Augmented Reality-Assisted 3D Visualization for Urban Professional Users

Christine Perey, PEREY Research & Consulting and
Graziano Terenzi, ARmedia

Research & Consulting

Abstract

This white paper explores how Augmented Reality-assisted 3D visualization can be used to cost-effectively address diverse needs of professionals performing routine and emergency tasks in urban environments. The current status and future scenarios of three simple use cases are compared with AR-assisted systems deployed with existing (2013) enabling technology.

We describe key technology enablers that will be introduced commercially over the next 12 to 24 months to improve performance of systems and strengthen the business case for AR use.

We suggest different factors that managers of professional city services can consider when estimating the Return on Investment (ROI) from introduction and use of AR with their IT systems. Finally, we recommend that different stakeholders of the municipal IT systems converge toward a shared vision and make a plan to achieve their goals.

The following audiences are considered most likely to benefit from this paper:
- GIS managers and 3D information managers working for municipalities and similar agencies or businesses serving data to professional users, and
- Experts in the fields of energy, utility management, water management, environmental management, safety and emergency response, architecture, construction and engineering, among others who would be included in a coalition or task force for AR introduction and deployment.

Table of Contents

Professional Uses for Urban AR

Digital information, structured data and especially that which uses 3D models, has many uses in a city or urban environment. With this section we briefly introduce only a few potentially valuable use cases in order to illustrate how visualization of data in "camera view," overlaid on the buildings and other objects in a user's proximity can provide benefits.

Call before digging

Most urban environments have infrastructure on the ground, where people can see it, and below ground, out of sight. The underground infrastructure is maintained by organizations designated to this task by the municipality.

When any project requires digging, implements could damage buried cables, pipes or devices. For this reason it is frequently against the law to dig without first obtaining clear and accurate information about the position of the utilities.

Local, regional and national agencies are in place to provide up to date information to professionals as well as to citizens who plan to dig. For example, in America, anyone can call a local "One Call Center," or dial 811 to reach the relevant One Call Center. Within a few days of providing an address, a professional will mark the ground with paint or flags showing the position of electricity, clean and waste water, gas, telecommunications and other services. The same or similar service is provided in virtually all well-managed urban centers.

It is highly likely that in the future One Call service professionals will have a tablet-based service that permits, as if they had "X-ray vision," to see the position of all underground systems when marking an area. The system might first be deployed on a dedicated device that can be provided and supported by a city service department or the company managing services. On this device the city can ensure that the required information is always accurate/up to date and accessible immediately to the application, regardless of the mobile data network access. Having data locally will also reduce power consumption in the field. Using a geospatial data synchronization protocol such as OGC GeoPackage, the city's database can, upon connecting with the mobile device, be updated, indicating which data sets were accessed, by whom and when.

Figure 1. Utilities under road. Image courtesy of Augmented Technologies Ltd.

Once positioning technologies are sufficiently precise and in areas where the position of utilities pipes are well known, such services may become available to lay people and non-professional users.

Building navigation for rescue/evacuation

Emergency responders are often in the line of danger. When in transit to and planning to enter a building that is damaged or in flames, they must carefully study paper maps showing where there are dangerous materials, emergency exits and other relevant landmarks. Frequently the maps are not available in digital formats (e.g., simplified floorplans appear at the entrance but may be inaccurate). Digital floorplans may be too costly or difficult to maintain up-to-date as building space usage changes. The lack of access to building floorplans when moving where vision is impaired (smoke filled rooms) is another hazard.

In order to reduce the risk to human life and increase the speed with which evacuations can be carried out safely, emergency responders are or will soon be testing devices that help them to see the most current digital floorplan on a mobile device in which an embedded accelerometer and compass can detect the user's orientation.

Unfortunately, indoor navigation with only the support of an Inertial Measurement Unit (IMU) is immature and not sufficiently precise. Furthermore, the systems based on commercial (mass market) platforms do not have barometers and other indicators of elevation, so the device is unable to automatically choose the floor plan for the correct floor. Finally, holding the mobile device for rapid consultation reduces the ability of the emergency responder to use both hands for other tasks.

