iTACITUS – Novel Interaction and Tracking Paradigms for Mobile AR

Michael Zoellner, Dr. Didier Stricker, Gabi Bleser, Yulian Pastarmov

Fraunhofer IGD, Germany

Abstract

iTACITUS is a sixth framework programme project and aims to privide a mobile cultural heritage information system for the individual. By combining itinerary planning, navigation and rich content on-site information, based upon a dispersed repository of historical and cultural resources it enables a complete information system and media experience for historical interested travellers.

This paper covers the development of a "Mobile Augmented Reality (AR) Guide Framework" for Cultural Heritage (CH) sites. The framework delivers advanced markerless tracking on mobile computers as well as new interaction paradigms in AR featuring touch and motion capabilities. In addition to visual components like annotated landscapes and superimposed environments the framework will feature a reactive accoustic AR module.

The markerless tracking is based on optical flow. Only pure camera rotation is estimated, assuming the user is staying in place. The system is capable of running at real-time frame rates on mobile computers which allows for wide range of applications both indoors and outdoors.

The visual AR application gives users the ability to explore the digital information about the site in a very straight way. By holding a mobile computer in front of a point of interest the user immediately gets further information as AR overlay on the screen. This effect appears like looking through the mobile computer's display and seeing the real world enhanced with virtual objects or information. Virtual Objects are reacting on the user's position. Direct interaction with the virtual objects is possible for example by touching the display or shaking the mobile computer.

Spatial accoustic AR transports a place's original ambience. While walking through a room the user gets an accoustic impression about how the place has been before, about people and their activities by listening to the conversations and the environment sounds. Due to the 3D position of the sound the user will create himself an spatial image of the situation.

CR Categories and Subject Descriptors: I.3.1 [Input Devices]: Motion Detection, Touch Screen; I.3.6 [Interaction Techniques]: Motion Interaction, Touch Interaction; I.3.7 [Three-Dimensional Graphics and Realism]: Annotated Landscape, Superimposed Environments; I.3.8 [Applications]: Mobile AR Framework for CH Site

1. Introduction

iTACITUS is a sixth framework programme project and aims to provide a bespoke cultural experience for the individual, based upon a dispersed repository of historical and cultural resources, enabling both location-based and location-independent services. The services developed will include audiovisual, Augmented Reality (AR) and organisational components, and will be delivered in a flexible and timely fashion.

iTACITUS investigates new methods of exploring cultural heritage sites with mobile computers. Using emerging technologies, iTACITUS aims to provide the mobile user with the most complete Augmented Reality experience possible, allowing dynamic, user-directed walk-throughs of the historical place of interest on their mobile computer. These advanced features are combined with itinerary planning, navigation and on-site information into a full featured cultural heritage guide based on open standards and public information.

Our part in the project is the development of a Mobile Augmented Reality Guide - an easy to use and flexible information system for retrieving information on-site based on advanced novel interaction and tracking paradigms.

With Augmented Reality we are entering a complete new way of visualization in spatial environments. Visualization in AR is no longer a projection on separate flat surfaces like monitors, PDAs or projection screens, but places context sensitive digital data into the real world and thus creates annotated landscapes and superimposed environments. Like placing signs and markers in cities and on roads one is able to place information right on the spot where it affects. Further more these digital signs are able to react to external influences like changing data sources or user interactions. The mobile computer's display becomes a filter screen that reveals hidden information by looking through it. In addition to the visual module a spacial acoustic Augmented Reality module recreates acoustic scenes and moods of places in the past. This is a big step but also a challenge for designing user interfaces and visualizing information.

These features are working at speacially prepared spots at Cultural Heritage Sites. These areas will be visually accentuated and will automatically trigger the mobile computer to give the user feedback as soon as he enters the prepared area.

There have been several approaches of developing mobile information systems for cultural heritage sites in the past. Early projects had to fight with technology and hardware restrictions. Like one of the well known early projects Archeoguide [VIK*02] where ancient greek architecture and ancient Olympic sports events were displayed as visual overlays through a head mounted display direct at the site (Figure 1). Large and heavy bagbacks with high technology gear, large GPS antennas and heavy uncomfortable head mounted displays were needed in order to reach the goal of positioning the user at the site and displaying context sensitive information on the spot. Others tried it with small and mobile PDAs had to struggle with low computing power, short battery life and the need for developing custom software.

