Mixed prototyping environment with different video tracking techniques

F. Bruno, A. Angilica, F.I. Cosco, M.L. Luchi, M. Muzzupappa

Università della Calabria - Dipartimento di Meccanica - Rende (CS)

Article Information

Corresponding author:

Fabio Bruno
Tel.: +39 0984 494623
Fax.:+39 0984 494673
e-mail: f.bruno@unical.it
Address: via P.Bucci 44/c 87036 Rende(CS)

Abstract

Mixed Prototyping has been recently proposed as an interesting design tool able to combine virtual and real objects, in order to allow both designers and final users to evaluate and assess multiple design choices. This paper describes a Mixed Reality environment that deals with several issues that usually arise in Mixed Prototyping: the reliable simulation of the product behaviour; the mutual occlusion between real and virtual objects; the interpretation of the user's gesture while she/he is interacting with the elements of the product interface; the composition of virtual graphics within the real scene through a fast and robust tracking methodology. In particular we focus on the tracking methodology considering two different solutions based on fiducial markers and natural features, respectively. We have considered different scenarios of use, referring to four typologies of products. Each one presents its intrinsic challenges: small hand held products (e.g. mobile phones, remote controllers, etc.), medium sized products (e.g. small appliances), large interactive products (e.g. major appliances) and large non-interactive products (e.g. furnishings). We propose a comparative study that aims to identify the tracking system that better fits each application.

1   Introduction

The potentialities of Virtual Reality (VR) in the area of product design have been investigated for several years. A lot of work has been addressed to the development of new devices, setup, techniques and methodologies able to allow engineers and designers to make a profitable use of VR for the development of new products. All these works have put in evidence the advantages that one can obtain through the virtualization of the prototyping process and through the use of VR for improving the interaction with the virtual prototypes. On the other hand, several problems and challenges have arisen, in particular about the ergonomics of the devices and, consequently, about the acceptability by the users. VR technologies are often not so comfortable or easy to use and this make them difficult to use in such analyses (e.g.: usability) that involve final users.

Recently, Mixed Reality (MR) technologies are emerging as a promising solution for the problems that are typically related to VR environments and in particular to some applications in the area of virtual product validation. MR environments are characterized by the simultaneous presence of virtual and real objects that are somehow able to interact with each other. Augmented Reality (AR) may instead be considered as the augmentation of a real world through the superimposition of digital objects over a real scenario.

AR and MR are often implemented through video see-through displays equipped with one or two cameras. These cameras are used both to acquire the real environment that represents the background of the augmented scene and to track the position and the orientation of the user head with respect to a world reference system. The video tracking techniques, based on the recognition of fiducial markers, are the most wide-spread technology used for AR applications. The availability of an Open Source software library like

ARToolkit, that is sufficiently reliable and easy to use, has made possible to implement AR application with a limited effort. On the other hand, marker-based video tracking presents several limitations related to the need to prepare the environment and suffers in terms of precision and reliability. An alternative to marker-based tracking is natural feature tracking, a technique based on the recognition of natural occurring features that are used to track the movement of the camera with respect to the world. Natural feature tracking gives better performance, does not require any preparation of the environment but requires more computational resources.

The goal of this paper is to analyze the two video tracking approaches (marker-based and feature-based) in two specific contexts: Mixed Prototyping (MP) and Augmented Prototyping (AP) practices. These consist on the analyses of digital prototype using MR and AR environments, respectively. The paper will also try to give some guidelines for the development of MP and AP applications, putting the emphasis on the issues related to the video tracking.

The paper presents four case studies that are the starting point for our considerations. We have identified four macro-typologies of products with which the user can interact:

- small (as mobile phones or remote controllers);
- medium (as small appliances);
- large and interactive (as major appliances);
- large and non-interactive (as furnishings).

In some cases, we have supported our opinions with the quantitative measurement of some performances (like calculation time and fault time).

This paper is organized as follows. The paper starts from a state of the art about MP and video tracking approaches. Section 3 presents the guidelines that we have defined for the development of applications based on MP and AP practices focusing on the choice of the video tracking technique. Section 4 and 5 describe the hardware set up and the approach that was followed to

implement the case studies that are described in section 6.

2   Related Works

The MP approach was recently presented as a practical solution able to involve the sense of touch in the interaction with a virtual prototype [1-8].

This design approach makes use of both physical and virtual components. The physical components, usually realized by means of Rapid Prototyping, give to the user the possibility to touch the object. On the other side, the MR displays superimposes missing parts (details, additional data, appearance of the objects surface, etc..) coming from the virtual model, improving the perception that the prototype really exists and works.

A correct tracking of the user position is crucial for MP applications. In order to perceive a consistent visualization of the mixed prototype, a robust registration of the user and the physical prototype, both immersed in the real scene, is needed. In facts, any misregistration will prevent the user from perceiving virtual and physical channels/stimuli as blended in a realistic prototype [9].

In the field of Mixed/Augmented Reality, the topic of tracking was deeply investigated during the last decade, and tracking techniques can be classified in three main categories: sensor-based, vision-based, and hybrid [10].

Among all, the vision-based techniques are widely used because they don't require specific and more expensive hardware. In fact, they can estimate the camera pose directly from the same imagery observed by the user. The vision-based category can be divided into two main areas: *marker-based* solution and *markerless* solutions.

