Towards an augmented reality guiding system for assisted indoor remote vehicle navigation

Héctor Martínez¹,* , Seppo Laukkanen¹

¹SenseTrix. PL 20, 00101 Helsinki, Finland

Abstract

Scientific facilities usually require the use of vehicles controlled remotely by an operator relying on the information provided by a camera. However, the operator lacks of depth perception, which makes difficult the navigation. Providing additional information by means of Augmented Reality (AR) may enhance the safety and accuracy of the navigation. AR has been used in outdoor navigation systems. However, the technologies for outdoor systems cannot be directly ported to indoor environments and hence new techniques are required. Although several works for AR-based human indoor navigation have been proposed, to our best knowledge there are no similar studies for assisted indoor remote vehicle navigation. In this paper, a new prototype for assisted indoor remote vehicle navigation is proposed. The prototype provides the operator with an augmented view of the remote scene and complemented with additional features. The prototype has been tested in a small-scale setup with promising results.

Keywords: assisted navigation, augmented reality, indoor navigation, remote vehicle navigation.

Received on 29 October 2014, accepted on 06 December 2014, published on 25 February 2015

Copyright © 2015 Héctor Martínez and Seppo Laukkanen, licensed to ICST. This is an open access article distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/3.0/), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/inis.2.2.e3

1. Introduction

Vehicle navigation is a task that requires continuous careful attention from the driver, especially in indoor areas full of equipment. Either remotely or from inside the vehicle, the driver needs to pay attention to all elements and events in the path where the vehicle is going through. Therefore, providing additional information to enhance the navigation is always sought.

Augmented Reality (AR) is a technology that merges virtual elements with real environments in real time. The real environment is augmented with properly positioned virtual elements and the final scene is displayed on the field of view of the user. Thus, AR can be a suitable tool for providing visual content to the user in the user’s environment. This is particularly useful in those cases in which the user needs additional information or help in his or her field of view, for example, when following maintenance instructions or when driving a vehicle.

AR has been utilised as an aiding tool for navigation in cars. In [1], the authors already proposed the car windshield as a suitable device for augmentation of driving instructions. The idea is to display virtual arrows over the real view (or over a view captured by a camera if a side display is used instead of the windshield) showing the path to follow by the driver. The same concept has been utilised by other researchers in later works.

Several examples of AR guiding car navigation systems can be found in the literature (e.g. [2], [3]) and even in commercial products (e.g. [4], [5]). These examples rely either on the information of the GPS system to determine the positioning of the virtual objects (e.g. [1], [5]) or on the use of computer vision algorithms to detect visual patterns that are found in most of the roads, such as crosswalks or speed signs (e.g. [2], [4]).

AR-based guidance has been also applied to different domains apart from cars. Human navigation in unknown environments is hard without additional information. AR can overlay walking instructions to the user in order to locate the final destination. Some studies in this direction can be found in [6], [7], [8]. Remote vehicle navigation is another branch of navigation that has been enhanced by using AR systems. For example, in [9] an AR system for micro aerial vehicle navigation is proposed.
The aids offered to the user in these studies cover a wide range of possibilities, including arrows aligned with the road, bird’s eye views, HUDs, depth cues, etc. The purpose of these aiding elements is that the additional information helps to enhance the navigation at the same time that is presented in a friendly and non-disturbing manner.

However, indoor environments differ from roads in two main aspects:

- **GPS systems do not work in indoor environments, except in some specific places using High-sensitivity GPS (HSGPS).**
- **The visual patterns of the indoor facilities are not a standard and, therefore, common patterns cannot be found.**

Therefore, different solutions from the proposed for outdoor navigation are needed for indoor navigation. Several approaches have been already proposed for indoor localisation. The techniques include a vast range of technologies, such as inertial sensors [10], RFID [11], WiFi [12] or even combinations of them [13], for citing some examples.