These projects were milestones in mobile cultural heritage visualization via Augmented Reality but never reached a product stadium due to the technical restrictions, the cost of the hardware setup per user at that time and thus the logistics needed at cultural heritage sites for distributing and maintaining the complicated hardware. But the success of these applications was only a matter of time: increasing computer power, affordable standard devices (i.e. ultra mobile computers) with a long battery life and the development of advanced markerless tracking algorithms.

Figure 1: *2002 - Archeoguide Application Scenario*

With a new generation of mobile computers like the UMPC and its successors the ultra mobile computers Sony UX and OQO the goal of mobile Augmented Reality at cultural heritage sites finally seems to become achievable from the hardware side. These devices feature standard operation systems like Linux (x86) or Windows (x86) and thus do not need separate development like on embedded hardware like PDAs with Windows Mobile Operation Systems. But these devices will only fill the gap between todays mobile computers and full featured mobile phones in the future which everyone will carry around and thus will be the ideal platform for mobile Augmented Reality.

Industrial Augmented Reality applications like digital manuals or CAD verification are often playing in a well known area or in a prepared environment. Unlike at Cultural Heritage sites it is normally not possible to change the environment in order to prepare the area for user tracking. There are comprehensible restrictions for changing the site's appearance by adding visual prominent elements like markers or flashing devices. Furthermore in most cases the area is too large and the budgets too small to install ultrasonic or magnetic tracking systems. These are powerful arguments for a markerless visual tracking method for Augmented Reality on cultural heritage sites. It is the only solution that is completely independent from artificial elements in the environment and thus fits the needs and restrictions of cultural heritage sites.

The technology we are developing in order to create a seamless Augmented Reality experience on mobile computers is a markerless tracking based on optical flow: A "Visual Feature Tracking Compass". Point features are tracked using the well-known Kanade-Lucas-Tomasi Tracker (KLT). Only pure camera orientation is estimated, assuming the user is staying in place.

The visualization is based on open standards (X3D). Not only to guarantee seamless integration of existing media like 3D reconstructed architecture. But also to design the applications with the long term availability and usability for digital archives in mind.

2. Project Overview

iTACITUS aims to provide a bespoke cultural experience for the individual, based upon a dispersed repository of historical and cultural resources, enabling both location-based and location-independent services. The services developed will include audiovisual, VR and organizational components, and will be delivered in a flexible and timely fashion.

iTACITUS will investigate new methods of representing historical sites virtually. Using emerging technologies, iTACITUS aims to provide the mobile user with the most complete AR and VR experience possible, allowing dynamic, user-directed walk-throughs of the virtual place of interest on their mobile computer.

iTACITUS will allow the customization of personalized walking and public transport tours in urban areas. A tourist will initially have an interactive session with iTACITUS, during which an itinerary is designed on the basis of the expressed interests and preferences, including the relevant public transport information and walking directions. Once the tour has been created the tourist will be guided through the areas of interest whilst relevant information will be presented by iTACITUS on the user's mobile computer at the appropriate points. iTACITUS will prompt the user should they fall behind schedule or wish to spend more time at a particular attraction, and shall be flexible enough to allow for replanning of the itinerary at any point.

iTACITUS will develop a system that local authorities will be able to employ to encourage cultural tourism.

The European project partners feature a broad range of knowledge and experience in the area of navigation, transportation and cultural heritage: British Maritime Technology (BMT, United Kingdom), TXT e-Solutions (Italy), Fraunhofer Institute for Computer Graphics (IGD, Germany), National Technical University of Athens (NTUA, Greece), Hampshire County Council (HCC, United Kingdom), Centro Studi sui Sistemi di Trasporto S.p.A. (CSST, Italy).

3. Scenarios / Situations at Cultural Heritage Sites

In order to evaluate the project's applications we have chosen two Cultural Heritage sites that will use iTACITUS software from January 2008 on: Winchester Castle in Hampshire County, Great Britain and Reggia Venaria Viale near Turin in Italy (Figure 2).