The *marker-based* approach relies on putting in the scene some artificial marker, that can be easily detected through image processing methods. *ARToolkit* [11] is one of the most commonly used open-source libraries implementing a marker-based optical tracking. It was lately upgraded to *ARToolkit plus* version [12] and recently superseded by *Studierstube* project [13]. Similar solutions were proposed by Fiala in [14-16]. Despite these evolution efforts, marker-based approach still presents some challenging problems like the lack of robustness against rapid motions/rotations of the user's head, low accuracy and jittering.

Conversely, *markerless* solutions are a set of promising technologies, that may solve all these problems in the near future, without adding artificial markers to the scene and therefore preserving the real scenario.

A suitable approach is given by the natural features tracking methods [17-22], where the basic idea is to use naturally occurring features instead of the artificial markers. In a few recent works, some researchers have proposed a set of descriptors, more robust and efficient if compared to geometrical fiducials such as points, lines, edges or textures: the SIFT [23], SURF [24], or Ferns [25-26].

Another suitable alternative is given by the model-based tracking methods, that explicitly use a model (such as a CAD one) in order to retrieve the object features to be tracked. A widespread solution is the use of *SLAM* methods [27]: they construct edge-based models from image sequences and use them for tracking without needing any prior knowledge of the world.

An alternative to *SLAM* was proposed by [28]. Their Parallel Tracking and Mapping (*PTAM*) method is able to produce detailed maps, with thousands of features tracked at frame-rate. Recently it was successfully extended in [29] where a complete 3D model of a cluttered indoor environment can be reconstructed from a live handheld camera in a few seconds, and then used as a structured reference for the tracking.

Despite to their affirmation in the field of AR/MR, most of MP researches are based on marker-based approaches.

In [4] the authors present a MR set-up where a tangible dummy, realized in RP, is animated through the augmented reality visualization. The set-up uses an hybrid tracking approach fusing *ARToolkit* [11] with an electromagnetic tracker. *ARToolkit* is used to track the position of the dummy. The magnetic tracking sensor placed on the right thumb allows the system to draw an avatar of the tracked finger, that is used both for resolving the occlusion problem and to determine the intention to press a button of the product. Their tests have given good results, but are limited to the case of hand held products.

Similarly, in [5] the authors use rapid prototyping to create a physical mock-up and they paste the markers needed for *ARToolkit* tracking on the physical model. The interaction with this augmented tangible prototype is done through a paper pen on which some *ARToolkit* markers are placed.

In [6] the virtual object is superimposed to a blue foam mock-up, but it doesn't simulate the behaviour of the prototypes. A similar application for the evaluation of car interiors is proposed in [7]. The authors put in evidence that the possibility for the user to see his/her own hands while touching the steering wheel and seeing the virtual images enhances the feeling of reality and the sense of presence.

Verlinden and Horváth [1-2] described a multiple case study addressed to assess usability and impact of AR based prototyping technologies. They put in evidence how the shortcomings in using physical prototypes and the bottlenecks in the design process technologies may be reduced using a MP approach. They adopt the projection-based AR that uses projectors to cast computer imagery directly on physical objects. They do not discuss about the tracking issue because they put the rapid prototype on top of a rotating table equipped with an encoder that allow to compute the projected model according to the rotation of the table.

In [8] a reference framework for the mixed prototyping practice is presented. The authors report several examples where this practice has been effectively used for rapid design review of new products and in particular for information appliances. They mainly investigate the problem of positioning information appliances within systems, and for the evaluation of ergonomics aspects of interactive devices. All their solutions rely on a video-see-trough HMD and a commercial optical tracking solution.

3   Guidelines for Mixed and Augmented Prototyping

Starting from the reference framework described in [8] and the famous Milgram continuum [30], we have reinterpreted these representations, collocating the four different typologies of prototyping as shown in figure 1.

In [8], the authors propose a 3D framework based on three independent domains: user, prototype and interaction. On the other hand, Milgram classifies different techniques using a single criterion: the environment.

For our purposes, we propose a 2D representation in which the two domains used are: environment and prototype. In this way, we can classify the four different typologies of prototyping practices.
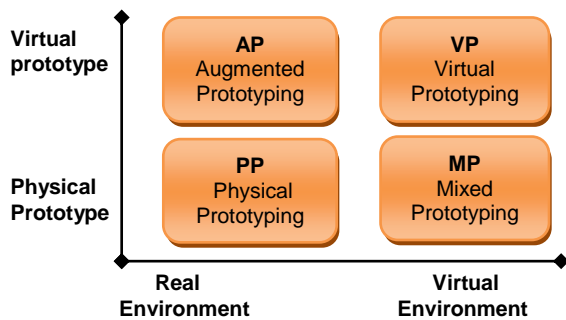
Fig. 1 Classification of various prototyping practices

Physical Prototyping (**PP**). This is the typical activity used in product development process, in which the environment and the prototype are real. This practice is the most reliable because the prototype is a very good approximation of the final product and also the interaction between the environment and the product can be analyzed with a good realism. In this case, there are no limits for the application of this technique regarding: dimension of prototype, types of assessment and interaction and numbers of users involved.

Virtual Prototyping (**VP**). This situation represents the typical practice used during the early phases of the design process. The prototype and the environment are virtual, the interaction is commonly based on vision (indirect interaction [8]), direct interaction can be obtained through data-glove [31] or through haptic device [32] so involving also the sense of touch. In VP the number of users that can be involved in the analyses depends on the kind of VR technologies adopted. Virtual room with large display are used for design review sessions with tens of users at the same time, immersive set-up with HMD, instead, can be used only by one user at a time. VP typically deals with several kind of different analyses (structural, CFD, crash, ergonomics, etc.) so it can be used to assess several elements including aesthetic and usability.