Figure 2. Emergency responders will use AR to find and follow the safest route through a building. Safety management operators will be prompted to manage a risk situation in real time (Image courtesy of Hopital La Magdalena and the ANGELS Consortium)

In the future, other relevant information (e.g., temperature, water pressure in pipes, air quality in rooms) will be graphically superimposed on the real world in real time when in the building regardless of the conditions and in perfect alignment with the dangers or exits. Hands-free displays with AR support, as are already in use for military applications, will permit people to perform their duties more quickly and safely.

Urban design and planning

The urban landscape is never static. There are always changes being planned or built. Whether in densely built areas or where structures are sparse, planners are continually assessing what changes will be needed to meet the needs of businesses and citizens. People working in the Architecture, Construction and Engineering disciplines have a rich array of information technologies and tools. Some help with the design of future structures or modifications of existing buildings. Many are mission critical, permitting the precise measurement of a site, the assessment of conditions and even safety of workers.

In this very brief paper, it is not possible to include all the possible uses for AR in all the steps of the life cycle of a building. We will only focus on design and planning.

Today, when a designer communicates a plan to other stakeholders, it is done with several different tools such as a 3D digital model viewed in a computer software program or a movie (fly through), a scale model built in lightweight materials and set up in a manner than can be seen from all angles in a "birds eye view" and on-site there may also be a temporary physical framework.

In these different environments (e.g., virtual, physical off site, physical on location), the architect, customers, citizens and any other stakeholders can provide feedback verbally or through voting or other mechanisms. Unfortunately, even the best of these tools are limiting. Models are frequently insufficient to fully grasp the height of the proposed change or new building.

Figure 3. Novartis campus and the river Rhine. Image Source: Planning Department of Basel-Stadt and LifeClipper2 project.

By using Augmented Reality in the planning stages, community members can visualize the proposed changes while on site and make their opinions known to urban officials. For indoor applications, simulations are often insufficient. Rendered visualizations can also quickly become costly for accurately assessing various elements of the design such as if a hallway will be too dark or a room too small for the client requirements.

The future occupant of a building, or the occupant of an existing building wants to be able to look out from the window of his living room and see the view as it will look in the future. With Augmented Reality-assisted urban design and planning systems, a planner's proposal can appear as if in full size on the display. A synchronized audio file can explain the designer's goals, project stages, the meaning of some architectural choices and other relevant information, such as how choices are constrained by the regulations of that zone of the city.

State of the Art of Urban AR in 2013

During preparation of this paper, we have developed an inventory of dozens of mobile AR visualization projects involving 3D models and illustrating the opportunities for AR to bring benefit to end users and to those people and businesses who they serve in a professional capacity.

Although the urban AR projects in the inventory are highly diverse and many use the latest research results, all have shortcomings that prevent them from being widely implemented today by professionals in the field. The barriers are organized in five domains: technological limitations, obstacles stemming from the digital assets (data), computer-human interface design factors, policy and legal issues, and, finally, financial constraints.

Technology Limitations

Augmented Reality relies on a variety of inputs from sensors and, to produce experiences there must be displays for information overlay. In this section we will identify the most important technological barriers to widespread use of AR by urban professional users in 2013.

First, the sensor and computational tasks for AR in most applications require continuous use of power so the battery life of a smartphone or tablet using AR can be as short as 2 hours of uninterrupted use. This short battery life means that the user with more than a few hours of work will have to carry a backup battery or make other changes to workflow.

Technology	Limitations
Mobile Devices	- Battery life - Compass errors - Sensor drift - Screen readability in sun light
GPS	- Accuracy limited to 5-30m - Signal blocking / reflection - Weather effects - Satellite fix
Markers	- Sensitivity to light changes - Limited tracking range - Requires environment engineering
2D NFT	- Limited tracking range - Slower than marker tracking - Sensitivity to light changes (more robust than Markers)

Table 1. Technological Limitations of Existing Systems for use by Urban Professionals for AR.

In addition, the sensors for absolute and relative position, orientation and movement are highly variable in terms of their accuracy for positioning the user's focus of attention and are not considered sufficiently reliable for professional uses described in many of the use cases above. For example, the compass (used for orientation) is highly sensitive to interference from magnetic bodies such as cars, elevator shafts, bodies of water, or high concentrations of motors used in electronics.

When outdoors the ambient illumination is frequently so high that the screen is dim by comparison. In order to use 3D models in AR view, the professional will need to turn on back lighting, which will consume even more battery, or find lower light (illumination) conditions.

