A Visualization Paradigm for 3D Map-Based Mobile Services

Mário Freitas[1] and A. Augusto Sousa[12] and António Coelho[12]

[1] FEUP, Rua Dr. Roberto Frias, s/n 4200-465 Porto, Portugal
[2] INESC Porto, Campus da FEUP, Rua Dr. Roberto Frias, s/n 4200-465 Porto, Portugal

Abstract. Nowadays, there is a wide range of commercial LBMS (Location-Based Mobile Services) available in the market, and a trend towards the display of 3D maps can be clearly observed. Given the complete disparity of ideas and a visible commercial orientation in the industry, the study of the visualization aspects that influence user performance and experience in the exploration of urban environments using 3D maps becomes an important issue. A generic conceptual framework is proposed whose main purpose is to objectively evaluate the impact and contribution of the major visualization elements involved. An online questionnaire was developed and administered to 149 test subjects in order to measure the real impact of each element. Combining the experimental results with the current state-of-the-art, a new visualization paradigm is defined in a dual specification: "layers" providing relevant visual content to the map, and "functions" providing the necessary functionality.

1 Introduction

The LBMS technology, namely in the form of GPS-based navigation systems, has just recently reached a state of technological maturity, enabling the development of 3D map-based graphical interfaces. Nowadays, there is a wide offer of LBMS solutions in the market, especially in the form of automotive navigation systems. Motivated by commercial interests, many of these products promise to offer the "best visualization experience ever", in search for a differentiating factor from the competition. By looking at the variety of visualization paradigms being proposed, one can clearly notice a great disparity of ideas without a clear notion of its usefulness. Provided the non-existence of an objective state-of-the-art generalizing theory capable of unifying and evaluating all the visualization elements and properties, the main motivation of this work is to study the most relevant of these features and how to adjust them appropriately, in order to maximize the usability of mobile maps and to improve the navigation experience, in accordance with the following objectives:

1. Elicit and assess the state-of-the-art contributions on visualization paradigms of 3D maps, with particular interest on mobile services and devices;
2. Develop a methodology for evaluating the different issues that influence user experience and performance when exploring an urban environment with mobile maps;
3. Define a new visualization paradigm of 3D maps for urban environments.

1.1 State of the Art

1.2 Visual Perception of Realism

The variety of free and commercial products featuring three-dimensional map-based mobile services available to the masses, usually ranges from very abstract to reasonably realistic and immersive visualization paradigms. However, there is a common misconception on what is *Image Realism*, how is it visually perceived, and how can it be effectively measured. In [1], a scientific experiment was conducted to understand what aspects of an image can make it look "real" or "not real", i.e., whether it is perceptually indistinguishable or not from the corresponding photographs. The results showed that subjects were not convinced by the increasing number of light sources and shadows nor the variety or number of shapes. The same could be said for "perfectly sharp" shadows or "perfectly polished" surfaces. In [2], an experiment was carried out with 75 test subjects to classify 90 images of the virtual landscape of Brunnen / Schwyz (Switzerland) from three different viewpoints in a degree of realism from 1 (very low) to 5 (very high). The results generally demonstrated that the variable that most contributed to the sense of realism was – by far – the high-resolution orthophotographic imagery, and the second most important being texture-mapping. In other works like [3], the importance of perception-based image quality metrics is studied, such as the ones given by the VDP (Visible Differences Predictor) and the VDM (Visual Discrimination Metric). These two metrics aim to analytically predict the differences between a computer-generated image and the photograph it depicts, taking into account the limitations of the human eye described by the HVS (Human Visual System). The VDP quality metric takes the two images as input and generates a *difference map* that predicts the probability of the human eye finding differences between the two pictures, as demonstrated in [4].

A simplification of the VDM quality metric was provided by following a similar approach [4]: instead of finding a *difference map*, a *just noticeable difference map* was proposed which corresponds to a 75% probability of a person detecting a difference between the two images [3]. Because of some controversy and no agreed-upon standards for measuring realism in computer-generated imagery, a conceptual framework for measuring image realism and evaluating its usefulness was proposed in [5]. The framework distinguishes three different varieties of realism: *physical realism*, *photo-realism* and *functional realism*. However, this framework does not seem to be enough to encompass the extents to which reality or virtuality can be "augmented". Accounting for such circumstances, the concept of *Virtuality Continuum* was introduced in [6] (see Fig. 1).

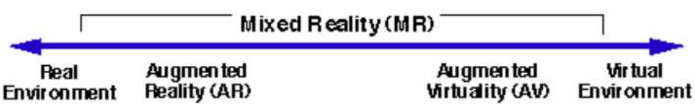

Fig. 1. The Virtuality Continuum

At the left end, we have the “completely real” *Real Environment*, which is made up of “real” objects: “any objects that have an actual objective existence”. At the right end, we have the “completely computer-simulated” *Virtual Environment*, which is made up of “virtual” objects: “objects that exist in essence or effect, but not formally or actually”.

1.3 User tasks

The underlying basic equation that can help us find the “perfect” balance in map-based mobile services is what could be called of *Mobility Equation*. This equation was first formulated by Leonard and Durrant-Whyte for mobile robot navigation [7] but can be equally extended to human navigation. The equation is made up of the following three questions:

– ‘Where am I?’
– ‘Where am I going?’
– ‘How do I get there?’

In [8], the tasks are classified into 4 different groups of high-level user tasks that have a strong relationship with these questions, as described in Table 1.

Table 1. The primary tasks that 3D maps are used for

Task	Description
Locator	Identification of the user’s own position and other objects. Answers ‘Where am I?’ questions.
Proximity	Inform the users of nearby facilities. Implied by ‘Where am I going?’ questions.
Navigation	The most tangible example is routing from one location to another. Answers ‘How do I get there?’ questions.
Event	Time/Location dependent objects, allowing the users to know what is happening and when/where. Answers ‘And now what?’ questions.

1.4 Location-Based Mobile Services

In this work we have analyzed and studied several state-of-the-art contributions on LBMS which provide a wide variety of visualization paradigms, in order to understand the current tendencies in the industry and to formulate hypothesis regarding their validity and usefulness. The contributions range from pilot studies to commercial products, within the scope of road and pedestrian maps, as follows: TellMaris, m-LOMA, LAMP3D, TomTom, Navigon, NDrive, iGO, Google Earth, INSTAR, Virtual Cable^{TM}, and Enkin.