However, in this paper we do not deal with these two approaches, but we will refer only to applications based on Mixed and Augmented Prototyping practice.

Mixed Prototyping (**MP**). As shown in figure 1, MP consists of a physical prototype in a virtual environment. As already noted in section 2, the virtual environment does not refer to the scene but rather is the prototype that is "augmented" in order to give to the physical prototype the appearance and the behavior of the real product. The MR displays, in fact, superimpose to a rapid prototype the missing parts (details, additional data, the appearance of the objects surface, etc.) coming from the virtual model, augmenting the illusion that the prototype really exists and works. The MP approach offers several advantages compared to the VP, because it makes the interaction of the user with the product interface easier, involves realistically the sense of touch and enhances the possibility of ergonomic evaluations.

MP has the advantage that the user can have a tactile interaction with the prototype; therefore it is useful when there is a physical interaction between the user and the product interface. In these cases, MP allows the designer to evaluate not only visual aspects of the product, but also usability properties.

The implementation of the MP practice requires to solve at least three problems: simulating the product behaviour; managing the occlusion between the user's hands and the virtual object; tracking the movement of the user's

fingers in order to make him/her able to interact with the mixed prototype. These aspects, which will be described in the next section, allow us to define some guidelines for the realization of a MP environment. MP in general requires a rapid prototype of the parts that the user may touch during the interaction with the digital product. In some cases it is convenient to build a physical prototype only of the user interface (fig.2). Only one user at a time can interact with the prototype because MP is usually based on HMD. The choice of the rapid prototyping technique influences the possibility to simulate the tactile perception of a surface, because materials and roughness deeply changes among the techniques available on the market. Finally, with regard to the video tracking approaches (marker-based or feature-based), it is possible to implement both, but the marker-based approach is more convenient when the prototype is small, interactive and mobile (such as hand held objects). The benefits of the feature-based approach are greater if the size of the prototype is increasing (see sect. *Case Studies*).

Fig. 2 The mixed prototype of an hot mixer.

Augmented Prototyping (**AP**). This is the case of a virtual prototype superimposed to a real environment (fig. 3). Generally this practice is useful for evaluating the relations between the environment and the product. So it can be used for the aesthetic assessment of home appliances, that can be presented in their real context of use. There are no limits about the dimension of the prototypes. As for the MP, the size of the prototype can determine the best choice among marker or feature-based tracking: for small prototypes the marker-based approach is recommended; if the prototype is large and fixed into the real scene, then feature-based approach can be more effective, because it gives to the user the possibility to move in a large area without covering the environment with a set of markers. The augmented prototype can be interactive, but it is necessary to define a proper interaction technique that includes dedicated hardware: e.g.: a haptic device. Also for AP, only one user at a time can be involved in the design process because of the output device.

Fig. 3 An example of augmented prototyping.

In the following sections, we describe the hardware set up and the approach followed to implement the case studies, based on Mixed and Augmented Prototyping, described in section 6. In particular, the link between simulation software and VR development tool in order to implement

an interactive virtual prototype; the strategy for managing the occlusion of user's hand and tracking the position of the user's index finger tip; finally, in section 5, we describe the new approach for video tracking based on PTAM software.

4 Mixed Prototyping Environment

The hardware set-up used in this study (Fig. 4) is very easy to realize and it can be implemented with a very low budget. A video see-through HMD is used for augmented visualization. The HMD is composed of an iWear VR920 and a CamAr, both from Vuzix. The CamAr has a resolution of 800x600 pixels at 30Hz. The camera is used both for the video see-through function and for video tracking through the ARToolkitPlus and PTAM library. For the MP applications, a coloured glove, equipped with a micro-switch placed at the index finger tip, is used both to solve the occlusion problem and to improve the interactivity. The micro-switch provides a force-feedback reaction when the user is pressing a virtual button, so he/she can feel the sensation of touching a real button on the physical prototype. The micro-switch signal is transmitted to the PC trough a wireless connection. For more details about the rendering pipeline, see [3].
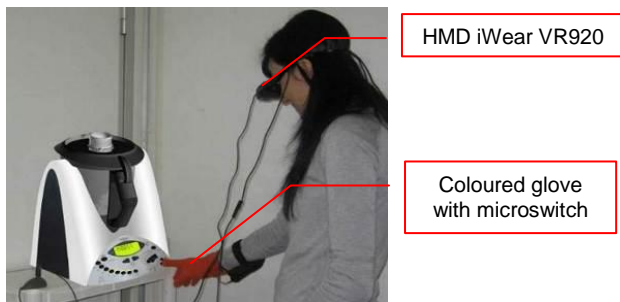

HMD iWear VR920

Coloured glove with microswitch

Fig. 4: The MP set-up.

4.1 The link between simulation and VR

In the following case studies, we have implemented interactive virtual prototypes by connecting a simulation software with a VR development tool. To do this, we have used the Matlab/Simulink environment to model and simulate the product behaviour and Virtools Dev to create the virtual environment where the user interacts with the virtual product. The link between Simulink and Virtools has been implemented through an Inter-Process Communication (IPC) channel using TCP socket. On the Simulink side we have developed two custom S-Functions. These ones can be used within a Simulink model as a conventional Simulink block, with user-defined behaviour and set of actions. In this case, our S-Functions are responsible for the communication between Simulink and Virtools through the IPC channel provided by the SimLib. One S-Function receives the events generated by the user that interacts with the virtual product and switches different simulation parameters according to these events. The other S-Function sends the data that describe the state of the product and each change that occurs during the simulation. The data sending is based on an asynchronous channel: the S-Function sends data to Virtools, without stopping the simulation. In order to obtain a consistent visualization, in fact, it is necessary to simultaneously send all data at each time-step.