Mobile AR applications for outdoor applications largely rely on the smartphone GPS. Global Positioning System (GPS) provides the user position based on triangulation of signals captured from at least 3 or 4 visible satellites by a GPS receiver. Standard GPS systems have 5m to 30m accuracy.

While they have the advantages of being ubiquitous and easy to integrate in AR systems, consumer-grade GPS technology also suffers from difficulties such as:

Figure 4. Sensor readings must be backlit to be seen in context with the surrounding environment. Image Courtesy of Inglobe Technologies.

- Being unavailable (or slow in obtaining position) when satellite signals are absent such as underground, indoors, in urban canyons, and when meteorological conditions block transmission, and
- Satellites can provide erroneous information about their own position.

Once the focus of attention is identified, an AR system must also track the user environment. Some solutions we have documented in our inventory use marker-based or 2D natural feature recognition and tracking techniques to identify and track patterns in the environment. The markers appear, for example near doorways during the construction phase of a building in order for the application to retrieve information for the correct space.

Marker-based approaches are highly sensitive to lighting changes, and the tracking range is limited. They can also be removed, switched or occluded. Two Dimensional Natural Feature Tracking (2DNFT without marker) approaches are more robust in certain respects (lighting, water, etc.) but they are slower than marker tracking and are still not appropriate in many important scenarios.

Finally, AR-assisted services based on 2D NFT do not reliably recognize deforming shapes (most plants and animals, fabrics, water, thin paper), or highly reflective surfaces. They are also not appropriate for reliably tracking recognized objects when there are also moving objects (e.g., machinery, humans) in a scene.

Data Preparation, Access and Use

Without digital data to associate with the physical world, there cannot be AR experiences. There are obstacles associated with the preparation of 3D models (and other data) for AR use, the storage and transmission (accessibility) of data and, finally, many problems having to do with placement of digital information within a 3D scene by AR-assisted systems.

Phase	Constraints

Data Preparation	- 3D models are
	    o expensive to acquire
	    o diverse (can be of different resolutions, formats, etc)
	    o frequently too complex and large in size
	- Systems for AR may not support native 3D formats

Storage and Transmission	- Enterprise Content Management Systems lack AR support - AR CMS do not interface with Enterprise CMS - Transmission may not be permitted without security precautions - File transfer slow and costly, and subject to interruption
Producing a realistic 3D scene	- Occlusion - Lighting and Shadows

Table 2. Data Preparation, Storage, Transmission and Use Considerations

Some application scenarios re-use existing data sets while others require creation of new, bespoke 3D models. Development of data sets, especially with 3D models, can take days and sometimes weeks but are used in multiple ways (not only for AR visualization). Use of existing data created by professional software users (e.g., designers of industrial machines, architects), generated by 3D scanners (e.g., LiDAR), and/or collected and maintained by city departments or private businesses are going to reduce the time and cost of data preparation for an AR-assisted project.

There are, however, special considerations. In order to reduce delay associated with large files on mobile networks, and issues associated with mobile device and memory management and buffer, some vendors recommend baking the lighting conditions into model textures. This greatly reduces model flexibility.

In addition to these general data preparation guidelines, each use case presents unique challenges. For example, before the data from utilities can be used with mobile AR routinely to reduce risk of damage to underground infrastructure (aka Call Before Digging) there needs to be comprehensive data validation. In the past, before there were accurate measurement tools, many pipes and other infrastructure were buried without accurately measuring their depths. Where underground pipes cross/intersect the depth of one or both changes.

For systems that are used by emergency responders in planning or executing an evacuation, it is essential that any changes in room accessibility (e.g., if a door is blocked) be reflected in the digital building maps. In many areas there are multiple sources of data available. Should two or more geospatially-referenced data sets be part of the experience objectives, precautions need to be taken to ensure that they have consistent data structure. Choosing which data to use for a particular objective, validating the source of data, the accuracy of the data and making corrections where possible, through the client application writing in database fields, are all beyond the capabilities of current systems.

Today's AR-assisted systems frequently involve deployment of software on a consumer/mass market mobile device (aka smartphone or tablet) and cloud-based platforms with which the mobile device synchronizes. This requires that the device be continually connected to the network for the assets to be queried and sent to the device when and as necessary. In most urban environments this is not a barrier, however, secure connections between device and service are another step in the architecture that current providers have not implemented (or not widely).