Some studies present sparse localisation of the user and provide navigation instructions at specific points. In [14], for instance, a computer vision AR navigation system based on sparse localisation is proposed. In this case, several *info points* are distributed in the facility. When the user reaches one of these points, he or she obtains an augmented view of the facility map and when he or she starts moving, virtual arrows (which are not registered in 3D with the environment) are displayed showing the path to follow. However, the problem of these sparse approaches is that they do not react to user’s deviations from the path.

One popular approach for indoor human navigation is to use computer vision algorithms to detect traditional fiducial or any other kind of AR markers and provide augmented instructions. Several markers are located along the path in order to update the position of the user. Some examples can be found in [15], [16], [17], [18], [19], [20]. In [23], a different approach is proposed: the markers are fixed to the user’s head and external satellite cameras track the position of the user by detecting the markers.

With the advance of processing power of modern devices, new computer vision approaches based on the detection of feature points in the image have been also proposed (e.g. [21], [22]). However, these approaches present some difficulties when they are applied to open spaces or to self-similar locations and when changes in the furniture of the infrastructure are carried out, as explained in [23].

The advantage of computer vision approaches is that they are cost-efficient and scalable. The majority of these studies are, however, oriented to indoor human navigation. There are also studies related to the use of AR techniques for localisation of the vehicle targeted to autonomous navigation (e.g. [24], [25]). However, to our best knowledge, there are no studies related to assisted indoor remote vehicle navigation.

In scientific facilities, the work needs to be frequently done remotely. In some areas inside these facilities, radiation is emitted and, therefore, the presence of human workers is not advisable. In some cases, autonomous vehicles can be used in these facilities. However, in other cases, due to the high cost of the equipment and singularities of the facilities, human control is needed for vehicle navigation and autonomous navigation is not possible. Therefore, it is common to find movable vehicles controlled by remote operators. In these cases, operators need to rely on the visual information provided by a camera mounted on the vehicle. The understanding of the environment is not always easy due to the lack of depth and position information.

Providing an enhanced view to the operator may potentially increase the performance of the navigation and reduce possible risks and collisions. As seen in previous paragraphs, AR is a technology capable of providing this additional information and helping the remote operator to understand the three dimensions of the environment with an augmented view. For these reasons, this paper proposes the first prototype of a new AR guiding system for indoor remote vehicle navigation.

### 2. Materials and methods

The proposed prototype aims to be the first step towards a final AR-based application for vehicle guidance in corridors and tunnels in scientific facilities. The vehicle contains a mounted camera which provides a video feed that is augmented to the remote operator in order to provide him or her with guiding hints and warnings.

The localisation of the vehicle in the corridor has been carried out by fiducial marker recognition. Once the position of the vehicle is known, the augmentation takes place. A virtual representation of a top view of the scene is also provided to the user. The distances between the vehicle and the walls are additionally calculated and used to warn the remote operator if required.

The main techniques for the final application have been developed and are presented in this paper. For the deployment of the first prototype, a small scale setup has been built. The following sections explain the details of the implementation.

#### 2.1. Multimarker configuration

Marker based AR applications rely on the detection of markers visible in the image. When the application detects one of these markers, it calculates the position and orientation of the marker respect to the camera in order to properly display the virtual objects. One common problem in these kinds of systems is that the marker can be partially or totally occluded. In those cases, the system
is not able to recognize the environment and the virtual content cannot be displayed.

In order to cope with this problem, several researchers have implemented multimarker configurations consisting on several markers printed in the same plane with distances that are known beforehand (e.g. [26], [27]). The application detects all markers that are visible in the image and provides the information for the augmentation. The major advantage of this setup is that it allows the applications to keep the position of virtual objects if one marker is hidden, as long as some markers of the configuration are visible.

Marker based navigation guidance inside a corridor or a tunnel requires a large number of markers, as markers in the walls or in the floor disappear when the vehicle is moving. Also obstacles in the corridor may occlude some markers in some positions of the vehicles. Finally, some markers may be not detected due to poor light conditions. Therefore, a large number of non-coplanar markers is needed along the corridor or tunnel. Due to this reason, the proposed prototype contains a multimarker implementation in order to provide a more reliable solution.