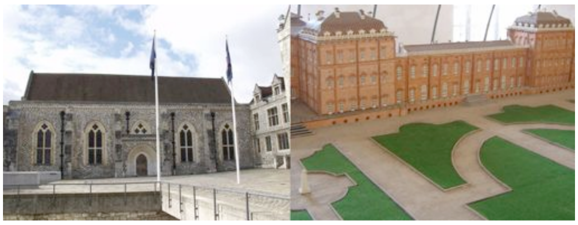

Figure 2: *Winchester Castle and Reggia Venaria Viale*

The typical scenario will look like this: A visitor equipped with a personal or a rental mobile computer enters the cultural heritage site. The site features completely recreated buildings and artifacts but also some partly recreated buildings and ruins. He walks around and reads some public available websites or listens to a podcast about the site on his mobile computer until he reaches a marked spot on the site where he immediately gets a signal about a point of interest (vibration, sound, visual alert) by his mobile computer. Looking on the screen he recognizes that the iTACITUS software switched from the browser based information mode into full screen Augmented Reality mode. The video of the computer's camera on the front is displayed in real time on the screen. It seems like it is possible to look through the display.

While moving the UMPC around and looking through the display the visitor recognizes elements on the screen that are not existing in the real surrounding. They seem to stick on certain objects in the real world. These elements are virtual information overlays in the form of icons describing the hidden content. As soon as the user is focusing on an overlay above one of the statues in the area, deeper information about the underlying object pops up. The information consists of images showing the original appearance of the statue and a short readable text. At this stage the information overlay and the camera image can be "freezed" in order to concentrate on the content. The information overlay disappears when the user moves the computer away.

In front of a ruin the computer again tells the visitor about available information. Again the display shows the camera image. While looking through the display and moving it over the ruin the visitor recognizes the difference between the real surrounding and the image on the screen. On the screen the ruin appears in its original appearance. All the broken statues are standing again on the pillars in front of the building. In its center water is running from a small fountain.

In order to create this imagination photo realistic 3D models of the former building and the statues are mapped on the real objects.

While walking through a long hall inside a building the computer again remembers the visitor about available information. This time the display does not show the video image. Instead of visual overlays the visitor hears virtual sounds that are located in the area. The sounds are changing their intensity as he walks around in the hall. He listens to conversations between invisible people from the past. Doors and the floor are creaking from different directions. As soon as the visitor leaves the room the sounds are disappearing.

4. Interaction Paradigms

During the last years some interesting interaction paradigms digged up the computer games industry. Intuitive interactions like touching, moving, tilting, shaking have been around for years and used in many experimental applications. Nintendo put these features in affordable hardware and created exciting and simple applications around it and thus established a new era of computer games. The Nintendo DS handheld gaming console with its touch screen interface and the Nintendo Wii with its motion sensitive “WiiMote” controller are outselling their competitors.

In iTACITUS' Augmented Reality module we are adapting and improving these successful and now well known interaction paradigms bringing them to another level with Augmented Reality. People using the application at cultural heritage sites will simply know how to interact with digital overlays out of the box without an additional learning process because they already have done it before at home. Furthermore we are introducing a spacial acoustic augmented reality mode that recreates an interactive sound simulation of scenes from the past.

4.1 Visual Augmented Reality Application

The visual Augmented Reality application gives users the ability to explore the digital information about the site in a very straight way. By holding a mobile computer in front of a point of interest the user immediately gets further information as an Augmented Reality overlay on the screen. This effect appears like looking through the mobile computer's display and seeing the real world enhanced with virtual objects or information.

Within the cultural heritage site several spots are prepared for this application. The user will recognize these visually marked areas. Once he arrives at the spot he will get a message from the system and the iTACITUS application switches from the browser based information mode into Augmented Reality full screen mode. From here on the visitor is enabled to explore the area around him with his mobile computer.