The Virtools socket connection for the IPC has been implemented through two user defined Building Blocks

(BB), that are able to send and receive data through the IPC channel. The first BB sets the 3D model parameters in agreement with the data coming from Simulink. The second BB sends to Simulink the messages generated by the user interaction in the VR environment (e.g.: a button pressed, a knob rotated, etc.).

4.2 Managing the occlusion of the user's hand

The correct visualization of the user hand in the MR environment requires the recognition of the skin that has to be separated from the background. Tens of works have approached this issue by analyzing the camera images in different colour spaces. Several techniques are proposed in literature and there are some comparisons that put in evidence advantages and disadvantages of each algorithm [33-34].

We adopted the HSI (Hue, Saturation, Intensity) colour space to overcome the illumination problems like high intensity at white lights, ambient lights or surface orientation relative to the light source. The first operation of the algorithm is a transformation from RGB to HSI [35]. We have adapted the Single Gaussian model described in [36] to classify the hand region, simplifying the equation to keep the algorithm easily implementable into a shader program.

In Fig. 5a the result of the adopted algorithm is shown. It is apparent that the segmentation is not perfect, but it is acceptable for our purpose. At the moment we have preferred to focus our attention on the performance than on the accuracy of the segmentation: the shader implementation needs less than 1ms to process and display the hand-region.

Although the technique is able to detect the skin region, we have used a coloured glove during the tests in order to improve the precision of the hand recognition, thus having a more stable application (fig.5b). Moreover, the glove is needed to place the micro-switch at the finger index tip, as described in the next section.
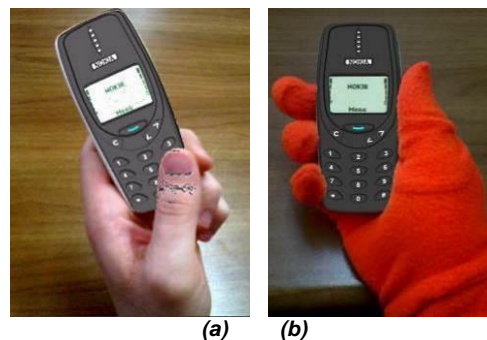

(a)   (b)
Fig. 5. Managing the occlusion. (a) skin detection; (b) with coloured glove.

4.3 Tracking the position of the user's index fingertip

As evidenced in the state of the art, the interactive mixed prototyping requires the possibility to interact with the elements of the product interface. There are three methods currently investigated in literature: a tracking device mounted on a fingertip [4]; a tracked pen [5]; haptic-simulated elements placed on the physical prototype [8].

We observed that in most cases, e.g.: with electrical appliances, when the user approaches the product interface to push a button, he/she tends to maintain the natural position of Fig. 2, with the index finger completely

extended. So we can assume that, for the cases in which the user may interact with buttons only through the index finger, we can track the position of the index finger tip in order to detect the user intention while he is interacting with the product interface.

The tracking strategy that we have adopted determines the position of the user's index finger tip in the screen space coordinate system through a blob detection algorithm. The 2D position of the index finger tip is used to determine where the user is pointing, through a classical ray casting method. When the user attempts to push a button of the virtual product, his/her index finger impacts on the physical prototype, activating the micro-switch mounted on the tip. The activation of the micro-switch triggers the ray casting in order to detect if the user is activating any button of the product.

In order to speed-up the finger tracking process, we execute the blob detection algorithm on the hand-mask produced during the hand detection step described in the previous section. Moreover, the algorithm attempts to reduce the search area to a square of 200 x 200 pixels, centred in the position of the blob in the previous frame. This considerably reduces the computational time: when the finger is found inside the square the computational time is reduced from 10 ms to 1.6 ms, thus increasing the frame rate of the application up to 60 fps.

5   Video tracking for MR applications

In order to realize a test with different video tracking strategies in MR applications, we had to integrate the video tracking functionality inside a VR environment. As already stated in section 4, we have adopted *Virtools* as the development environment for our interactive applications, so we had to extend its features by allowing the video tracking with the two approaches that we would like to evaluate: the marker-based and the feature-based. For the former case we used *Artoolkit Plus* [12], while for the latter we used PTAM [28]. *Artoolkit Plus* can be integrated in *Virtools* through the *Virtoolkit* plug-in. This is composed by some BB for the tracking based on single marker and multi-marker detection. In figure 6, the script used for a multi-marker setup is shown. The script starts with the *Video Player* BB that initializes the video capturing from the HMD camera. The *Video Capture* BB acquires one frame from the camera and the *Marker Tracking* BB initializes ARToolkit, which specifies the calibration file of the camera and a configuration file of the multi-marker setup. If a true Camera Correction is set, the BB calculates the projection matrix of the camera using the focal length of the calibration file. Alternatively, it is possible to choose the "field of view" of virtual camera in *Virtools*, acting on the parameters of the active virtual camera (fig.7). The configuration file contains the information on single marker of the setup, placed with respect to the origin of the world system. The markers, recognized by *Artoolkit Plus*, are several thousand; each marker is identified by its own ID and dimension, and by the transformation matrix.