The detection and tracking of the individual markers in this prototype is based on ARToolKit [28], a well-known library for AR. The approach followed for the multimarker detection is similar to the one presented in [26], but adapted to enable non-coplanar markers. To achieve this, when calculating the transformation matrix between markers using the known geometric relations in real world, our implemented algorithm takes into account not only the real distances between markers, but also the real rotations. Therefore, the transformation matrix $T_{i,k}$ can be represented as shown in (1), where $i$ is the $i$th marker, $k$ is the reference marker, $R$ is the rotation matrix between markers and $d$ is the distance vector between the centers of the markers.

$$T_{i,k} = \begin{bmatrix} R & d \\ 0 & 1 \end{bmatrix}$$ (1)

The relationships between markers and the corresponding matrices are displayed in Figure 1. The multimarker implementation proposed in this paper calculates the transformation matrices ($T_i$) of all markers visible in the image and uses them, together with the transformation matrices between markers ($T_{ik}$), to calculate the rectified transformation matrix of the reference marker, even if it is not visible in the image. Furthermore, the 3D model of the whole corridor is associated to the reference marker of the multimarker configuration. With this approach, the augmentation will be consistent along the whole corridor even if the reference marker is not seen during the navigation.

**Figure 1.** Relationships between markers and corresponding matrices in a typical corridor setup containing two walls and one obstacle. $T_i$ is the transformation matrix of every individual marker respect to the camera (calculated by the marker detection system) and $T_{i,k}$ is the transformation matrix between marker $i$ and the reference marker $k$. Both types of matrices are used to calculate the final rectified transformation matrix of the reference marker respect to the camera and, subsequently, the rectified transformation matrices for the other markers.

Compared to traditional single marker approaches, the impact of the multimarker calculations on the framerate should be minimal. The detection of the markers is exactly the same in both approaches and only few mathematical calculations are added in the proposed approach. Therefore, the proposed multimarker configuration can be used seamlessly in real time.

### 2.2. Augmented information

The scene displayed to the maintenance operator consists of an augmented view of the corridor where virtual objects are displayed over real objects. For the first prototype, the 3D models of walls and obstacles have been used for the augmentation. Although the display of these models does not provide additional useful information to the operator, they have been used to analyse the feasibility of the system.

Apart from the 3D models of walls and obstacles, two additional elements have been added to the final view: a collision warning system and a virtual reconstruction. The details of these two elements are explained in the following sections.

### 2.3. Collision warning system

The collision warning system provides the operator with continuous information of the distance to the walls of the corridor or tunnel. This information is measured based on the information obtained from the markers and the calculated position of the vehicle and is displayed in the final view on a HUD. If the vehicle moves close to a wall,
the system detects the event and warns the operator with a message, also displayed on the HUD.

For the calculation of the distances to the walls in the first prototype, the distance between the camera and the marker plane has been considered. For each wall, the transformation matrix is used to calculate the plane of the marker (2), which is also the plane of the wall. The distance is finally calculated as the distance between the camera (i.e. origin of coordinates, p in Figure 2) and the plane. The distance is calculated using the equation (3). A representation of the values defined in the calculations is shown in Figure 2.

\[ a x + b y + c z + d = 0 \]  
\[ D = \frac{|a x_0 + b y_0 + c z_0 + d|}{\sqrt{a^2 + b^2 + c^2}} \]  

Figure 2. Calculation of the distance between the vehicle and the wall. The equation of the wall plane is defined by the plane with normal vector \( n(a,b,c) \). This plane can be calculated from the information provided by the transformation matrix of the marker (wall and marker are coplanar). The point \( p(x_0,y_0,z_0) \) is the centre of the camera, which is mounted on the vehicle. The distance between the car and the wall (D) is defined by the minimal distance between the point p and the plane defined by n.