2 Conceptual Framework

In this section, a generic evaluation framework is proposed which can be used as the main methodology for the specification, development and evaluation of new or existing solutions in the visualization problem domain. This framework is proposed in order to simplify the evaluation process to the most relevant features, to the detriment of other classical analysis methods that can be used to obtain a more thorough evaluation. This framework defines the concept of *feature vectors* comprising *orientations* and *magnitudes*. The *orientation* defines the idea or concept the visualization paradigm represents, and *magnitude* the degree/level to which the paradigm "amplifies" the vector. An example can be seen in Fig. 2 to describe a possible feature vector for transportation. An orientation of this feature vector is the mode of transport, while pollution, cost and speed are magnitudes.

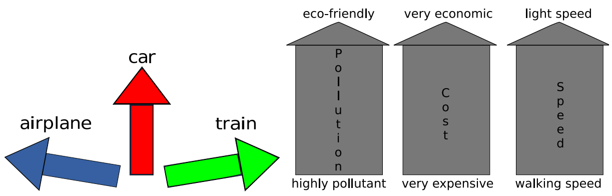

Fig. 2. A possible feature vector for "Transportation"

The framework is composed by six feature vectors as described below. These feature vectors are not intended to characterize the complete set of visualization features, but the most relevant ones observed from the current state of the art described in Sect. 1.4.

2.1 Image Realism

Image Realism is the feature vector that is concerned with how real, i.e., free from any idealizations or abstractions, is the image of the map presented to the user. Taken into account what was previously mentioned on this matter (see Sect. 1.2), the suggested *magnitudes* for this vector will be based on the framework proposed in [5] and the concepts on *virtuality continuum* defined in [6], with a few modifications. Firstly, a "relaxed" version of *physical realism* will be adopted, i.e., it is assumed that current displays are considered perfect in the sense that they can emit the actual energy we want them to reproduce. Secondly, this framework will be incorporated into the *virtuality continuum* as illustrated in Fig. 3, adapted from the above work.

Photo-Realism is located to the left of *Functional Realism*, not because it is considered "less virtual" than *Functional Realism* but because it is closer to the *Physical Realism*, and consequently providing a more "realistic" environment. In terms of *orientations*, this vector includes the visualization elements that represent the real world visual information, namely *3D Buildings* (city buildings, landmarks), *Map Vectors* (roads and polygons), and *Surface Model* (ground surface elevations).

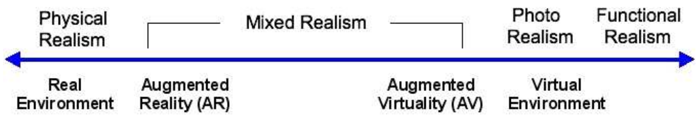

Fig. 3. An illustration of the proposed framework combining the Virtuality Continuum spectrum with varieties of image realism (adapted from the previously mentioned work)

2.2 Object Labeling

Object Labeling encompasses the kind of visual techniques and strategies that are followed to label map elements such as rivers, streets, cities, and so on. In [9] and other studies, the importance of two types of labeling, namely *static labeling* and *dynamic labeling*, is discussed. This is relevant to distinguish since, depending on the case, we might be dealing with dynamic maps, i.e., maps that support continuous zoom (changing the scale) and continuous panning (usually by dragging the map). Based on the framework proposed in the previous study, the *magnitudes* for this vector will include the concepts of *Static/Dynamic Selection* (visibility) and *Placement* (size, position and orientation) of labels. One of the possible approaches when labeling objects is to project the labels oriented towards the current perspective, analogous to a billboard in Computer Graphics. This approach is followed by all the contributions except *Google Earth* where labels are flattened and laid down on the maps surface. Based on the works of [10, 11] and the previous discussion on adaptiveness to the current perspective, the proposed *orientations* for *Object Labeling* are *Perspective-Adaptive* (oriented towards the current perspective), *Point Positioning* (point symbols), *Line Positioning* (polygonal chains, such as rivers), *Area Positioning* (areal features such as countries), and *General Positioning* (a combination of the 3 previous methods).

2.3 Visual-Spatial Abstraction

Visual-Spatial Abstraction measures the complexity of mental operations that are required to perform the visual matching of the real environment that can be observed and the one on the screen. This vector is specifically focused on the mental viewing transformation that is required in order to have a perfect correspondence between both images: the reality and the screen. The proposed *orientations* for this vector are presented, regardless of the elevation angle of the "camera", namely *Ground Level* (when it is only possible to observe the current street and its junctions), *Local-Area Level* (when streets that may not even be part of the route can be observed), and *Wide-Area Level* (when municipalities and an overview of the route are visible). The proposed *magnitudes* reflect the adaptiveness of the camera to the users' behavior. We define *Adaptive Level* and *Adaptive Orientation* when the camera adapts to the user's movement (according to some variable like speed), and whether it adapts to his looking direction, respectively.

2.4 Route Indication

Route Indication provides a classification of the visual techniques and strategies for showing the itinerary path in the road maps, and the kind of maneuver indicators or way points that are presented in the display. The proposed *orientations* for this vector, can be regarded as the visual indicators that are generally used by the majority of the contributions to display the route, namely *Arrows*, *Cords*, *Way Points* and *Carpet*-like shapes to indicate the route. These indicators can be used with different “immersion” levels which are considered the proposed *magnitudes* for *Route Indication*, namely *Instructive* (when indicators are merely instructive) and *Simulative* (when they resemble real world indicators).

2.5 Landmark Symbology

Landmark Symbology evaluates the cartographic symbology that is used to portray the world using a pictorial language, represented by “map symbols”, often accompanied by a legend. This vector is also related to *Image Realism*, in the way that both should be complementary, i.e., excessive realism may distract the users, but a great lack of symbology may completely blur their sense of orientation. New concepts and design guidelines for the cartographic visualization of landmarks in mobile maps are proposed in [12]. Based on these concepts, the *orientations* for this vector will reflect the kind of buildings represented by symbols, specifically *Shops referenced by name* (e.g., KFC, McDonalds), *Shops referenced by type* (e.g., hotel, pharmacy), *Buildings with unique name / function* (e.g., Tokyo Tower, Statue of Liberty), and *Buildings with unique visual properties* (e.g., “the large yellow house”). Additionally, the first proposed *magnitude* for this vector will define in itself, the concept of levels of abstractions for landmarks, according to a scale (from the most abstract, to the most concrete): *Words*, *Sign*, *Icon*, *Sketch*, *Drawing*, and *Image*, as defined in the previous study. There are other parameters that influence the decision of whether an abstraction level should be used in a mobile map for a given situation. For instance, some cartographic generalization procedures (like scaling down a landmark object to an appropriate size suited for its representation in a map) might raise some problems such as *congestion*, *coalescence*, and *imperceptibility* [12]. To account for these restrictions, the proposed *magnitudes* consist of *Adaptive Zoom* and *Adaptive Complexity*, respectively, whether the abstraction level of landmarks adapts to the current zoom level, and whether they change with the varying complexity of features.