After this initialization phase, there is a loop that captures a frame and processes it through the *ARToolkit* routines included in the *Marker Detection and Transformation* BB. This BB receives a texture that is the current frame of the live video of the HMD camera, detects the marker, and calculates the transformation matrix of the marker with respect to the camera. This matrix is assigned to the virtual object.
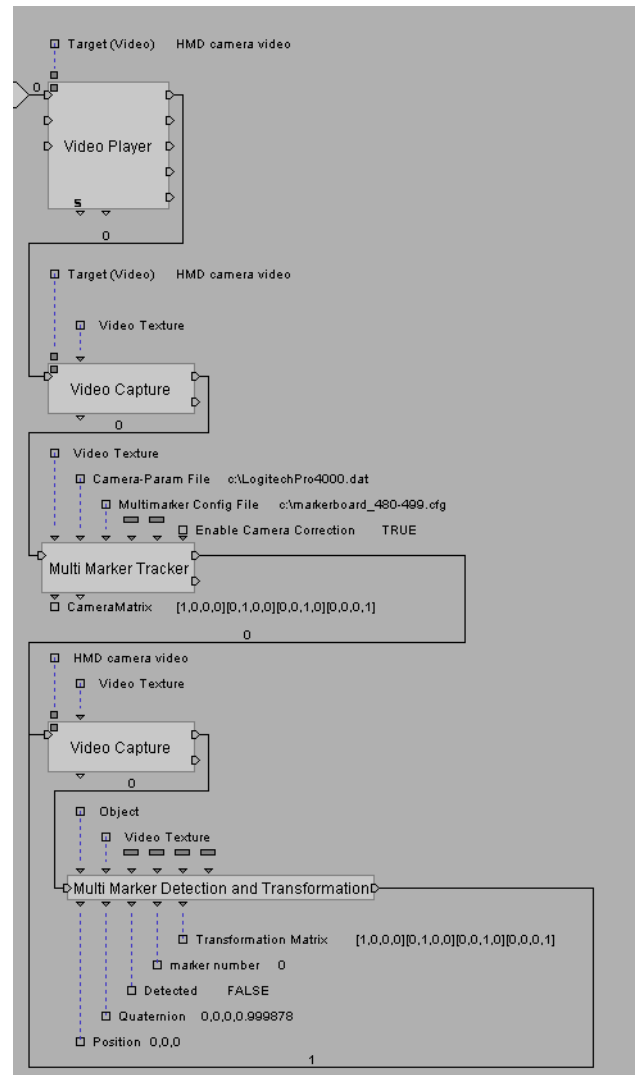

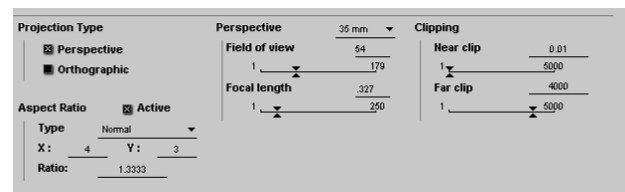
Fig. 6. The Virtools script for marker-based tracking.

Fig. 7 Virtual camera setting.

The integration of *PTAM* functionalities inside Virtools has been realized through the implementation of a single BB that receives the current frame from the HMD camera video stream and two transformation matrices that are needed for the correct orientation of the world coordinate system. The *PTAM* BB updates the world transformation matrix that is applied to the root node of the scene (Fig.8).
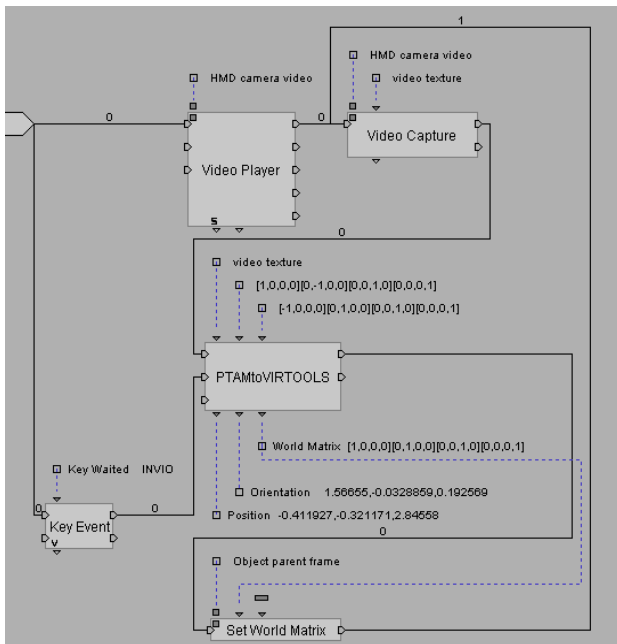
Fig. 8 The Virtools script for feature-based tracking.

The initialization of *PTAM* is done by the same BB that is triggered when it receives a specific key event. The initialization procedure employs the native interface of the *PTAM* library shown in Fig. 9. The user has to move the camera sideways between the first two key-frames: this allows *PTAM* to define a baseline between the corresponding features in the first two key-frames.

Fig. 9 PTAM initialization sequence

The *PTAM* Library is not able to retrieve the real distances among the features recognized during the tracking. Moreover, at each initialization it defines a different origin of the world reference system, because it does not have any information about the environment. This means that the reference system, defined during the initialization, cannot be predetermined and it is different for each initialization. This is a limit for MP and AP applications, because in these cases it becomes essential to fix the correct dimension and position of the digital object superimposed to the real environment. The second version of *PTAM* partially solves this problem, as it allows the user to create and save the feature map created during the stereo initialization phase. This new feature makes it possible to carry out the calibration phase just one time, but it still requires a manual calibration of the reference system.