The distance to the wall is continuously displayed to the operator. If the distance to the wall falls below a secure threshold, a warning message is additionally displayed.

2.4. Virtual reconstruction

The final goal of this prototype is to aid the guidance of vehicles in a corridor or a tunnel of a scientific facility. As the geometry of the facility is known beforehand, it is possible to design 3D models of the whole corridor or tunnel and use them for creating an augmented experience, as it has been already mentioned. These models are aligned to the reference marker of the multimarker configuration and they are properly displayed respect to the camera which, in the end, is mounted in the vehicle. This means that the position of the vehicle (i.e. the position of the camera) respect to the corridor or tunnel objects is always known once the markers are detected.

Taking the previous considerations into account, it is possible to create, in real time, a whole virtual reconstruction of the situation in the corridor or tunnel. The virtual reconstruction of the situation seen from the point of view of the camera does not provide additional information compared to the augmented view, as both views are, in essence, the same view. However, if the point of view of the virtual reconstruction is modified, the information provided can potentially be valuable for the operator. Adding a view of the scene (i.e. the virtual reconstruction) from the top of the corridor or tunnel can provide a whole overview of the current situation and aid the operator in describing the correct path that could be more difficult with the only help of the augmented view.

An example of a top view of a corridor using the virtual reconstruction can be seen in Figure 3. The concept presented here is similar to GPS systems. However, as GPS systems do not work in indoor environments, the positioning of the vehicle with respect to the corridor or tunnel is calculated from the known geometry and the information obtained from the detection of the multimarker configuration.

Figure 3. Example of a top view of the virtual reconstruction of a real scene with a vehicle navigating through a corridor. The virtual reconstruction is a totally virtual view of the real scene generated automatically using the information obtained by the multimarker configuration. The vehicle is represented by a white crossed box. The virtual corridor walls and obstacles are also displayed in the view. The vehicle is always displayed in the centre of the view in a similar disposition as in traditional GPS systems. The virtual reconstruction is constantly updated with the position of the vehicle respect to the corridor and obstacles obtained by the multimarker detection system.

For the virtual reconstruction proposed in this prototype, the same virtual scene used for the augmented view has been used to recreate the top view of the corridor. A crossed box has been drawn as a representation of the
vehicle. The coordinates of the vehicle respect to the walls are continuously updated with the information obtained from the multimarker configuration, keeping always the vehicle in the centre of the virtual reconstruction view. The virtual reconstruction view has been integrated as a HUD inside the augmented view in order to complement the second one.

2.5. Experimental setup

The experimental setup for the first prototype has been designed as a desk-size scale environment. Although the final deployment should be done in a real scale environment, the setup utilised for the first prototype serves as a testbed to prove the feasibility of the system and to analyse the main problems encountered.

The vehicle used for the navigation is a Radio Control (RC) car. Due to the high speed of the RC car, the car has been driven by hand instead of using the controller for the tests. The speed in the real facilities will be generally slow in the movements and never as fast as the RC car.

A webcam attached to a standard computer running the AR application has been mounted in the car. In the final implementation, the camera would be accessed remotely from the operator’s computer, using the same communications that are currently being used for the remote navigation.

For the simulation of the corridor, a block-game has been utilised. The blocks have been used to build a two wall corridor with two obstacles. Four markers have been attached to strategic places in the corridor (one marker for each obstacle and one marker for each wall). The goal of these strategic places is to ensure that there is always at least one marker visible in the image. In the real deployment, a larger number of markers would be needed in order to cover the longer distance or a real corridor. Figure 4 shows the experimental setup.

3. Results

This paper presents an on-going research work and therefore, final results cannot be formally presented. However, some preliminary results can be already explained at this stage. The problems encountered during the development are also explained in this section.