When digital assets, such as 3D models are integrated into a scene involving real world assets, additional complexities emerge. Full realism in the composed scene—having the digital asset and physical world appear seamless—is not always required, however, it is frequently recommended. Realism can be diminished if the physical world obstacles (a stationary landmark or the passage of a person) fail to produce occlusion in the digital asset placed further than the obstacle. Lighting is another element of realism that is dynamic, depending on the latitude, time of year and time of day. Lighting also interacts with textures and produces shadows that, when missing, break the realistic effect of AR.

In some use cases the assets are below the surface of the land or behind a wall. A realistic projection of infrastructure below street level (pipes, parking garages, etc.), involves introducing additional elements that are not part of the scene but provide depth cues for the user.

Human Factors

If a tourist holds up their smartphone at eye-level in order to view points of interest in proximity, they may do this while standing still and rotating for 15 or 20 seconds. This process can be repeated briefly over the course of a tour in a cultural heritage site, or when looking for a restaurant. Maintaining an eye-level position of the display is not acceptable for all use cases and cannot be expected for more than brief intervals.

Urban professionals can do their tasks better or more quickly with the assistance of mobile AR-assisted technology, however, many also need their hands free and for the information to be available on a continuous basis.

Devices such as the Google Glass head-worn display suggest that there will be alternative display technology available, however, the solutions designed for mass market may not be appropriate for the professional user.

Professional users will need to have the ability to see the AR in camera view continuously and also intermittently, depending on the stage of a workflow. Eyewear will need to be securely attached so that it does not drop from the user's head when moving; it will need to be lightweight for use over extended periods of time; it will need to be made with materials that do not shatter upon impact and have other safety attributes prior to being introduced widely in certain professional use scenarios.

Policy and Legal Issues

It is very early in the introduction of AR into systems that will help public servants or professionals perform services better, however, we anticipate that prior to becoming fully integrated there will be policy and legal issues to address.

For risk reduction with technology, new professional information management (IT) systems must be certified to perform reliably and accurately, alone and in combination with existing and future IT components. The certification process may be lengthy and costly, therefore, prior to introducing new technologies into mission critical IT systems, their use must be proven to policy makers and, in some cases, also approved by lawmakers.

When AR-assisted systems are in use in public environments and must comply with existing legislation, there must be precautions taken to protect the privacy of citizens. At this time, AR platforms are not integrated with any facial feature removal technology.

The situation is different, but there are also policy and legal concerns for the use of AR in non-public workplaces. Workplace health and safety policies will need to be changed. For example, on a construction site, there are risks with machinery, moving objects and unsecured platforms that could face the user of AR-assisted systems that obscure vision or auditory signals.

Financial constraints

We have already highlighted that there are costs for developing and integrating AR into professional user workflow and data sets. Just overcoming the technology barriers alone will require large investments.

In addition to the costs of research, development and productization, the enterprises and agencies that introduce AR into their workflows will also need to purchase new hardware, train their personnel on the safe and reliable use of the new technology, contract with service providers to maintain the systems in working order, and other operational expenses. Investments will need to be made based on the calculation of a Return on Investment (ROI) on the capital.

Key Technology Enablers

In order to successfully address the requirements of the users in the scenarios described above, many companies are working on new technologies in limited environments and some of these are even being tested in deployments. In this section we introduce the most important technologies emerging to address the challenges we have just described when using the current state of the art tools.

