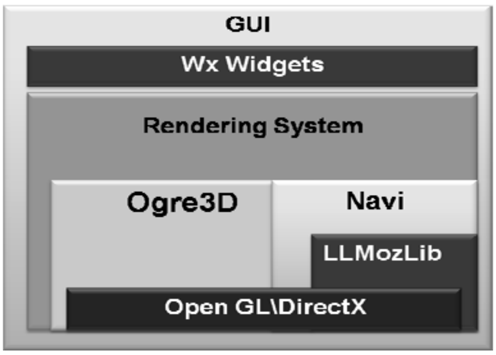

Figure 2. Rendering architecture.

The 3D sub-module has been developed using the open source Ogre3D [9] library. The 2D rendering submodule has been developed using an existing experimental Ogre3D-based library called "Navi" [10] which uses the LLMozLib library to embed Gecko (Mozilla's rendering engine) into traditional 3D applications.

3.2. Video Plug-in

The video plug-in manages the video acquisition and display. The stream is directly delivered to a node of the scene graph for AR overlay and for other services (i.e. optical tracking). Different plug-ins have been developed to acquire data from different sources (camera, files, network, etc.).

3.3. Interaction plug-in

The Interaction plug-in allows the user to manage different 2D and 3D input devices without the need to change the application code. As shown in Figure 3, the data flow starts from the input devices and reaches the application through an Input Router component which is responsible of translating user inputs into proper commands.

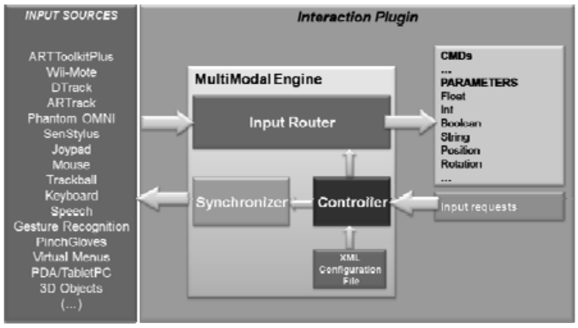

Figure 3. Interaction Plug-in Architecture.

The Multimodal Engine receives the events generated by the connected devices. Each event identifies itself and notifies its attributes and properties (such as type of data, range, time of acquisition). On the other side, the Application notifies the Multimodal Engine about the kind of command and/or parameters it is waiting for. The Input Router matches the received data and the input requests in order to provide the application with the correct command and parameter. The matching process is supervised by the Controller which uses the XML configuration file to determine which device is feeding the appropriate input for the current application request.

The Synchronizer component is used to ensure consistency between different input devices. As an example, it would update a graphical input box of a mobile device with a numerical value entered by speech recognition, in order to use both input techniques seamlessly and simultaneously.

3.4. Communication plug-in

The Communication Plug-in synchronizes the current application instance, running on the client, with the shared database (CAD and engineering analysis data). It detects changes in data and collects chat messages sent by other users connected to the database server. This plug-in allows a collaborative AR visualization, since each user can share a common set of digital data.

3.5. Multimedia Plug-in

Most applications in AR usually need to visualize large amounts of text/image/video annotations associated with objects.

Usually browser embedding into applications has been relegated mostly to provide user help or simple internet browsing functionalities. We have taken it a step further: our idea is to use existing and well-known web browser technology to build 2D interfaces. This gives the possibility to add and browse easily multimedia contents (Web 2.0) attached to engineering models.

To improve the collaborative environment the system may also present users a transparent chat window on top of the scene. This allows users to exchange ideas and comments while discussing the overall project.

3.6. Engineering Tools Plug-in

This plug-in handles the communication with any engineering tool coupled with the AR visualization environment. Typical engineering tools may include: CAD modellers, engineering simulators (heat transfer, structural CFD, etc.) In our framework all the data (geometry and feature models, simulation dataset, multimedia added information) are contained in an engineering database.

This plug-in queries the database for sensible data. These data are used to create new nodes in the scenegraph for visualization. The plug-in has also direct access to the coupled engineering tools (CAD modeller, FEM solver, etc.) and sends actions to allow geometry of simulation changes. Data and actions are synchronized.

4. TEST CASES

In this section we illustrate two engineering applications of the presented framework. The main goal is to evaluate the integration and user acceptance of the AR multimodal technology in real industrial scenarios. We ran our test using experts from local SMEs, who were not familiar with AR technology.