2.6 Contextual Awareness

Contextual Awareness measures the extent to which a visualization paradigm is applied to get additional information on a contextual or situational basis. It is important to distinguish the three groups of application areas in which virtual urban environments can be valuable, according to the spatiotemporal nature. These groups constitute the proposed *orientations* for this vector, depending on whether they focus on the past, present or fiction, according to [13]: *Reconstructional* (reconstruction of urban environments that were totally or partially lost), *Recreational* (urban design, urban planning, etc.), and *Fictional* (creation of imaginary realities). Levels of awareness regarding the current location, time, and situation can vary from contribution to contribution. In [14], it is claimed that a passive contextual-awareness approach is generally more flexible than an active approach. In the latter case, if the user is constantly presented with unwanted information it can become “too obtrusive”. Contrarily, in most automotive navigation systems, direction instructions or location-based information such as nearby Points of Interest (POI) are automatically presented, i.e., without the need of the user’s intervention. For these reasons, the proposed *magnitudes* for this vector will reflect the different autonomy levels of “contextual awareness” an application can demonstrate in different contexts and tasks, as previously denoted by [15], specifically *Active Awareness* (without the need of user intervention), and *Passive Awareness* (when the user shows interest for getting context-based information). Table 2 summarizes the evaluation framework, according to the proposed *magnitudes* and *orientations*.

Table 2. Structure of the proposed evaluation framework

Feature Vector	Orientations	Magnitudes
Image Realism	3D Buildings, Map Vectors, Surface Model	Physical Realism, Mixed Realism, Photo-Realism, Functional Realism
Object Labeling	Perspective-Adaptive, Positioning (Point, Line, Area, General)	Static / Dynamic Selection / Placement
Visual-Spatial Abstraction	Ground Level, Local-Area Level, Wide-Area Level	Adaptive Level, Adaptive Orientation
Route Indication	Arrows, Cords, Way points, Carpet	Instructive, Simulative
Landmark Symbology	Shops (referenced by name), Shops (referenced by type), Buildings (with unique name / function), Buildings (with unique visual properties)	Abstractness (Words, Sign, Icon, Sketch, Drawing, Image), Adaptive Zoom, Adaptive Complexity
Contextual Awareness	Reconstructional, Recreational, Fictional	Active Awareness, Passive Awareness

3 Methodology

An interactive online questionnaire was developed and several hypotheses were formulated, in order to assess the real impact of each visualization feature described in the conceptual framework. An interactive online questionnaire was developed specifically for this study, enabling the measuring of time for each answer and a more adequate visual aspect definition. However, due to the intrinsic limitations of the proposed questionnaire, and in order not to make it perceived by potential participants as “too exhaustive”, only the features for which there are no significant indications from the state-of-the-art (regarding their impact and relevance) were evaluated with the questionnaire. Moreover, there are some components that were not possible to evaluate, given the limitations imposed by this kind of questionnaire. The questionnaire was divided into 3 parts. In the first part, the exercises were mainly based on the *pointing task paradigm* as previously performed in other studies [16]. In the second part, a similar approach was followed, but instead of evaluating the matching of the two realities, the main objective was to measure how well users perform a given task (see Sect. 1.3). In the last part, users were asked about their preferences regarding the visualization of map elements such as landmarks.

3.1 Image Realism

All *Image Realism orientations* were tested along with the various degrees of *magnitudes*, in accordance with the vector instances (*orientations* and *magnitudes* combined) found in the state-of-the-art contributions. These instances were considered eligible for the evaluation through the questionnaire, since there are few or no indications, with regards to their impact:

- *Simple Textured Buildings* versus *Photo Textured Buildings*
- *Colored Map* versus *Orthophotomap*
- *Flat Model* versus *Terrain Model*

It was hypothesized that, in the absence of *Simple Textured Buildings*, test subjects will have to rely on their ability to match the 3D geometry of the real building with the geometry of the 2D polygon representation on the map. At the same time, it is supposed that by providing the three-dimensional (yet simple) geometry of the whole building, in the presence of this component, test subjects will make fewer mistakes and, as a consequence, will require less time matching both realities (see Fig. 4, left part). In the case of the *Photo Textured Buildings* component (see Fig. 4, right part), it was hypothesized that, by simultaneously providing the 3D geometry of a building along with photographic façades, test subjects will be able to detect features (e.g. windows, doors, unique wall patterns, etc.) more accurately and faster than in the case of *Simple Textured Buildings*.

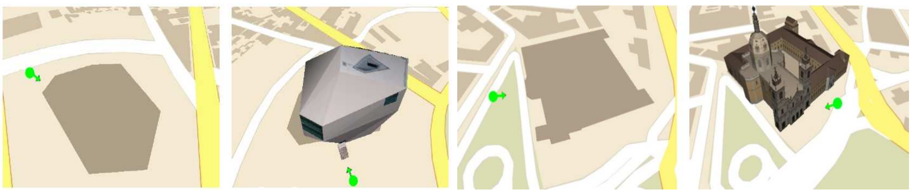

Fig. 4. The 4 images supporting the questions that evaluate Simple vs. Photo Textured Buildings

Regarding *Map Vectors*, it is assumed that an *Orthophotomap* can provide subjects a much more enriching visualization experience than the one provided by a *Colored Map* (see Fig. 5, left part). The hypothesis rests on the belief that an *Orthophotomap* component can make easier for users to discern the true features of the map's surface, by giving a realistic view rather than a rough generalization. There are many situations were colored vector polygons are not enough to represent features like a tiled pavement; a group of trees arranged in a special and unique way; and several "static" features like public benches, zebra crossings, and many others that are impossible to find in a colored vector map. In terms of *Surface Model*, it was hypothesized that by using a *Terrain Model* rather than a *Flat Model* component, users will be able to perform the spatial matching of both reality and virtuality in a much more immersive and natural way (see Fig. 5, right part). It is expected that by providing the *Terrain Model* component, users will be able to use elevated reference points, and to understand and visualize occlusions caused by the varying landscape elevation.