Figure 5 shows an example of a typical view of the prototype. The main view is the augmented view where the virtual corridor elements (i.e. walls and obstacles) are displayed over the real ones. In the bottom right part of the figure, a top view of the virtual reconstruction of the scene is displayed. The distance to the wall is also displayed as text on the bottom left part of the HUD. In this particular example, a warning message explaining the risk of collision is also displayed as a result of the close proximity of the car to the wall.

![Figure 4](image-url)  
**Figure 4.** Image of the experimental setup. The camera is mounted on a hand-controlled RC car navigating through a corridor with obstacles built with blocks. The camera is attached to the computer running the application and the RC car is moved manually for the tests. The markers are placed on the corridors and obstacles.

![Figure 5](image-url)  
**Figure 5.** Image of the final view. In the image, the augmented view is displayed in the main window, showing virtual walls and obstacles over the real ones. The virtual reconstruction (a top view of the scene) is displayed as a HUD in the bottom right part of the window. Text messages including the distance to the wall (bottom left) and a warning message explaining the risk of collision (top) are also displayed on the HUD.

The detection of the multimarker setup has proven to work as expected. The 3D models have been always displayed in the augmented view (and therefore in the virtual reconstruction) regardless the number of markers visible, as long as at least one marker is visible.

As expected, the multimarker calculations have not affected significantly to the framerate of the application, being able to provide a stable framerate of 30-35 fps with an image resolution of 800x600. The hardware used for the tests is a standard computer (AMD Turion II P560 Dual-Core 2.5GHz, 6GB RAM) with a standard webcam (Logitech HD Pro Webcam C910).

Due to the movement of the vehicle, it may appear a case where all markers are outside the field of view of the camera which would be translated into a loss of the vehicle position respect to the corridor. Although the position would be recovered as soon as one marker is seen...
again in the image, the sudden change in the augmented view and in the virtual reconstruction may confuse the operator. The integration of a system that analyses the movement of the camera in those cases where no markers are visible in the image could be a possible solution. The camera motion could be used to determine the movement of the vehicle since the last marker was seen and use that information to properly display the augmented view and the virtual reconstruction.

The virtual reconstruction of the corridor has demonstrated to be a helpful tool as it increases the understanding of the situation by the operator. It is especially useful to understand the depth of the vehicle navigation, as it is usually more difficult to perceive depth distances from a frontal view.

The main problem of the virtual reconstruction is that it totally relies on the detection carried out by the multimarker setup. Therefore, if all markers are outside the field of view of the camera, the virtual reconstruction view appears empty. This is a problem that would be solved if the aforementioned camera motion system is integrated in the system.

4. Discussion

This section introduces the conclusions obtained after the development and testing of the first prototype and the future work.

4.1. Conclusions

In this paper, a new prototype of an AR guiding system for indoor remote vehicle navigation has been proposed. Although previous research studies have dealt with outdoor navigation and human indoor navigation, to our best knowledge this is the first study for AR-based guiding remote vehicle navigation system.

The developed prototype provides the remote user with an augmented view of the scene in which the vehicle is moving. The augmented view is complemented with a top view virtual reconstruction of the corridor and vehicle and with text information and warnings in a HUD.

The proposed approach provides the following advantages for assisted indoor remote vehicle navigation:

- Visual clues and warning messages directly displayed over the real environment in forms of augmented content and/or screen-aligned text.
- Improved depth perception of the environment by the operator. The virtual reconstruction provides a point of view (from top) that enhances the understanding of the current situation.
- Automatic detection of potential collisions based on distance calculations from the vehicle to the walls and obstacles.

The prototype has been developed and tested in a small-scale scenario to analyse the feasibility and to detect the potential problems of a system of this kind. The results obtained during the tests show a great potential to this concept and further analysis are expected to provide more accurate results.

The deployment of the system in a real-scale scenario should be feasible without major problems. The camera would be mounted on the real vehicle and the markers should be scaled according to the real size of the corridor or tunnel. Instead of the current direct connection of the camera to the computer running the application, the existing communication system used for the remote navigation would be also used for the transmission of the video feed of the camera mounted on the vehicle to the computer used by the remote operator. This computer would be the responsible for running the proposed application and displaying the final view to the operator.