Localization and Positioning Technology

The professional can be in a well-known and tightly defined space, for example a warehouse or a hospital, as well as in any open space: a neighborhood, a town or a geographical region; the user may be indoors and underground. Some systems must know the absolute user location (usually in Global Coordinate Reference System) while others can work entirely in relative location (relative to a known point within a coordinate reference system). There are many options for positioning the user and there will never be a silver bullet sufficiently powerful to identify precisely the location of individuals or assets in all environments and for all use cases. Choices must be made.
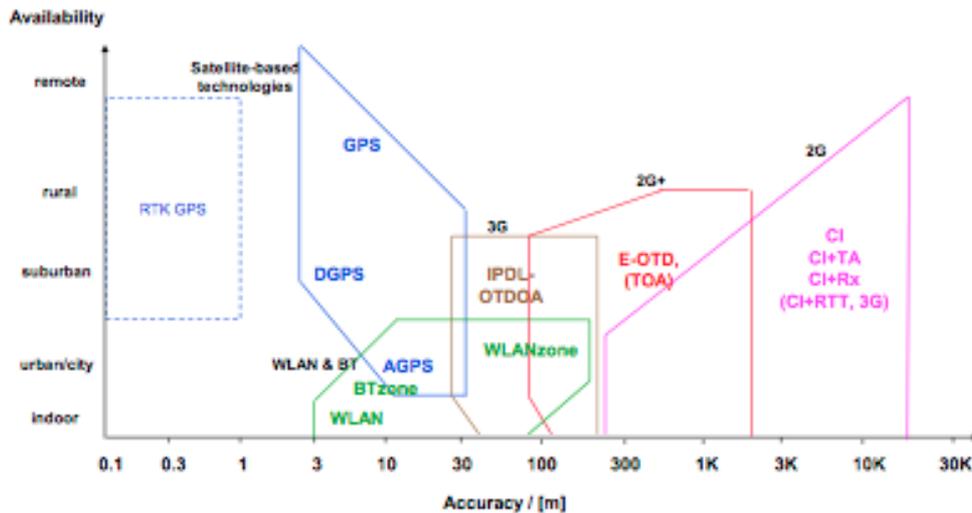

Figure 5. Accuracy of WiFi and GPS by comparison with alternate technologies. In order to select technologies appropriately to match the user needs, the primary conditions must be defined and a compromise reached with respect to positioning the user of an AR-assisted system.

Outdoor positioning technologies
Differential GPS (DGPS) relies on a network of fixed, ground-based transmitting stations to correct the error of the satellite positions and is considered superior. Precision of DGPS within 100-300km from the transmitting station is in the range of 5 to 10 cm. Their higher accuracy makes DGPS popular for urban planning and survey as well as in the military/defense and aerospace sectors where cost sensitivity is lower than for consumer/mass market devices. This could become an option for urban AR positioning.

Cell tower and WiFi antennas are also being used for positioning. Using triangulation, these technologies provide a position for an object endowed with a WiFi antenna as long as there is access to a map of WiFi access points. The precision of these systems ranges between 5m and 100m, depending on the WiFi density and accuracy of the map (access points can be moved).

Frequently, the AR tracking systems (see below) also are useful for positioning. By leveraging the readings of the gyroscope, accelerometer and compass, an Inertial Measurement Unit is simulated. The virtual IMU then positions the user with respect to the last registered (known) location. Pedestrian Dead Reckoning (PDR) is the use of algorithms that, in combination with IMU, provide trajectory for a user when walking.

Indoor positioning technologies
Due to the attenuation of signal, the GPS/DGPS system does not perform adequately for AR experiences indoors. There are many alternate methods and systems to identify the position of an item or person indoors.

Using the same principle as outdoors, WiFi signal-based positioning, virtual IMU and Pedestrian Dead Reckoning are still options. The best implementations can provide a precision of approximately 1m, however, setting up an accurate indoor WiFi map may be labor intensive and involve a calibration phase. In some cases the process can be automated at the expense of accuracy. Also, any alteration of the magnetic field could affect the precision of location identification, and in some circumstances, a recalibration of the system would need to be performed.

Other indoor positioning systems can be built upon Radio Frequency Identification (RFID) tags and readers, optical (e.g. Infrared, LED) or acoustical (e.g. ultrasound) as well as a combination of WiFi + RFID and WiFi + LED. Once the transmitters are excited and signals are detected by a suitable receiver, the indoor positioning system performs all the calculations that are required to locate a user or an asset in a given environment. These systems are in limited deployment and, where available, can be reliable.

Indoor positioning can be also achieved by means of suitable computer vision and pattern recognition techniques. The simplest solutions rely on previous knowledge of the 3D structure of a building and on markers (e.g., QR codes) attached to specific points within the building. The markers serve as location encoders and then are used to retrieve information associated with that point and its surroundings. The more computationally-complex implementation of this same principle uses the 3D structure of a building, specific landmarks and visual cues (e.g., doors, edges of signage) which can be identified and then used by algorithms to establish where the camera is pointed to calculate the position

within the environment. This requires that semantically-relevant information be part of the video (see Recognition and Tracking technology).

Finally, it's worth mentioning that using autonomous systems (e.g. robots) with camera and machine learning algorithms moving in a space are being shown to acquire the 3D structure of an indoor environment quickly and accurately. This is useful when the environment into which an AR-assisted system needs to work may be mapped in advance.