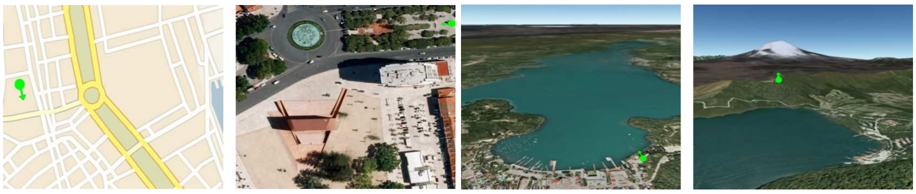

Fig. 5. The 4 images supporting the questions that evaluate Colored vs. Orthophoto Maps and Flat vs. Terrain Models

In the end, it is expected that users will be able to perform their tasks in less time, since they just need to think “outside the box”. On the other hand, by using a *Flat Model*, users would understand that the image on the screen does not account for occlusions, and therefore, they would have to do that job themselves.

3.2 Object Labeling and Route Indication

With respect to *Object Labeling*, it was hypothesized that, when users are analyzing labels (e.g. of streets, rivers, cities, and so on) which are not oriented towards the current viewing direction depicted in the device, they will feel much more difficulty reading the words, due to the decreased visibility, especially when looking in a direction which is parallel to the map’s surface (see Fig. 6, left part). In such case, users will not be able to read labels as faster, and will pan the map closer to the camera so it becomes easier to read. Particularly in the case of labels which are almost parallel to the camera’s viewing direction, some users will wish to skip words, if they find them “too difficult” to read. In terms of *Route Indication*, it was hypothesized that, when a user is presented with an image which looks more familiar to him, given the current context, he will be able to perform his task with lesser effort (see Fig. 6, right part). It is assumed that users won’t make more mistakes using one approach or the other, but that a significant difference in the time they require to complete their task may arise, i.e., that a *Simulative* component will result in faster responsiveness than an *Instructive* approach.

3.3 Landmark Symbology

For this feature vector, it was hypothesized that users will require *Adaptive Zoom* functionality, i.e., that the majority of them will choose an abstract landmark representation

Fig. 6. The tasks that evaluate Perspective-Adaptive Labeling and Instructive vs. Simulative route indications

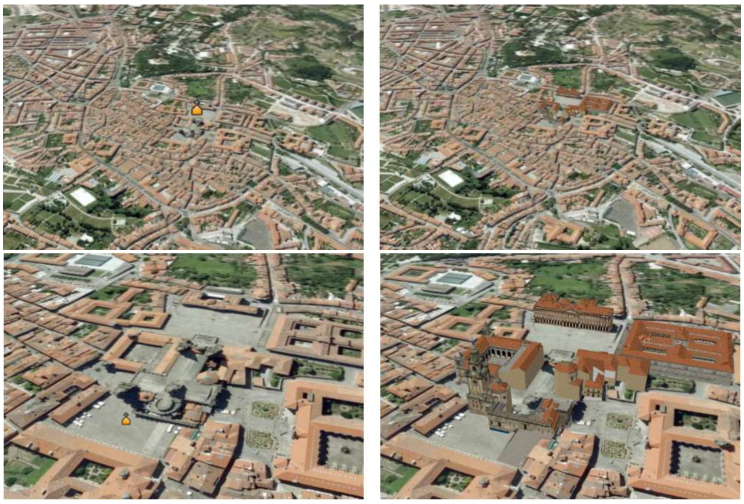

Fig. 7. The preferences that evaluates the users' need for an Adaptive Zoom approach, when a map which is zoomed out far / zoomed in close to the ground is used

of a given building, when a map which is zoomed out far from the ground is used, but a more concrete representation when at close range (see Fig. 7).

The basis of such hypothesis rests on the various issues raised by the cartographic generalization procedures, as previously explained in Sect. 2.5. For instance, even if a concrete landmark is used rather than an abstract representation, there are certain zoom levels of a map which do not allow users to perceive enough features of that landmark, in order to identify it with a significant confidence level.

4 Results

In total, 149 test subjects answered the questionnaire, mostly from a student population in Computer Science and Informatics: 89% were male, and 78% were in the 18 to 25 age group. In general, prior to answering the questionnaire, subjects considered themselves fairly capable of using both maps and GPS navigators, given the approximate 50-50 ratio shared between "average" and "experienced" users. Only 3% of the participants reported they were unfamiliar with either maps or GPS navigators.

4.1 Image Realism

Regarding the impact of the presence and absence of *Simple Textured Buildings*, there were 91% and 77% correct answers, respectively, in both situations. Although slight, the difference between the two cases shows the advantage of the presence of *Simple Textured Buildings* over its absence. Test subjects required, in average, 11s (7.3s standard deviation) to answer when buildings were shown, and 15s (6.4s s.d.) using a classic 2.5D map. While in this case there was just a 14% difference in the number of correct answers, in the case of *Photo Textured Buildings* component there were 88% and 30% correct answers respectively. Despite this difference between both questions, the number of correct answers in the presence of *Photo Textured Buildings* was almost the same as in the case of *Simple Textured Buildings*. In terms of answers times, 95% of the subjects had already answered before the first 21s in the presence of *Photo Textured Buildings*, about 4.4s less than in the presence of *Simple Textured Buildings*. When the buildings were all removed from the exercise with *Photo Textured Buildings* (i.e., in its absence), 95% of test subjects answered before the first 42.7s (avg. 17.8s, s.d. 14.4s) against 25.4s (avg. 15s, s.d. 6.4s). This clearly demonstrates that the results with *Photo Textured Buildings* are more stable, considering the increase in difficulty of the exercise. In the presence of a *Colored Map*, the number of participants who were unable to answer the question was quite high (14%). The same happened with the number of wrong answers being quite different from the *Orthophotomap* (67% and 7%, respectively). Nevertheless, subjects had no apparent difficulty in finding the correct answer, in the presence of the *Orthophotomap* component, as 92% chose the correct answer in similar conditions (as shown in Fig. 8). Besides being more effective, the *Orthophotomap* proved also to be more efficient as subjects took an average time of 9.3s (s.d. 18.4s) to answer the question, considerably faster compared to the 23.5s (s.d. 16.8s) in the case of the *Colored Map*.