On the display of the mobile computer the user sees a live video from the camera on the front. It seems that he can look through the device's screen. He moves the computer around until a virtual overlay appears (Figure 3). These overlays are divided into two visual categories. On the one side there are objects that fit into the scene like recreated paintings, statues or architecture. On the other side there are abstract context sensitive information overlays showing images, texts and videos about that spot. The most simple overlays will reveal detailed information on the screen as soon as the user stops and focuses the overlay. The information automatically disappears when the user moves on and explores other parts of the area. The selection task works simply by focusing on an object and does not need an additional action like clicking or pressing buttons on the device. Therefore no learning and introduction process is needed in order to show the user how to reveal the underlying information. The first time he rolls over an information layer all questions about how to use the application will be answered.

Figure 3: *See through AR with UMPC*

Other more advanced overlays are demanding more interaction by the user in order to create a much more fetching story by interacting with the environment itself. This is achieved by using the mobile computers touch sensitive screen and the motion determination of the "Visual Feature Tracking Compass".

Touch interaction enables a direct physical interaction with the virtual object in the scene. This interaction starts with a simple tip for its activation. Like tipping on a real person's shoulder for starting a conversation this activation starts for example a conversation with a virtual character like a virtual tour guide or presses a button on a wooden chest and opens its lid.

Another feature is the rubbing and shaking effect. It enables the user to rub hidden information free like removing dirt from an archaeological artifact with a brush or shaking it until the dirt falls off. By moving with the finger fast over the overlay more and more of the digital dust withdraws and exposes the underlying information. In a variation this rubbing effect becomes a tickling effect when a virtual character tries to hide information.

The well known drag and drop interaction works in this Augmented Reality environment exactly like it works on our computer desktop. Thus applications using this feature are obvious. Putting one virtual object onto another activates a certain task defined by the second object and articulated by its appearance. This feature enables possibilities to pick up virtual objects like keys, take them with you and use them in another place for opening doors and chests. Also games about solving puzzles by sorting objects in a certain way are possible.

Figure 4: *Interaction via motion determination*

While touching the display and thus interacting with virtual objects acts as a direct interaction, motion determination and its effects works as indirect manipulation of the scene. Several unique gestures can be detected via motion determination by the "Visual Feature Tracking Compass" and are implemented in the framework as standard interactions: Shaking, nodding, leaning and rotating (Figure 4).

These physical interaction modes do have multiple effects depending on scene and story:

Shaking: Like the other modes it does not have a fixed effect but can be interpreted in different ways. Shaking the head and communicating disapproval could be a reaction on a question asked by the system or a virtual character. But it can also be interpreted as shaking away dirt from an archaeological object like the rubbing effect of the touch interface. In this case both interface modes (touching and physical) can be combined by selecting an object with the finger and shaking it.

Nodding: The nodding effect works quite similar in order to approve a question or select an object and move it up and down.

Leaning: By leaning the mobile computer to the left or to the right side it is possible to define a direction the selected object should move to. Here the physics functions of our Mixed Reality framework (instantreality) come into play and improve the level of realism of the scene.

Rotating: Moving the mobile computer circularly in the right or left direction defines a turning or a roll up motion of the object. This will for example be used to roll up a cable winch, opening a safe or controlling a rope bridge.

The technology for the position tracking and motion determination of the mobile computer and thus for the placement of the virtual objects is called "Visual Feature Tracking Compass". It uses computer vision algorithms to calculate the orientation of the mobile computer without any preparation of the environment (like markers, etc). The tracking software is an enhancement of Fraunhofer IGD's VisionLib computer vision software.

4.2 Acoustic Augmented Reality Application

Acoustic Augmented Reality has always been there. At airports, train stations or in shopping malls automatic messages have been transported via loudspeakers into the area. This acoustic information layer overlays the real sound layer like visual augmented reality overlays the users' field of view.

Figure 5: *Acoustic and visual illustration of the scene*

Our application uses spatial acoustic AR for transporting a place's original ambience. While walking through a room the user gets an acoustic impression about how the place has been before, about people and their activities by listening to the conversations and the environment sounds. Sounds that don't exist anymore or that are too silent to hear with all the civilization noise around are revived. Like the church bell some streets away or the noise of wooden chariots in the cobblestone lane where today cars are driving with rubber tires on asphalt. Due to the 3D position of the sound the user will create himself an individual spatial image of the situation. Like every reader of a book has a slightly different image of the story in his mind each visitor will have another image of the soundscape in his mind. But in order to assist the visitor's fantasy original paintings, drawings and documents are displayed on the screen (Figure 5). By showing abstract impressions and additional information they are transporting a glimpse of the scene and the time.