In order to support the manual calibration, we have implemented a quick procedure that allows the user to move and scale the digital object until it reaches the right dimension and position in the real scene. The user should

place in the scene a real object, with known dimensions, that is used as a reference. In particular we adopted a chessboard that is easily recognized by *PTAM* and allows us to quickly orient the world reference system. So the user has only to translate and scale the digital object in order to obtain the correct collocation in the environment.
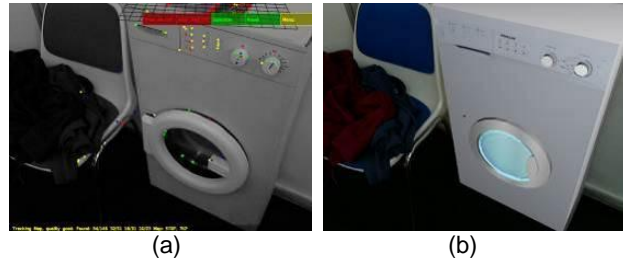

       (a)   (b)

Fig. 10 (a) PTAM tracking, (b) Object in Virtools.

6 Case Studies

In this paper, we have developed four case studies based on MP and AP practices. In particular, we have simulated a mobile phone, a hot mixer, a washing machine and a multimedia bed. Each case describes a specific application related to MP or AP. These case studies are used for a comparative study that aims to identify the video tracking system that better fits each application.

6.1 Hand held products

For hand held product, we mean a small sized interactive product . This category includes products such as mobile phones, remote controllers, mp3 players, etc. As put in evidence in section 3, for this typology of product the marker-based technique is recommended (fig.11). In fact, in this case the user always focuses his eyes on the prototype, making the solution with *ARtoolkit* particularly stable and robust (during our test, we have not found significant losses on the images).

In order to complete the prototyping, we have implemented the occlusion and the functionalities of the virtual model. But, if the prototype is very small, it is important to consider that several problems may arise in hand tracking, due to the strategy adopted for the hand detection algorithm. In these cases, the resolution of the HMD, the dimension of the keys and the proximity of the prototype to the point of view are crucial for the effectiveness of the simulation.
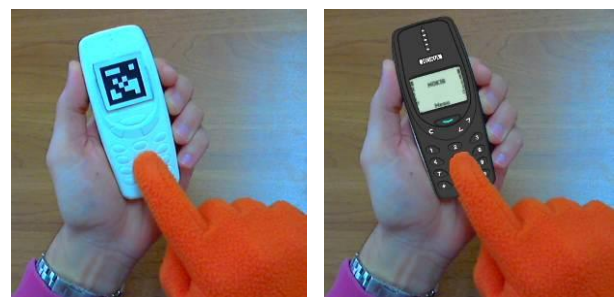

Fig. 11: Mixed Prototyping of a hand held product.

For hand held products, the PTAM solution is not usable because the tracking needs many fixed points to define the world coordinate system. If the prototype is moving, PTAM is no longer able to track the virtual model within the scene.

6.2   Small appliances

In this category, we include all the small sized products with an electronic control interface (see examples in figures 2 and 4) that are fixed in the environment. For these devices, the interface plays an important role in the overall design of the product and usually it occupies more than one third of the main side. Hand recognition and tracking do not present any particular challenge for this kind of product, because the size of the interface elements is, in general, large enough.

For small appliances, both video tracking approaches may be used. So the most appropriate technique can be chosen considering the performances that each one is able to obtain in this kind of application. For this reason we have carried out a comparative test that aims to evaluate the robustness of the two techniques under the same conditions. The test has been carried out applying the tracking algorithms on the same pre-recorded video, thus having in output a video of the resulting tracking and a measurement of the total time of tracking failure. This failure time has been calculated as the sum of the time intervals in which the tracking is not able to determine the position and the orientation of the camera with respect to the environment. The source video has been recorded taking into account that the *ARToolkit* markers have to be always visible, but most of the conditions that often cause a degradation of the tracking performances (bad light conditions, fast movements of the camera, etc.) were present.

A qualitative comparison has been done by comparing the output video generated by the two tracking techniques. The results showed that the feature-based solution is much more stable and vibration-free. Moreover, we have detected that the time of tracking failure is 4 seconds for *PTAM* and 30 seconds for *ARToolkit* for a 2' 10" video.

We have also compared the calculation time required by the two libraries to compute the position of the camera with respect to the world. Obviously *PTAM* requires a longer calculation that is approximately 3 times bigger of the one required by ARToolkit. In particular on the test case we have measured 5 ms for *ARToolkit* and 15 ms for *PTAM*.

6.3   Major Appliances

We consider in this category all those products that have an electronic interface control that occupies only a small part of the device. This feature, that differentiates this kind of product with respect to that described in section 6.2, directs the choice for the video tracking to PTAM for two reasons: it is not necessary to prepare the environment with a multi-marker configuration; it is possible to see, without problems, both the interface and the whole device. Also in this case, the HMD resolution could lead to some problems in the recognition of the texture of the interface. Moreover, the video tracking with PTAM could be problematic if the scene is completely white.

Fig. 12. Mixed Prototyping with PTAM.

6.4   Furniture

The type of objects considered in this section includes large sized products generally non-interactive. Since these products are usually not involved in usability analyses the functional simulation is not necessary and the issues related to the user-product interaction are not present. In these cases Augmented Prototyping, as widely discussed in literature, has demonstrated to be a good choice in order to evaluate the style of a product placed in a real environment. We suggest to use feature tracking in this kind of application, because it allows to contextualize the prototype in a real scene without preparing the environment with many markers. The technique is very robust and vibration-free, even when the user moves in the scene.