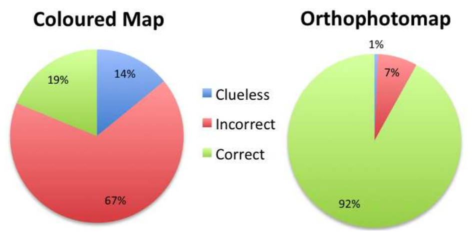

Fig. 8. Answers in the presence of a Colored Map and an equivalent Orthophotomap

In terms of *Surface Model*, there was just a 5% difference in the number of correct answers between both cases, with advantage to the *Terrain Model*. However, the *Terrain Model* was much more efficient, as the average response time was 7.5s (s.d. 5s), compared to the 15.3s (s.d. 13.8s) obtained with the *Flat Model*. These results point out that image realism can improve the task of matching the 3D map with reality, both maximizing effectiveness (lesser mistakes) and effectiveness (lesser time).

4.2 Object Labeling

With respect to *Object Labeling*, when labels were oriented towards the camera, the subjects took lesser time to perform the task (avg. 11.8s, s.d. 5.2s) than when labels were not oriented according to the camera (avg. 15s, s.d. 6.4s), as shown in Fig. 9.

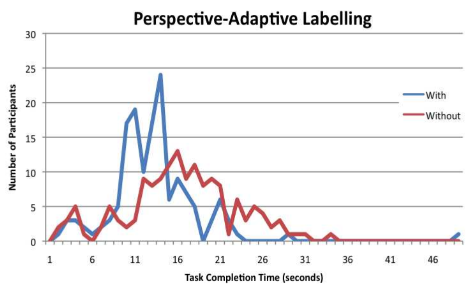

Fig. 9. Answer times in the presence and absence of Perspective-Adaptive Labeling component

From these results a conclusion can be made that *Perspective-Adaptive Labeling* can increase readability of labels in 3D maps.

4.3 Route Indication

With respect to the *Route Indication* there was no relevant difference in terms of answer correctness between *Instructive* or *Simulative* components. In terms of average answer time, the *Instructive* component resulted in 11.8s (s.d. 7.9s) against 8.6s (s.d. 5.9s). Provided with the *Simulative* approach, 95% of the participants had already found the matching route indication, after about 19.8s, compared to 26.6s in the opposite case. Although both techniques can achieve similar levels of correctness, the *Simulative* approach can speed-up the task of matching reality with the 3D map. This can be of great importance when supporting activities that demand short response times, such as driving.

4.4 Landmark Symbology

A vast majority of participants (87%) answered they would more easily identify and recognize the presence of a given distant landmark, when an abstract representation of that landmark was used. Approximately 86% of them indicated their preference towards the use of concrete landmarks at close range. Different zoom levels over 3D maps will encompass also different levels of visual complexity, and as such, *Adaptive Zoom* functionality is of great importance for maximizing readability.

5 Feature Vector Analysis and Discussion

This section presents an analysis of the evaluation performed on the conceptual framework, by combining the results of the questionnaire with a discussion of the feature vector components that were not evaluated, based on state-of-the-art studies and empirical knowledge of 3D map-based mobile services. The best approaches to follow are discussed, regarding the individual impact of each feature vector component.

5.1 Image Realism

The results of the questionnaire clearly identified the three best approaches, by means of identifying instances (combining orientation and magnitude) of *Image Realism*:

- *Photo* rather than *Simple Textured Buildings*;
- *Orthophoto* rather than *Colored Map*;
- *Terrain* rather than *Flat Model*.

It is safe to assume that *Mixed Realism* or *Photo-Realism* cannot be used interchangeably at any time, i.e, it cannot be affirmed that their use does not affect users performance nor is limited by any restriction. *Virtuality*, if used alone, is enough to allow a user to perform every task, although not as immersive as the experience provided by an *Augmented Reality* (AR). However, there is an intrinsic limitation to this kind of paradigms, because they cannot provide information apart from what the built-in cameras can capture. This kind of paradigm is ideal when the user is driving a car, i.e., performing a navigation task [17]. The problem arises when the user tries to apply this paradigm outside the reduced visibility range provided by the live imagery, like when planning an itinerary, or to get a route overview. Therefore, it is believed that the merging of both paradigms could prove to be the best choice. The advantage given by *Mixed Realism* can make users tasks easier, but the scope of these tasks is very limited. In general, it is believed that the combination of both, *Photo-Realism* and *Mixed Realism* magnitudes, will provide the best results.

5.2 Object Labeling

Based on the study of the state-of-the-art contributions, it can be observed that the current tendency in this feature orientation is towards the adoption of *General Positioning* algorithms. Rivers and other polygonal chains are best described using a *Line Positioning* approach; countries and other polygonal features are more easily recognized using an *Area Positioning*; and finally, cities and municipalities are better perceived if they are indicated with a *Point Positioning* approach. Additionally, and provided that most applications have evolved from simple city guides to general purpose mobile maps, the combination of the 3 strategies represented by the *General Positioning* algorithms becomes the most appropriate choice. The results of the questionnaire also demonstrated that a *Perspective-Adaptive Labeling* approach, i.e. labels oriented towards the current perspective, results in faster responsiveness from the users. The reason for this, is that users generally feel a higher degree of difficulty reading labels that are laid down along the street vectors, and have therefore reduced visibility and readability, instead of being oriented towards the camera. Regarding the proposed magnitudes for object labeling (*Static / Dynamic Selection / Placement*), the desiderata described in [9] should be taken into account when creating dynamic maps, otherwise it will be very difficult for users to perform the matching of labels with the corresponding point, line or polygonal features of the map. For example, suppose a user is reading a given city name label; he/she pans the map around just a bit, and then the city label vanishes completely or suddenly 'pops up' (i.e., in a discontinuous manner) in a completely different position on the screen. This is why, the current tendency is to find stable and fast algorithms, while pursuing *Dynamic Selection* and *Dynamic Placement* of labels.