4.1. Engineering re-design

In this scenario, we would like to test the advantages of AR multimodal interface in providing product design insight and team working support. Industrial components are analyzed in practice using CAE on desktop (i.e. mouse and monitor) for simulating structural, thermal and fluid dynamic behaviour. Those tools clearly do not support well discussion and exchange of ideas among experts from different disciplines often distributed geographically. We tested a practical scenario of component re-design: a thermoplastic pot for household applications.

The first step of the design process was to acquire the digital model file and to import it in the framework. The model geometry was a shell structure of about 2mm thickness, manufactured by injection moulding process. The CAD file was originally generated using SolidEdge [12] modeller, and then imported in finite element analysis (FEA) software for evaluating the structural behaviour. We developed a specific “engineering tool plug-in” to retrieve the CAD data form the database, to run the FEA simulation and then to visualize the solution. Experts (6 technicians from local industrial compartment) could test different loads and constrains interactively.

During our test a discussion interested the pot handles (see Figure 4). Users applied loads on the handles and on the bottom surface as suggested by practical usage. The result of the simulation were displayed to all the team members to discuss critical points which can result in structural failure (see Figure 4). People from the design department (using CAD) and from the manufacturing section (using CAM) could use their own plug-in to access their software for visualization. The team had to decide if to increase the overall width or to design ribs stiffeners (not shown). Increasing width meant an extra cost (CAD software calculates total volume and weight), while adding ribs could be an issue during injection (as verified by CAM software).

The users can discuss the shared engineering data by annotating the model using shared labels. This task can be performed differently by each user according to his/her preferences and to the available devices. The multimodal plug-in manages the user input automatically at runtime. For example, one user can navigate his\her own virtual representation using a tangible 3D tracked interface, if available, or simply by mouse, trackball, etc.

After about one hour of collaborative discussion, the final solution converged to the use of ribs inside the handles. All the users were positively impressed by the communication tools and usability.

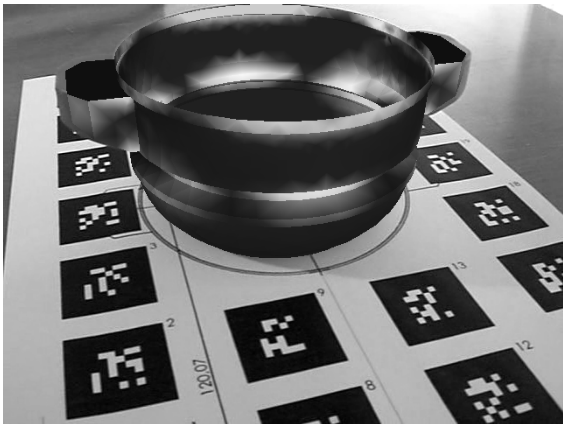

Figure 4. Real pot augmented with stress simulation.

4.2. Industrial plant augmented visualization

Design and maintenance of industrial plants can be very demanding in term of visualization and interaction, due to the complexity in geometries, extension in space, site integration and conflicts between different services and components. Our idea is to support the evaluation of industrial plans projects. The framework can assist teamwork by using the following scenarios: (i) meeting table, (ii) projection screen, (iii) on-site project visualization and (iv) augmented marketing.

In the first scenario experts gather around the meeting table and discuss a industrial plan layout printed on a paper sheet with tracking markers. Each user can augment the drawing using a laptop and its integrated camera. Important information can be added using a personalized visualization context. Each user can selectively visualize part of the design (i.e. hot water circuit), add labels or web pages and activate chat. Using a wireless connection each team member can move freely and his/her point of view is updated consequently.

The second scenario pictures a standing speaker presenting industrial layout to a technical audience. A large screen display is used and the participants access the presented data. They can suggest their opinion by voice, chat or shared annotations which are updated in real time among the members. In this specific scenario the multimodal plug-in allows to seamlessly integrate different interaction devices. For example, the presenter may use the Wii remote [11] as a mouse replacement when he/she is standing. With this device the user can specify a precise point of the model, add labels, select options and browse web contents. Text input can be provided in a multimodal way: PDA, wrist keyboard, voice recognition, etc.