The test case we have analysed is a multimedia bed called Hi-Can (High Fidelity Canopy) that has several functionalities like home-video projection, lights, electric blinds, etc.. All these functions can be controlled through an iPod or an iPhone that work as a remote controller.

Fig. 13. The Hi-Can: a multimedia bed used as test-case

For this particular test-case, the user interface can be completely real since it is implementable inside the iPod/iPhone, while the virtual prototype can be visualized in AP inside a real environment. The user does not need to touch the digital prototype because all the controls are placed on the mobile device in her/his hand. The test of the *PTAM* feature tracking within this case-study has shown that it works quite well as the environment has a sufficient number of natural features. In fact we have

experienced some problems when the AP application was running in a clean room without any object or texture over the wall and the floor. In these cases *PTAM* is not able to recognize a sufficient number of features to make the tracking stable. Instead, in an environment like the one shown in figure 14, *PTAM* works very well, making possible to observe the object from any point of view.
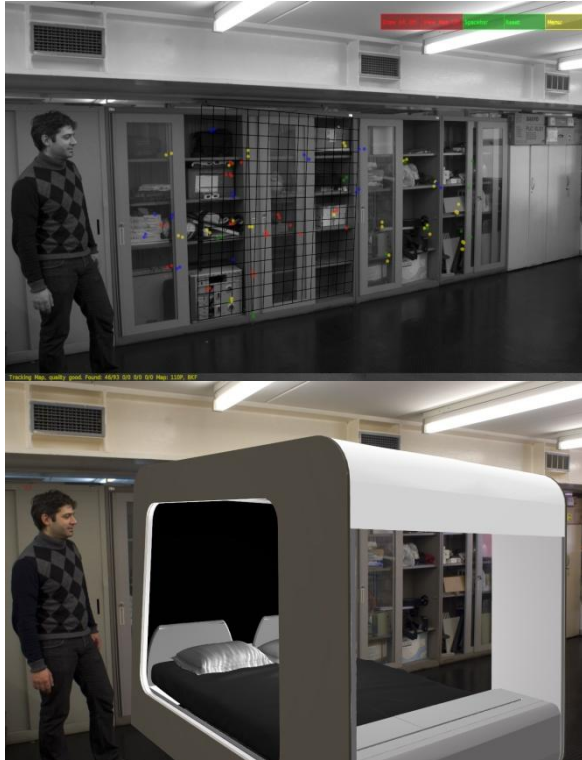

Fig. 14. Augmented Prototyping with PTAM video tracking.

7   Conclusion

This paper analyzes some of the most relevant issues in Mixed Prototyping applications, putting the emphasis on the choice of the right tracking technique. We have focused on video tracking because this solution is easy to implement and does not require any dedicated tracking device. The paper also describes how to solve other common issues related to the simulation of the product behavior, to the management of the occlusion and to gesture recognition.

Moreover, through the analyses of four test cases we suggest some guidelines for the choice of the better video tracking techniques in different situations.

In case of medium-sized static objects, i.e. objects not moving in respect with the scene, using natural features tracking instead of a multi-marker solution allows to improve stability without requiring the annoying set-up phase where each marker is carefully placed within the scene. We propose the same solution also for the case of immersive applications, where the visual presence of the markers within the observed scene can disturb and compromise user perception of the real environment, visually contaminated by the markers. We have supported this suggestion also with the numerical comparison of the performance obtained by the two tracking technique in a typical scenario of use.

Conversely, in the case of hand held objects the marker based tracking still provides a good solution, because the user can entirely observe the mixed prototype using one or few markers just fixed on the surfaces of the physical prototype.

Future works will be addressed to the development of a complete methodology that may assist engineers and designers in the use of Mixed Prototyping during the product development process. The guidelines and the considerations reported in this paper will be analytically formalized and will be enriched with other suggestions related to the realization of the mixed prototype. Particular attention will be given to the usability analyses that we consider to be one of the most promising application of AR/MR technologies in the industrial field.

References

[1]   J. Verlinden, I. Horvath. *Framework for testing and validating Interactive Augmented Prototyping as a Design Means in Industrial Practice*. Proceedings of Virtual Concept 2006, Playa Del Carmen, Mexico, November 26 – December 1, 2006.

[2]   J. Verlinden, I. Horvath, *Analyzing opportunities for using interactive augmented prototyping in design practice*. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, Vol. 23 (2009), pp. 289-303.

[3]   F. Bruno, F. Cosco, A. Angilica, M. Muzzupappa, *Mixed Prototyping For Products Usability Evaluation*. In Proceedings of the ASME 2010 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference - IDETC/CIE 2010, Montreal, Quebec, Canada, August 15-18, 2010.

[4]   H. Aoyama, Y. Kimishima, *Mixed reality system for evaluating designability and operability of information appliances*. International Journal of Interactive Design and Manufacturing, Vol.3 (2009), pp. 157–164.

[5]   H. Park, H.C. Moon, J.Y. Lee, *Tangible augmented prototyping of digital handheld products*. Computers in Industry, Vol. 60-2 (2009), pp. 114-125.

[6]   W. Lee, J. Park, *Augmented Foam: A Tangible Augmented Reality for Product Design*. Proceedings of the International Symposium on Mixed and Augmented Reality - ISMAR'05, October 5-8, 2005, Vienna, Austria.

[7]   T. Ohshima, T. Kuroki, H. Yamamoto, H. Tamura, *A Mixed Reality System with Visual and Tangible Interaction Capability*. Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality - ISMAR03, October 7-10, 2003, Tokyo, pp. 284-285.