5.3 Visual-Spatial Abstraction

There are already several directions provided by the state-of-the-art contributions, with regards to *Visual-Spatial Abstraction*. First of all, there is an evident tendency towards the use of all the camera level orientations in mobile map applications (*Ground Level*, *Local-Area Level* and *Wide-Area Level*). We define *Ground Level* when it is only possible to observe the current street and its junctions, *Local-Area Level* when we can observe streets that may not be part of the route and *Wide-Area Level* when an overview of the route can be possible. To understand when to use which level, one needs to understand the difficulties and requirements expressed by test subjects when performing tasks with mobile maps. For instance, due to the nature of the *Ground Level* perspective, one can argue that the walking view mode is among the best choices for *Locator Tasks*. In the top view mode, since buildings cannot occlude other buildings, it becomes easier to search for nearby facilities, which is the purpose of *Proximity Tasks*. On the other hand, both flying and top view modes can be used for orientation purposes, as the users are able to get a better overview of the surroundings and will more likely spot a landmark or reference a point which is part of their spatial knowledge. In fact, test subjects already mentioned that when they are performing a navigation task, it is sometimes very useful to zoom out (i.e., to raise the camera level) just enough to get a better overview of the surroundings, and possibly spot a landmark or some feature that is part of their spatial knowledge. We define *Adaptive Level* when the camera adapts to the user's movement (according to some variable like speed) and *Adaptive Orientation* whether it adapts to his looking direction. It is believed that an *Adaptive Orientation* approach is generally the best to follow. This approach is advocated in most AR-based prototypes and experimental studies. The results of these studies [18] not only demonstrated that the 3D maps provided shorter completion times, but also that the worst problematic misalignments between the 2D and the 3D maps could be found at the conditions of 90 degrees. Provided that misalignments and the wide range of possible viewing directions can pose a big problem to the success and effectiveness of a user task, by causing a lot of confusion among users, it is believed that an *Adaptive Orientation* approach is an important requirement, for the purpose of solving such issues. Regarding the *Adaptive Level* magnitude, there are already some options available that take contextual variables into account. For instance, when a user is driving slow, and suddenly presses the accelerator to achieve a higher speed, the automotive navigation system will generally adapt to this situation by increasing the camera altitude, providing a better overview of what is ahead.

5.4 Landmark Symbology

The main question regarding this feature vector is how and when should each building be represented. To answer this question, the Adaptiveness magnitude for this vector should be taken into account. According to the experience in practical use of maps, all the various abstraction levels are considered relevant, depending on the type of building, as demonstrated by [12]. Regarding *Adaptive Zoom*, the results of the questionnaire indicate that users prefer highly abstract representations for a landmark when features are not perceptible by the human eye, i.e. the scale of the map is too small. On the other hand, an abstract representation may not be enough, especially in locator tasks, when users are trying to compare a close view of a landmark, which is right in front of them. In such case a more detailed representation (ex. a 3D building, an image or drawing) is preferred. These circumstances indicate the need of continuously adapting the current landmark abstraction level as a function of the zoom level. Regarding *Adaptive Complexity* and looking at a typical view of a map provided by Google Earth there is a huge amount of icons and signs overlapping mutually and filling the whole map. Because of this it can be easily argued that the task of decoding all the represented information is very difficult and time-consuming. *Adaptive Complexity* is the key to solve this problem, by grouping / merging similar symbols (see Fig. 10).

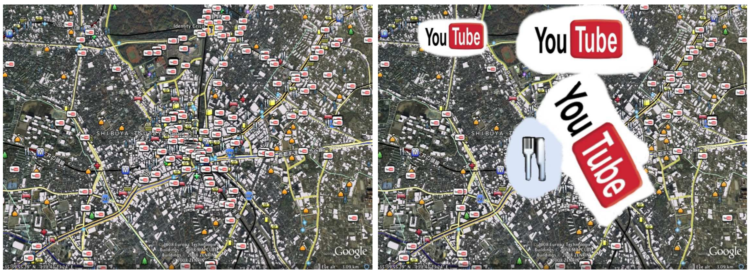

Fig. 10. Adaptive Complexity

5.5 Contextual Awareness

The proposed orientations for *Contextual Awareness* (*Reconstructional*, *Recreational*, and *Fictional*) should be chosen according to the context and scope of the mobile application. For instance, a mobile guide that aims to rebuild the ancient city of Atlantis will provide a strong reconstructional approach rather than the other two orientations, because it is related to the scope of the application. Similar examples can be given for the other two orientations. In terms of magnitudes, there is an *Active Awareness* variable that is common to all mobile map applications (or they wouldnt be called ‘mobile’), i.e. the location context. Apart from this, the other contexts can be ‘processed’ in a passive (when the user asks for more information) or in an active way (i.e., automatically) by adapting to the detected context. In some studies it is argued that an *Active Awareness* approach can become very obtrusive, especially when the detected context does not correspond to the real context and, thus, opposes the users expectations. Nevertheless, it is believed that every mobile map application should leverage the awareness levels appropriately, depending on whether the information is considered relevant or not, and hopefully provide means to toggle back and forth between both *Active Awareness* and *Passive Awareness* modes.

6 A Visualization Paradigm for 3D Map-Based Mobile Services

The analysis in the previous section delivered a set of recommendations for each feature vector, as summarized in Table 1. Based on these analysis, we developed a visualization paradigm for 3D map-based mobile services, taking into account the interactions between feature vector components as well as eventual conflicts that may arise.

Table 3. Set of *feature vector* components that individually maximise user experience and performance with mobile 3D maps

Feature Vector	Orientations	Magnitudes
Image Realism	3D Buildings, Map Vectors, Surface Model	Mixed Realism, Photo-Realism
Object Labeling	Perspective-Adaptive, General Positioning	Dynamic Selection and Dynamic Placement
Visual-Spatial Abstraction	Ground Level, Local-Area Level, Wide-Area Level	Adaptive Level, Adaptive Orientation
Route Indication	Arrows, Cords, Carpet, Way Points	Simulative
Landmark Symbology	Shops (referenced by name), Shops (referenced by type), Buildings (with unique name / function), Buildings (with unique visual properties)	Abstractness (Words, Sign, Icon, Sketch, Drawing, Image), Adaptive Zoom, Adaptive Complexity
Contextual Awareness	Reconstructional, Recreational, Fictional	Active Awareness, Passive Awareness

We present this paradigm as a dual specification:

- Visualization layers that can be turned on or off in order to add or remove features to the image being visualized;
- Visualization functions that should be provided to improve the effectiveness of the user interaction.

6.1 Visualization Layers

The first part of the visual paradigm specification is outlined in this section, regarding the multiple layers of visualization elements. These layers, as the name implies, can be turned on or off to compose the final image presented on a device. We have devised the following layers:

Regular Buildings Buildings are one of the most important layers of a map, since they often correspond to the source and destination in a navigation task. For a better user responsiveness in identification tasks, the buildings should be depicted as *Photo Textured Buildings* or, if not possible, *Simple Textured Buildings*, according to Sect. 5.1. Following what was stated in the same section, the questionnaire proved that by providing the third dimension (even with simple textured or colored façades) there is an improvement on the user performance. However there are cases where 3D buildings can cause occlusions along the route, making the navigation task more difficult. For instance, in a very pessimistic case, a user can be driving a car in a narrow one-way street, with lots of buildings on both sides, and approaching a maneuver where he/she has to perform a 90-degree turn (or more, for even worse cases). If the occlusions are not taken into account, it will very difficult to visualize the route. One solution for this can include rendering buildings near a maneuver point, in a translucent way, so that the route becomes visible, while keeping the presence of such buildings perceptible to the human eye.