The placement of acoustic overlays does not have to be as accurate as visual ones. Slightly misplaced sounds will not be recognized by the user. Thus in comparison to the visual modules the acoustic module is not fixed at certain prepared points with a small radius but plays in a broader area. The tracking technology for this application is dependent on the CH site's environment. Outside the rough positioning is achieved by GPS and the orientation by the "Visual Feature Tracking Compass". For indoor situations without GPS reception there are alternative concepts. A rough "Bluetooth Proximity Detection" is already complete. It consits of bluetooth beacons that are placed at the points of interest. Due to the beacon's adjustable field we can define its active radius. This is achieved by setting the bluetooth module's field strength and modification of its antenna. The next step will be a more accurate approach based on bluetooth triangulation.

In addition to the intuitive and fanciful aspects this module gives cultural heritage sites also a cheaper alternative to expensive 3D recreations and characters. The success of podcasts shows us that informative shows created by amateurs feature a very high quality in content and technique. But even the professional production of an interactive audio play goes with a very low budget.

4.3 Augmented Reality Information Layers

The visual Augmented Reality application concentrates on two types of established information layers for Augmented Reality visualizations: Superimposed Environments and Annotated Landscapes. Both methods have been evaluated in our research projects like the Messel Pit and the Augmented Reality Telescope Telescope.

4.4 Superimposed Environment

This visualization method is mostly used for integrating spacial virtual objects into an Augmented Reality scene. Typical applications are architectural visualizations, recreated cultural heritage artifacts and buildings and virtual characters.

Figure 6: *Fresco overlay fits in the scene*

There are two possible opposed visualization qualities that guarantee a convincing integration of the virtual objects into the real environment (Figure 6). The objects have either to be displayed in an non photo realistic way like cell shaded or sketch rendered in order to emphasize the artificial impression of the object and to create a high contrast to the real environment. The second method is to display the objects in a very high realistic way in order to create a seamless integration into the environment. Both methods have to assure realistic lighting conditions. Otherwise the object will silhouette against the environment and ruin the Augmented Reality effect.

4.5 Annotated Landscapes

Normally fully recreated and textured 3D models are not available and most cultural heritage sites can't afford their creation. Unlike Superimposed Environments Annotated Landscapes are based upon existing media like images, sounds and videos. This media is appended direct at real objects and points of interest in the environment. Abstract icons are placed next to the object giving a good impression about the kind of information they are containing (Figure 7).

Detailed information is hiding behind the icons in order to avoid an information overload in the scene because of too much information layers flying around. Therefore the underlying information is only displayed by selection. We have chosen a simple “roll over” effect like described in section 4. The detailed information contains media like images, sketches, paintings and movies placed in a frame next to the icon.

Figure 7: *Annotation hints at missing fresco*

Like with all Augmented Reality visualizations we never know the background lighting conditions of the scene the information layers are displayed on. Thus it is important to choose a visualization style that guarantees high contrast in different lighting and background situations. We had good results with drop shadows and outlines (Figure 7).

5. Visual Feature Tracking Compass

In order to correctly estimate the observer's position, optical flow tracking is used. Computing the displacements of well distinguishable points in the live image flow provides enough information to compute the camera pose with 3 DOF (degrees of freedom), representing it's relative orientation according to the camera's optical center. Once the camera orientation is successfully computed, virtual objects can be correctly overlaid on the live image at real-time speed.

The stability of the features engaged in this process is important for successful optical flow tracking. But using too complex structures can lead to high computation times, thus rendering the tracking unusable in real-time conditions. This is the reason why a very precise trade-off between complexity and stability has to be made. After evaluating different possibilities a simple but never-the-less precise and highly repetitive point detector has been selected. This point detector is described by Rosten at al. [RD05]. Features are selected at corners in the image, where at least 75% of the surrounding pixels have intensities distinctively lower or higher than the center point of interest. To prevent an overflow of features that can fit this description, we force a minimum pixel distance during the construction of the feature map, where candidate features are considered in decreasing order of their “cornerness”. For the purpose of successful optical flow estimation some reasonable number of features have to be present, practically at least 10 to 20 strong features must be detected for stable orientation estimation.