Recognition and Tracking Technology

Not only can environments be indoor and outdoor, they can also be known and unknown. In case the environment is unknown, the features of the environment need to be learned in some way.

Recognition

Once the attributes of a space are known, a suite of recognition technologies permit identification of an object, be it a person or an asset, and collect all the information associated with the same object. Recognizing a person or an asset in real world situations can be a challenging task, especially when environments change or evolve over time due to human or natural factors.

In order to overcome the difficulties of recognizing dynamic environments, several diverse pattern recognition techniques have been developed. The idea behind the best proposals is to find the invariants with respect to the transformation the object or the environment undergoes.

One of the technological solutions proposed in research is to use artificial neural networks -in a combination for example of ART (Adaptive Resonance Theory), Back-propagation and Recurrent architectures- to smoothly resolve this invariance problem. Another possible solution relies on isolating specific features of the objects that change at a lower frequency with respect to those of the changing environment.
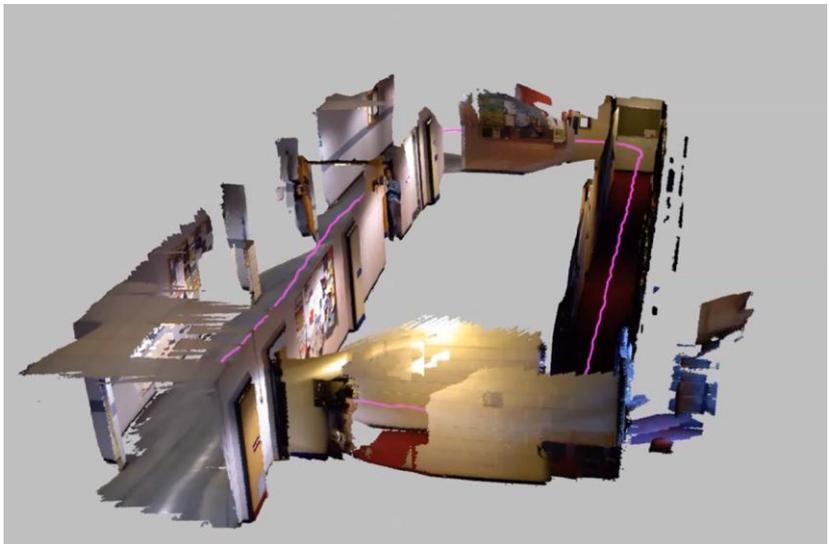

Figure 6. The environment in which an AR experience is delivered can be mapped and then easily recognized. (http://web.mit.edu/newsoffice/2013/3d-mapping-in-real-time-without-the-drift-0828.html) (http://web.mit.edu/newsoffice/2012/simultaneous-localization-mapping-kinect-0216.html)

Tracking

Recognition and detection of fixed locations or of an object in a given environment are not sufficient for certain purposes. In some urban AR use cases, the object of attention, perhaps a vehicle, is moving. The direction and orientation of a user or a target object's movement must be known within the reference system of the environment with respect to the observer's point of view (and the reverse). This kind of information is strictly required when users need to access relevant data in real time within an application domain, for example, in order to solve a maintenance problem or to successfully address a planning task in real time such as, for example, finding the nearest exit in emergency situations.

Tracking for AR is the act of continuously identifying the position and orientation of an observer's viewpoint with respect to a given 3D reference system over time. Tracking can be approached using one of three methods and a hybrid of techniques. Active tracking methods rely on sensors that are taking readings via a beacon/receiver system that must be supplied with power (e.g., depth sensing with infrared or laser, ultrasonic, GPS, WiFi). Passive tracking does not require use of power to emit or receive a signal. These are based on inertial sensors (e.g., accelerometer, magnetometer, gyroscope). The third approach for tracking is with computer vision (optical). This covers both marker-based, natural feature tracking approaches. Finally, there are hybrid tracking methods that combine inertial and computer vision-based techniques.

So, it is important to note that positioning and tracking technologies overlap with the main differences being that tracking must also:
- Register with the real world in 3 dimensions and
- Identify changes in the position and orientation of the user's viewpoint in time, using positioning sensors.

For professional users in the built environment, it is most useful to make distinctions between 2D tracking, 3D tracking and Hybrid Tracking. We have previously introduced 2D tracking and its weaknesses for the use cases proposed.