Roads and Polygonal Features Roads are a layer of major importance for a mobile 3D map, especially for car driving assistance. Other polygonal features, comprising urban, water and vegetation features, often represented by polygons, allow us to identify and to recognize the surrounding environment, while comparing it to the image presented on a device. For these two elements, and following the results in Sect. 5.1, the best approach consists on using aerial image component, since it allows a faster and more reliable identification of the ground features that surrounds the users. However, when using the *Orthophotomap* component, it may be more difficult to visualize the road network. One possible solution consists on coloring the vectors with a translucent color, over the aerial images. Following Sect. 5.2, the labels of the roads should be, as much as possible, oriented towards the camera (*Perspective-Adaptive Labeling*) for higher readability, so it makes no difference if we are reading labels in the horizon or the label of the current street. Roads and polygonal features should be labeled using a *General Positioning* approach, i.e., for roads, a *Line Positioning* approach is the most indicated, because a road is a linear feature; while polygons, like countries, cities or wards are better labeled with an *Area Positioning* strategy. Some problems may affect readability, when displaying a label of a street along its line, depending on the angle used to represent the road vector. If labels are considered difficult to read in such circumstances, it may be preferable to adjust its rotation angle a bit different from the road vector angle, so this problem disappears.

Rivers and Mountains Results pointed out in Sect. 5.1, following the questionnaire, point to the use of a *Terrain Model*, which should be combined with the aerial image, for a higher degree of *Photo-Realism*. When this layer is turned on, it will be possible to see a clear distinction between the altitude of rivers and the altitude of the peaks of mountains. In certain circumstances, especially when the user is in a mountain place, many occlusions may block him/her from observing whats ahead. Because of these restrictions, the ability to turn on and off the *Terrain Model* component may become an advantage for the user.

Points of Interest Points of Interest, also known as landmarks, refer to buildings with historical/cultural/social significance or other prominent objects in a given landscape. Users must be able to find all kinds of Points of Interest, including shops referenced by name and/or type, and buildings with unique name / function / visual properties. An extra advantage is that they work, many times, as anchor points to the surrounding area. For many specifications of a mobile map, it may be difficult to provide all the possible abstraction levels, as pointed out in Sect. 5.4. Nevertheless, at least 2 or 3 levels of abstraction should be used, from the most abstract to the most concrete:

1. *Sign*, *Icon* or *Sketch* for the most abstract representations;
2. *Drawing* or *Image* for the most concrete representations;
3. *Words* as an addition, especially in the case of landmarks referenced by their unique names.

However, when it is not possible to model all landmarks with an *Image symbol*, then it should be given particular attention to modeling at least the most famous/'visually significant' buildings and monuments. By modeling these buildings instead of less important ones, the probability of providing guidance for a random user will be maximized.

6.2 Visualization Functions

The second part of the specification addresses the high-level visualization functions that should be considered as a requirement for the great majority of 3D map-based mobile services, regardless of the layers involved:

Panning the Map It should be possible to pan the map around and see the labels in a continuous way (*Dynamic Placement*), as stated in Sect. 5.2. The user should not be confused with visual artifacts of labels popping in or out, especially in the limits of the screen.

Zooming the Map Under monotonic zoom, the labels of the map should appear also in a continuous way (*Dynamic Selection*). This implies that, for example, when zooming in, labels that are visible should not suddenly "vanish" but instead slide out of the view area [9]. The same reasoning can be used in the opposite circumstance. The size of labels should be decided as a function of zoom level. For example, when a label of a city occupies the entire polygon that defines its borders, the name should be reduced when zooming out, or enlarged when zooming in, until it is not visible for a given zoom level.

Showing Points of Interest When showing POIs, some care should be taken regarding the choice of the represented abstraction level. It is believed that when a landmark is far from the camera, an abstract symbol such as a Sign or an Icon should be used, rather than an Image. The opposite is also true, i.e., when the camera is close to the landmark, a more concrete representation like an Image should be used instead. In general, the abstraction level for a landmark representation should be higher, when its visual features are not visible enough to allow an appropriate recognition; and lower, when the addition of visible visual features could improve identification of nearby landmarks. The *Adaptive Zoom* component can help the visualization paradigm selecting the appropriate abstraction level of a landmark as a function of zoom, as seen in Sect. 5.4. At far distances, there are cases where the number of overlapping POIs becomes so high that is impossible to recognize what is where. For this reason, the visualization paradigm should be capable of aggregating close or overlapping POIs into abstract clouds representative of the group of POI s that were aggregated. This “clouds” will be broken into smaller clouds as we zoom in the map, since the number of POI s will be reduced along with the size of the view area.

Navigation *Active Awareness* approach should be used, especially in an automotive navigation experience, to provide users route indication instructions, as stated in Sect. 5.5. The main reason is that the user is concentrated driving and paying attention to the road, and thus not as willing to receive turn-by-turn instructions in a passive way as if they were a passengers. However, in some situations a *Passive Awareness* approach can be used more adequately. In an ancient part of a city, the user can indicate the desire to observe the same city as if it was 100 years before. In terms of *Route Indication*, and following Sect. 5.3, it can be argued that a *Cord* or a *Carpet* may not provide enough visual information, when used alone. Since it is a continuous visual cue, it does not give a clear indication for maneuvers, i.e. if you are driving a car you must know precisely if you are turning or changing direction in an intersection. Thus, for navigation tasks, other visual indicators, like arrows, should be used to describe discreet events such as maneuvers. When coming closer to a major interchange, drivers will be able to read (on the device) life-like signposts with destinations written on them, along with maneuver indicators provided by a *Simulative Route Indication* approach, allowing them to identify much faster and accurately the correct way to follow. The previous example also introduces another issue, which was addressed in Sect. 5.2: when a user is driving fast, it should be possible to see “more ahead”, in order to properly anticipate maneuvers. On the other hand, in a pedestrian navigation, users will be interested: in a *Ground Level* perspective that allows an easy identification of the buildings around them; in a *Wide-Area* perspective for acquiring an overview or planning their itinerary; and in an intermediate *Local-Area* Level perspective for confirming the presence of a landmark in the surroundings. Ideally, the map should perform this change automatically, i.e., it should provide an *Adaptive Level* camera. In a pedestrian navigation experience, it should be of extreme importance that the visualization adapts to match the same orientation than in reality, i.e., that it supports *Adaptive Orientation*. For instance, if a user is standing still, but looking to the buildings around him/herself, the visualization should reflect what in reality the user is facing to. In the end, the user will have fewer difficulties matching the perspective of both realities, when identifying the buildings around.