The third scenario envisages the visualization of the virtual 3D model on the real site using geo-referenced tracking. Design errors and incoherencies can be spotted at early stage of design, with huge saving of resources.

The forth scenario depicts a marketing oriented application. The preliminary design has usually to be presented for approval. The use of 3D augmented visualization may help and speed up this phase. Instead of 2D drawing layouts or desktop based technical software, the contractor can experience a higher level of interaction.

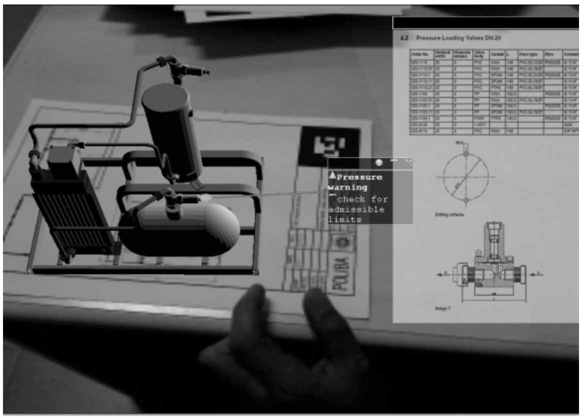

Figure 5. Industrial plant augmented visualization.

We tested a process plant layout and piping design of a chiller system (Figure 5). Basically chillers circulate fluid through a low-pressure piping system. All the users agreed about the benefits of the multimodal interface.

5. CONCLUSIONS

We have designed a novel framework for multimodal engineering applications using augmented reality. The main features of this architecture consist of: (i) plug-in based software architecture for high flexibility, (ii) multimodal module designed to manage and synchronize different inputs, (iii) hierarchical XML-based configuration.

We presented two applications of the proposed framework. We tested them with real industrial experts. After a first learning phase all the technicians involved in the experiment became very familiar with the interface. Wii remote navigation interface on large screen was appreciated by all users. Also the web browser potentialities were used above our expectation for retrieving technical data (i.e. technical pdf sheets). User complains regarding the marker based tracking jittering were successfully solved using multi-markers configuration. Users were interviewed by questionnaire with solid results. All of them were positively impressed by the communication tools and the usability of the interface. Most of them agree that the multimodal interface surpasses the traditional desktop based interface. Some of them asked for more functionalities in their specific applications (which can be done by improving plug-ins).

Future work will include the extension of the available plug-ins to meet user needs, and the support of different tracking systems (i.e., markerless tracking).

6. ACKNOWLEDGEMENTS

We would like to thank Stefano Cristiano for implementing\testing the visualization module.

7. REFERENCES

[1] Azuma R, Baillot Y, Behringer R, Feiner S, Julier S, and MacIntyre B. Recent Advances in Augmented Reality. IEEE Computer Graphics and Applications Nov 2001; 21, (6). p. 34-47.

[2] Halevy G, Hecht D, Reiner M. Multimodal Virtual Environments: Response Times, Attention, and Presence. Presence: Teleoperators & Virtual Environments (October 2006), http://www.mitpressjournals.org.

[3] ARVIKA Project. URL: www.arvika.de.

[4] Shin D. Communicating Designs through View Manipulation in Augmented Reality CAD. Independent Research Study, Civil Engineering, Purdue University; 2002.

[5] Dunston P, Wang X, Billinghusrt M, and Hampson B. Mixed Reality Benefits For Design Perception. Proceedings of 19th Int. Symp. on Automation and Robotics Construction; 2002; Gaithersburg, MD, US. p. 191-196.

[6] Vo MT, Waibel A. Modeling and Interpreting Multimodal Inputs: A Semantic Integration Approach. Technical Report, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, (December 1997) CMU-CS-97-192.

[7] Milota AD. Modality fusion for graphic design applications. Proceedings of the 6th international conference on Multimodal interfaces (2004) ACM Press

[8] Gutierrez M., Thalmann D., Vexo F. Semantic Virtual Environments with Adaptive Multimodal Interfaces. 11th International Multimedia Modelling Conference (MMM'05).

[9] Object-Oriented Graphics Rendering Engine, URL: http://www.ogre3d.org.

[10] NaviLibrary, URL: http://navi.agelessanime.com.

[11] Nintendo. URL:www.nintendo.com.

[12] Siemens. URL: www.siemens.com