[8]   M. Bordegoni, U. Cugini, G. Caruso, S. Polistena, *Mixed prototyping for product assessment: a reference framework*. International Journal on Interactive Design and Manufacturing, Vol. 3-3 (2009), pp.177- 187.

[9]   M. Bajura, U. Neumann, *Dynamic registration correction in video-based Augmented Reality Systems.* IEEE Computer Graphics and Applications, Vol. 15 (1995), n. 5, pp 52–60,

[10] F. Zhou, H. Been-Lirn Duh, M. Billinghurst, *Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR.* Mixed and Augmented Reality, IEEE / ACM International Symposium on, 2008, pp 193–202.

[11] H. Kato, M. Billinghurst, *Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System.* In IWAR '99: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, pp 85–94, Washington, DC, USA. IEEE Computer Society.

[12] D. Wagner, D. Schmalstieg, 2007. *Artoolkitplus for pose tracking on mobile devices.* Proceedings of 12th

Computer Vision Winter Workshop CVWW07, 2007, pp. 139-146

[13] D. Schmalstieg, *Studierstube augmented reality project: Studierstube tracker:* http://studierstube.icg.tu-graz.ac.at/handheld_ar/stbtracker.php, accessed 12 Feb 2011

[14] M. Fiala, *ARTag, a fiducial marker system using digital techniques.* In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2005, pages 590 – 596 vol. 2

[15] M. Fiala, *Comparing ARTag and ARToolkit Plus fiducial marker systems.* In Proceedings of the IEEE International Workshop on Haptic Audio Visual Environments and their Applications, 2005, pages 148 – 153.

[16] M. Fiala, *Designing Highly Reliable Fiducial Markers.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 32 (2010), no. 7, pages 1317 –1324.

[17] U. Neumann, S. You. *Natural feature tracking for augmented reality.* Multimedia, IEEE Transactions on, vol. 1, no. 1, (2002), pages 53–64.

[18] M.L. Yuan, S.K. Ong, A.Y.C. Nee. *Registration using natural features for augmented reality systems.* IEEE Transactions on Visualization and Computer Graphics, pages 569–580, 2006.

[19] T. Guan, C. Wang. *Registration based on scene recognition and natural features tracking techniques for wide-area augmented reality systems.* Multimedia, IEEE Transactions on, vol. 11, no. 8, pages 1393–1406, 2009.

[20] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, D. Schmalstieg. *Real Time Detection and Tracking for Augmented Reality on Mobile Phones.* IEEE Transactions on Visualization and Computer Graphics, 2009.

[21] L. Gruber, S. Zollman, D. Wagner, D. Schmalstieg and T. Hollerer. *Optimization of Target Objects for Natural Feature Tracking.* In 20th International Conference on Pattern Recognition, ICPR2010, pages 3607–3610, 2010.

[22] Zhuo Chen, Xinyu Li. *Markless tracking based on natural feature for Augmented Reality.* In Educational and Information Technology (ICEIT), 2010 International Conference on, volume 2, pages V2–126 –V2–129, sept. 2010.

[23] D.G. Lowe. *Distinctive image features from scale-invariant keypoints.* International journal of computer vision, vol. 60, no. 2, pages 91–110, 2004.

[24] H. Bay, T. Tuytelaars and L. Van Gool. *Surf: Speeded up robust features.* Computer Vision–ECCV 2006, pages 404–417, 2006.

[25] M. Ozuysal, P. Fua, V. Lepetit. *Fast keypoint recognition in ten lines of code.* In Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, pages 1–8. IEEE, 2007.

[26] M. Ozuysal, M. Calonder, V. Lepetit, P. Fua. *Fast keypoint recognition using random ferns.* IEEE transactions on pattern analysis and machine intelligence, vol. 32, no. 3, pages 448 –461, 2010.

[27] J. Neubert, J. Pretlove and T. Drummond. *Semi-autonomous generation of appearance-based edge models from image sequences.* In Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on, pages 79–89. IEEE, 2008.

[28] G. Klein, D. Murray. *Parallel tracking and mapping for small AR workspaces.* In Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on, pages 225–234. IEEE, 2008.

[29] R.A. Newcombe, A.J. Davison. *Live dense reconstruction with a single moving camera.* In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pages 1498–1505. IEEE, 2010.

[30] P. Milgram, F. Kishino. *A Taxonomy of Mixed Reality Visual Displays.* IEICE Transactions on Information Systems, vol. E77-D, no. 12, December 1994.

[31] F. Bruno, M. Muzzupappa, *Product interface design: A participatory approach based on virtual reality*, International Journal of Human-Computer Studies, Vol. 68, n. 5, May 2010, Pages 254-269.

[32] F. Ferrise, M. Bordegoni, J. Lizaranzu, *Product design review application based on a vision-sound-haptic interface.* Proceedings of the 5th International Conference on Haptic and Audio interaction design - HAID'10. Copenhagen, Denmark, September 16-17, 2010.

[33] P. Kakumanu, S. Makrogiannis, N. Bourbakis, *A survey of skin-color modeling and detection methods*, *Pattern Recognition*, Vol. 40-3 (2007), pp. 1106-1122.

[34] S.L. Phung, A. Bouzerdoum, D. Chai, *Skin segmentation using color pixel classification: Analysis and comparison.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27-1, (2005), pp. 148-154.

[35] R.C. Gonzales, R.E. Woods, *Digital Image Processing.* Pearson Education, 2008, pp. 410-411.

[36] J. Yang, A. Waibel, *A real-time face tracker*, Proceedings of the third IEEE Workshop on Applications of Computer Vision, WACV96.