With this kind of setup, the performance of the navigation would be increased in terms of time and risk. A better understanding of the corridor or tunnel helps the operator to perform the navigation in a faster and safer way. Furthermore, the values of distances to objects and the warning messages are a second aid to prevent potential collisions and damages, reducing risks and increasing, thus, the safety.

A great advantage of the virtual reconstruction is the high accuracy of the positioning system, as it can be as accurate as the tracking system is (i.e. the multimarker setup).

However, the operator cannot rely only in the virtual reconstruction, the same as a driver cannot rely only on a GPS system. Therefore, a combination of the augmented view and the virtual reconstruction is advisable to ensure a better performance.

The study presented in this paper aims to be the first steps to a guiding system suited for assisted indoor remote vehicle navigation. The paper proposes a new multimarker approach to provide a more robust AR system compared to individual markers and that allows non-coplanar markers to belong to the same multimarker configuration. The system also proposes a battery of visual helps to enhance the quality of the navigation such as variable measurements (e.g. distances to objects) and warning messages. Finally, the system implements a virtual reconstruction of the scene calculated from the multimarker detection system information, offering a second point of view to the operator, increasing thus the understanding of the environment.

The work developed in this prototype has been integrated into a larger platform oriented to AR maintenance applications for large scale scientific facilities. The details of the mentioned platform can be found in [29].

4.2. Industrial applications

Industrial and scientific facilities usually contain certain areas with extreme conditions (e.g. radiation, extreme temperatures, etc.). In such conditions, it is advisable to
substitute human work with other means of remote intervention.

Many times, this work is remotely carried out by the use of robots and vehicles. Nowadays, the number of autonomous robots and vehicles in industrial and scientific facilities is large. However, there are still some cases in which their use is not possible due to the facility restrictions. In these cases, the robot or vehicle needs to be controlled remotely by an operator.

The proposed system tries to help operators in the navigation of the robots and vehicles in these circumstances. Although the proposed scenario focuses on tunnels and corridors, the proposed system could be applied to any kind of industrial setup in which assisted remote vehicle navigation is required.

The proposed approach leaves an open door to integrate the currently available sensors of the facility into the system. The information retrieved by the sensors could be used to offer additional values and warnings to the operator in the augmented view.

The use of these kinds of systems would help industry to improve the functioning of the facilities in terms of faster interventions and less risks which would also reduce costs and optimize productivity as the different works would be performed more efficiently.

4.3. Future work

The basic prototype has already been developed and the next step in the workflow should be a first implementation in a real scientific facility (for example, using the facilities of our partners in this project – CERN and GSI – as testbed) in order to analyse the size and suitable disposition of the markers in a real environment.

Once the prototype is built in the real scenario, user tests are needed to analyse the performance of the system in real conditions and to prove its usefulness as an aiding tool for assisted remote vehicle navigation.

Although at this stage of the prototype the augmented view of the corridor has included only the 3D models of the walls and obstacles, which do not provide additional information to the operator, further virtual objects and animations can be included in a second prototype to provide new means of help for the operator.

For example, flashing objects in the corridor that are too close to the vehicle may warn the operator that he or she is too close to the object and that a rectification is needed. The flashing objects can be also displayed in the virtual reconstruction in order to avoid collisions in rear area.

Moreover, these visual warnings can be complemented with voice warnings. Also, if a path planner is included into the system, virtual 3D arrows displaying the correct path can be also used in the augmented view.

Acknowledgements.

This work was carried out in project “Preventing hUman intervention for incrReased SAfety in infrastructures Emitting ionizing radiation (PURESAFE)” PTN-GA-2010-264336-PURESAFE, funded by the 7th Framework Program of the European Union under the Marie Curie Actions - Initial Training Networks (ITNs).

References