Tracking Type	Categories	Sub-categories
Optical	- Natural Feature	Edge-based tracking
		Template-based tracking
		3D feature tracking
	- Marker-based	
Inertial		
Ultrasonic		
Magnetic		
Mechanical		

Table 3. Types of Tracking for Mobile AR in Urban Professional Systems.

3D Tracking

Professional users of AR-assisted systems must have 3D tracking. There are many approaches described in academic literature to solving this challenge. They include edge-based tracking and template-based tracking, just to mention a few.

The idea here is to use features that are already present in the environment to extract information that is useful for tracking. Edge-based techniques are based on processing algorithms that are able to identify object boundaries and edges in the camera stream. The edges are then matched towards previously available 3D models to get the correct pose of the camera with respect to the object (e.g. RAPID, one of the first edge-based tracking approaches ever introduced - C. Harris, Tracking with Rigid Objects. MIT Press, 1992). Template-based tracking is another interesting solution when it comes to tracking real objects using a camera. A template refers to the combined set of projected model features and its corresponding position parameters of the model with respect to user's viewpoint. Templates are created by projecting a 3D model in an off-line process providing a discrete approximation of the observer's viewpoint. By comparing the camera signal with templates, the system can estimate the observer viewpoint based on the matching with the available templates and the images.

Both approaches provide tracking for complex 3D objects, but they can be considered as complementary in specific circumstances. The advantage of the edge-based techniques over the other approaches is that they are more robust at tracking reflective or transparent objects. Both approaches are generally good for 3D object tracking as opposed to more standard 2D image tracking. The drawback is that they are both more computationally expensive than traditional tracking techniques. This is not considered a limitation for professional user systems.

Hybrid Tracking
Finally, some professional application developers are beginning to use the combination of optical tracking methods and inertial methods, i.e. optical-inertial sensor fusion. This permits continuous tracking in conditions where information about the environment is so sparse that neither method will work alone at a satisfactory level. Applications based on such approaches are currently being developed in sectors related to safety and maintenance in urban settings.

Integration with Existing Data Infrastructure

Taken together, the developments described above will soon permit AR-assisted systems to be introduced and tested in combination with existing IT and information management systems. Data collected and maintained by GIS and other municipal services as well as by the agencies providing planning, clean water, energy, waste removal and other essential services can be used to guide people in the field in real time.

There cannot be a custom integration performed in each city and for every data set in every agency. Standards for querying and interfacing with data are the only way that these opportunities can be addressed scalably and cost-effectively.

There are a few widely-recognized data formats and information management standards that serve the urban data sector. Both public and private institutions (companies) use Geographic Information Systems (GIS) to store and retrieve data while some disciplines are using Building Information Modelling (BIM) workflows. An important part of the data infrastructure relevant in the urban sector is the Spatial Data Infrastructure (SDI), which is currently under construction in USA, Europe and United Nations.

The final stage is to add AR to existing urban IT platforms. In light of the fact that there are well-developed urban data formats and there are no standards for displaying or using AR, it is up to those with the data and the requirements to communicate these to their IT suppliers and systems integrators.

Predicting Return on Investment

In this paper we suggest many ways that municipalities and providers of services in the urban environment can increase the accuracy or reliability of services they provide to professionals, and in turn to citizens and visitors, with AR-assisted IT systems. In order to estimate the return on investment for implementing AR programs, several factors need to be considered.

City size and age

To estimate the ROI, we may factor in the city size and age. In our experience with AR adoption, cities with more than 100,000 inhabitants rarely make decisions and investments as single units. Rather, they operate as a collection of loose autonomous units with human and financial resource limitations. A city's centralized cadaster and land survey service bureau is the best place to begin as this group may have the most accurate maps and even 3D city models. Cities with less than 1,000 people are challenged with having insufficient revenues to have dedicated staff for IT as well as for other services for centralized planning or environmental monitoring.

City age and architectural style also impacts the ROI of AR technology investments. Cities in which building facades are highly variable are easier to recognize using elementary computer vision algorithms but very expensive to model. On the other hand, a very homogenous, modern architecture style is more difficult for recognition, registration and tracking purposes.

Figure 7. A highly reflective city façade is difficult to recognize due to lack of unique natural features.