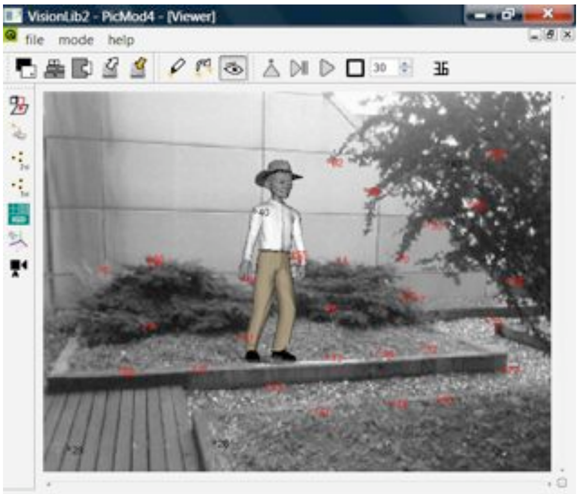

Figure 8: *Features found in the current image*

Feature tracking describes the process of following points in subsequent images that correspond to the same physical point in the real environment explored (Figure 8). This means correspondences between features from subsequent frames of the image stream must be built and maintained as long as possible. The longer the features are successfully tracked the more stable is the camera orientation estimation. This is why the point detector is required to be highly repetitive. Of course this condition is vital for the tracking algorithm, too. From the large palette of possible approaches to that problem the Kanade-Lukas-Tomasi tracker described in [TK91] and [ST94] has been chosen. The idea proposed by Tomasi at al. benefits from relatively low computation times and strong spatial correctness. In short, tracking is based on building an image pyramid from the live image with levels of decreasing scale and increasing Gaussian filtering for smoothing out smaller artifacts in the image. Features are sought in a local neighborhood around their previous positions. This process is done on all levels of the pyramid starting at the coarsest level down to the finest one. This enables to follow features over longer distances with high precision. In case not enough features are present for the pose computation new ones are selected from the feature map and used in the tracking further on.

The procedure described up to now provides an effective way of estimating the optical flow on a frame-to-frame basis. However, the lack of context requires some more temporal information to be collected and used to ensure that the tracking does not drift over time. Therefore some features detected in the beginning of the tracking process are selected and marked as *persistent* and these are not replaced or removed from the current set of features throughout time. This set of anchor features helps to stabilize the tracking process.

Once enough features have been tracked, the camera rotation **R** around the center of the camera coordinate system can be estimated. This can be assumed to coincide with the center of the world coordinate system the camera resides in. As we are interested in a pure rotation, camera motion can be represented by a *homography at infinity* H_{inf} as described in [HZ01]. However in our implementation the homography matrix is never computed explicitly. Instead a rotation is estimated based on non-linear minimization of the re-projection errors of the tracked features. The estimate of the previous frame serves as starting point.

The general model of a 6 DOF camera projection is described by the following equation :

$$p=K[RP+t],$$

where **P** is the position of a given point in world coordinates, **K** is the matrix of the camera intrinsic parameters, **R** is a 3x3 rotation matrix, **t** a translation vector and **p** the homogeneous projection of **P** in the image. The intrinsic camera parameters matrix has the following form :

$$K=\begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix},$$

where $\mathbf{f_x}$ and $\mathbf{f_y}$ express the focal length in x- and y-direction, **s** is the skew factor of the pixel matrix and $\mathbf{c_x}$ and $\mathbf{c_y}$ denote the optical centre in pixel coordinates.