7 Conclusions and Future Work

In this study, a generic *Evaluation Framework* was proposed as the main methodology for the specification, development and evaluation of new or existing solutions in the 3D map visualization problem domain for LBMS. *Feature Vectors* can individually describe a set of choices (*orientations*) and degrees of applicability (*magnitudes*). The proposed framework focuses on 6 *feature vectors* namely, *Image Realism*, *Object Labeling*, *Visual-Spatial Abstraction*, *Route Indication*, *Landmark Symbology*, and finally *Contextual Awareness*. These feature vectors encompass the most relevant visualization issues in 3D maps on LBMS, but there was no intent to cover them completely. A future line of research would consist in analyzing the totality of features that address visualization aspects, in the context of exploration of urban environments, using 3D LBMS as guidance. Although the state of the art contemplates some of the issues involved, the questionnaire gave a much more clear insight on them. In general, it is observed a greater tendency towards the need of *Image Realism* rather than *Image Functionalism*. In terms of *Perspective-Adaptive Labeling*, it was proved that users are at disadvantage, if they are given the task to read labels of a map, when these labels are not oriented towards the camera's viewing direction. The results also demonstrated that users can more easily identify the presence of a distant landmark with an abstract representation, and a close landmark with a concrete representation, which is indicative of the need of an *Adaptive-Zoom* behavior. Since there are several limitations on the kind of measurements that can be performed with the proposed questionnaire in order to evaluate *feature vectors*, it would be interesting to perform other kinds of tests, with particular focus on dynamic experiments, to get more information about other vectors such as *Visual-Spatial Abstraction* and *Contextual Awareness* which were not evaluated. An example of these experiments would include using a driving simulator to test the participants' reflexes, given a situation where they are approaching a maneuver, and deciding which way to go. We have presented a visual paradigm to support the development of 3D map-based mobile services. This specification contemplates visual contents in the form of visualization layers and required functionality as visualization functions. This definition was based on a comprehensive analysis on current state of the art and on an experiment made by using an online questionnaire. This analysis was directed by a conceptual framework developed to focus the study on the most relevant visualization elements (*feature vectors*) that influence user performance and experience. Given that the proposed visualization paradigm still has a lot of empirical knowledge within itself, the future research would focus on studying the interactions between feature vectors, to understand eventual conflicts that may arise, and how can they be combined to maximize usability and user experience.

References

1. Rademacher, P., Lengyel, J., Cutrell, E., Whitted, T.: Measuring the perception of visual realism in images. In: Proceedings of the 12th Eurographics Workshop on Rendering Techniques, London, UK, Springer-Verlag (2001) 235–248
2. Lange, E., Ch, Z.R.: The degree of realism of GIS-based virtual landscapes: Implications for spatial planning. In: D. Fritsch and R. Spiller (eds) Photogrammetric Week '99. (November 28 2003) 367–374
3. McNamara, A., Chalmers, A., Trocianko, T.: Visual perception in realistic image synthesis. In Coquillart, S., Duke, D., eds.: STAR Proceedings of Eurographics 2000, Interlaken, Switzerland, Eurographics Association (August 2000)

4. Bolin, M.R., Meyer, G.W.: A visual difference metric for realistic image synthesis. In: Proc. SPIE. (1999) 106–120
5. Ferwerda, J.A.: Three varieties of realism in computer graphics. In: In Proceedings SPIE Human Vision and Electronic Imaging ’03. (2003) 290–297
6. Milgram, P., Kishino, F.: A taxonomy of mixed reality visual displays. IEICE Transactions on Information Systems **E77-D**(12) (December 1994) 1321–1329
7. Borenstein, J., Everett, H.R., Feng, L.: ”Where am I?” – Sensors and Methods for Mobile Robot Positioning. The University of Michigan (April 1996)
8. Hunolstein, S.V., Zipf, A.: Towards task oriented map-based mobile guides. In: Workshop ”HCI in Mobile Guides” at Mobile HCI 2003. 5th International Symp. on HCI with Mobile Devices and Services. (July 29 2003)
9. Been, K., Daiches, E., Yap, C.: Dynamic map labeling. IEEE Transactions on Visualization and Computer Graphics **12**(5) (2006) 773–780
10. Wolff, A.: Automated label placement in theory and practice. In: PhD thesis, Freie Universität. (1999)
11. van Dijk, S., van Kreveld, M., Strijk, T., Wolff, A.: Towards an evaluation of quality for label placement methods. In: Proc. 19th Internat. Cartographic Conf. (ICC’99), Ottawa, Canada, Internat. Cartographic Association (14–21 August 1999) 905–913
12. Elias, B., Paelke, V., Kuhnt, S.: Concepts for the cartographic visualization of landmarks. In: Proceedings of Symposium 2005 Location Based Services & TeleCartography. (2005) 11
13. Coelho, A.: Expeditious Modelling of Virtual Urban Environments based on Interoperability and Geospatial Awareness (in Portuguese). PhD thesis, Faculdade de Engenharia da Universidade do Porto (2006)
14. Burigat, S., Chittaro, L.: Location-aware visualization of vrml models in gps-based mobile guides. In John, N.W., Ressler, S., Chittaro, L., Duce, D.A., eds.: Web3D, ACM (2005) 57–64
15. Chen, G., Kotz, D.: A survey of context-aware mobile computing research. Technical Report TR2000-381, Dept. of Computer Science, Dartmouth College (November 2000)
16. Nurminen, A.: The m-loma mobile 3d map project website. http://www.init.hut.fi/research%26projects/m-loma/ (2006) Last Checked: November, 2008.
17. Narzt, W., Pomberger, G., Ferscha, A., Kolb, D., Muller, R., Wieghardt, J., Hortner, H., Lindinger, C.: Augmented reality navigation systems. Universal Access in the Information Society **4**(3) (2006) 177–187
18. Oulasvirta, A., Nurminen, A., Nivala, A.M.: Interacting with 3d and 2d mobile maps: An exploratory study. Technical report, Helsinki Institute for Information Technology (HIIT); Finnish Geodetic Institute (April 2007)