Cities where the utilities were embedded under roads and buildings prior to the 1980s will have far greater challenges with data accuracy and reliability than those in which access to precision measurement and mapping tools was high during building phases. Cities with recently built or improved transportation systems also have a far higher likelihood of maintaining accurate measurements of the built environment, including 3D models produced from LiDAR or other scanning methods.

City IT infrastructure

As mentioned, successful deployments of AR in an urban environment will benefit greatly from an advanced IT infrastructure. First and foremost, access to geospatial data sets with standard encodings, metadata using well-documented semantics and managed in a modern database with Web service interfaces will accelerate the development of an AR-assisted solution using the city data.

Networking in the urban environment can also be a factor to consider in the ROI. Where there are already municipal WiFi network services, these can be used to relay observations from sensor networks and nodes. These also tend to encourage the citizens to use city services and to report where there are services in need of attention (roads, trash, etc).

Impacts of delays and errors

When estimating the ROI of an investment to support AR-assisted services it will be necessary to include time for delays in deployment as well as, in some circumstances, to wait for the technology to decline in cost. There are also costs that can be expected with erroneous data that is not discovered prior to using Augmented Reality to overlay the digital data on the physical world.

Finally, there may be need for citizen consensus building for widespread AR use as there is in some regions for the deployment and use of other fixed camera/video infrastructure. People who benefit from the presence of the technology will be in support while others will be suspicious and fearful of the impacts of technology on their freedoms and lifestyles.

Calculating the AR visualization Benefits

An ROI calculation must also estimate the actual benefits of AR visualization in terms of lowering the cost of doing business, reducing risk of property loss or risk to the lives of emergency responders and other factors. Today, the benefits of professional uses of AR are difficult to measure quantitatively.

The first step is to examine specific use cases. Such as those described in the first section of this paper. Within each use case, a systematic approach will include:

- Capture the number of users of a system prior to introducing an AR-assisted feature. With this background data, it will be possible to compare "before" with the "after."
- Determine current spending on services that are being targeted for AR. The budget will initially increase as an investment in AR is made (see next section)
- Estimate cost reduction through fewer city staff trips to the field to support professional users, fewer errors or accidents or the number of complaints.
- Estimate the benefits as a result of superior access to data visualization and citizen understanding, or any improvements in the quality of city data as a result of urban professionals finding errors in data when visualizing in the physical world.

Next, the impact of urban AR on professional users must be measured by the groups of users themselves when in the field or upon completing a post-use survey. In this case, the emergency responder, or the architectural firm's

employees should be requested to estimate the time saved, or the risk reduced as a result of the ability to visualize the digital information in the real world.

Costs of implementing and maintaining AR-assisted IT systems

AR-assisted IT systems are provided by professional systems integrators around the world, including SAP, Oracle resellers and others. The best systems will begin with a pilot or prototype project. This should include training of city IT staff so that an in-house expert (champion for AR) can provide assistance to those considering how to use AR as part of their department's services.

The cost of adding AR will probably be billed on a use case basis, where until these are commonplace, the service provider will design a user interaction best suited to the task. The design of AR interfaces should not be done outside the context of the larger workflow, making it intuitive for the user to continue their work without camera-view using power or interfering with the decision making process.

Next Steps

Professionals in the urban environment are already accustomed to using mobile technologies and the municipal departments are already maintaining city data. We recommend that cities exploring the possibility of supporting the urban professional with AR-assisted systems organize a workshop of stakeholders during which the project needs/requirements are identified and prioritized by a coalition of qualified people.

A workshop is also the first step to forming a group or committee (task force) that will be responsible for formulating a vision for the city's or organization's future IT support for AR visualization. A spokesperson may be appointed to communicate the AR program to different other agencies, businesses and the media. A committee will also be preparing a request for proposals to submit to IT systems integrators and soliciting financial support from decision makers' budgets.

As AR technologies become more open and interoperable with existing IT systems, it will also be beneficial to implement support for open standards in city GIS systems and other Urban AR content management systems.

For more information or to schedule a workshop, please contact the authors.
Christine Perey cperey@perey.com +41 79 436 6869 www.perey.com
Graziano Terenzi g.terenzi@armedia.it +39 0775 1886100 www.armedia.it

Acknowledgements

For more information or to schedule a workshop, please contact the authors.

Christine Perey   Graziano Terenzi
cperey@perey.com   g.terenzi@armedia.it
+41 79 436 6869   +39 0775 1886100
www.perey.com   www.armedia.it

Research & Consulting