Assuming that $\mathbf{t} = (0, 0, 0)^T$ we can derive a formula that relates the displacement of a given feature point in the image plane to the relative camera rotation. Let's assume we have a world point **P** with the image coordinates $\mathbf{p_{init}}$ taken from a camera with orientation $\mathbf{R_{init}}$. It's coordinates in the world coordinate system are then given by the following equation :

$$P=R_{init}^{-1}K^{-1}p_{init}.$$

Now let's assume $\mathbf{R_{act}}$ is the orientation of the current frame, hence if both $\mathbf{R_{init}}$ and $\mathbf{R_{act}}$ are correct the same point **P** must be present in the image taken from orientation $\mathbf{R_{act}}$ at coordinates $\mathbf{p_{act}}$ given by the following equation :

$$p_{act}=KR_{act}R_{init}^{-1}K^{-1}p_{init}.$$

Using this formula the re-projection error can be estimated as the distance between the actual feature coordinates and the projected coordinates. We make the assumption that the motion between consecutive frames is small enough so that the non-linear least squares minimization of the re-projection errors of all successfully tracked features converges to the correct orientation, if the orientation of the previous frame is used as an initial guess.

Minimizing the re-projection errors between the current and the last frame would reduce the required amount of storage significantly but inevitably lead to drift and jitter in the camera orientation estimate, due to the jitter in the localization of the features. Therefore, we minimize the error between the frame where a feature was first detected and the present frame. As the features usually originate from different frames the jitter is smoothed out. Moreover the presence of many features helps for a fast minimization convergence and keeps the whole processing time low enough to achieve a constant frame rate of 30fps on a P4 2.4Ghz and satisfying 15fps on a Sony Vaio Micro PC with Intel Core Solo 1.3Ghz - the intended target of application.

6. Software Platform

The complete iTACITUS software is browser based and thus platform independent. The main target platform is a Sony UX1XN Micro PC with Windows Vista but in the future mobile phones and novel handheld devices will be supported, too. The Augmented Reality Module bases upon two integrated software parts developed at Fraunhofer IGD's Department for Virtual and Augmented Reality: instantreality framework and VisionLib.

The instantreality framework is a high-performance Mixed Reality (MR) system, which combines various components to provide a single and consistent interface for AR/VR developers. Those components have been developed at the Fraunhofer IGD and ZGDV in close cooperation with industrial partners over the last years. This project has gone public in april 2007 with the free available instantplayer, a platform independent desktop X3D player.

The framework offers a comprehensive set of features to support classic Virtual Reality and advanced Augmented Reality equally well. The goal was to provide a very simple application interface while still including the latest research results in the fields of high-realistic rendering, 3D user interaction and total-immersive display technology. iTACITUS will feature instantreality's browser plugin that is in development right now and will be released in fall 2007.

VisionLib is a complementary tracking system for Augmented Reality. It combines different tracking algorithms modular in order to create the ideal result in different scenes. iTACITUS' “Visual Feature Tracking Compass” is a custom module for markerless tracking at cultural heritage scenes. VisionLib is integrated as instantvision in the instantreality framework. It enables the less experienced developer rapidly integrating advanced computer vision in X3D scenes.

References

[VIK*02] Vlahakis V., Ioannidis N., Karigiannis J., Stricker D., Daehne P., and Ameida L. 2002. Archeoguide: Challenges and solutions for a personalized augmented reality guide for archaeological sites. IEEE Computer Graphics and Applications, pages 34–47, Sept.-Nov.

[RD05] Edward Rosten and Tom Drummont. Fusing Points and Lines for Fast Feature Tracking. IEEE Internation Conference on Computer Vision, pages 1508-15011. ICCV, October 2005

[TK91] Carlo Tomasi and Takeo Kanade. Detection and Tracking of Point Features. Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.

[ST94] Jianbo Shi and Carlo Tomasi. Good Features to Track. IEEE Conference on Computer Vision and Pattern Recognition, pages 593-600, CVPR '94, June 1994

[HZ01] Hartley, R.~I. and Zisserman, A. Multiple View Geometry in Computer Vision. 2001, pages 327-328, Cambridge University Press, ISBN: 0521623049

Citation:

Zoellner, Michael; Stricker, Didier; Bleser, Gabriele; Pastarmov, Yulian:
iTACITUS - Novel Interaction and Tracking Paradigms for Mobile AR.
In: Arnold, David (Ed.) u.a.; The European Research Network of Excellence in Open Cultural Heritage (EPOCH):
Vast2007 - Future Technologies to Empower Heritage Professionals : Short and Project Papers from VAST 2007.
Budapest : Archaeolingua, 2007, pp. 110-117