GeoPointing: Evaluating the Performance of an Orientation Aware Location Based Service under Real-World Conditions

Rainer Simon, Peter Fröhlich
Telecommunications Research Center Vienna, Donau-City-Str. 1
A-1220 Vienna, Austria
{simon, froehlich}@ftw.at

The growing trend towards mobile phones with integrated GPS clearly suggests that navigation and location based services will be key applications for future mobile devices. New hardware features that are currently finding their way into state-of-the-art phones – such as digital compasses and tilt sensors – promise to drive the adoption of mobile geospatial services and to change the way people navigate, explore and interact with their physical environment: Location based applications that exploit attitude information to realize *orientation aware* interaction have been discussed in research for several years. Yet, little actual results on the achievable real-world performance of such systems exists in literature. In this paper, we report on a series of function trials carried out with a prototype *Geo-Wand* – a portable system that allows users to access geo-referenced information by physically pointing towards objects in the real world. The application was realized with a mass market mobile phone connected to a Bluetooth GPS and a custom-built orientation sensor module. We present test results for multiple types of urban terrain and discuss the possibilities and limitations of this next-generation mobile location based service technology.

1 Introduction

While the success of the early location based services (LBS) in the late 1990s has been rather disappointing, and user uptake has been less than anticipated, mobile geospatial applications like maps, driving directions or local search are now again being marketed as prominent features of top-of-the-line phone models such as the iPhone or the Nokia N95. The growing penetration of GPS in more consumer-oriented phones – combined with the increased availability of affordable tariffs for mobile data and Internet services – has the potential to create new momentum for LBS, and to promote new types of location-enhanced applications such as social and community LBS[1,2], or location aware games[3,4].

Recently, new types of mobile phones are appearing on the market that promise to add further to the attractiveness of future LBS: Equipped with digital compass or tilt sensors, these devices enable *orientation aware interaction*. Concepts for how such sensors can be used to enhance location based applications have been investigated in research for several years. For example, innovative interaction metaphors like the *Geo-Wand* – a portable system that allow users to identify physical objects by pointing towards them – have been proposed (Egenhofer 1999); prototype implementations have been used to demonstrated the use of orientation awareness in a variety of contexts, such as in map-based pedestrian navigation (Wasinger *et al.* 2003); as

[1] loopt (http://www.loopt.com/)
[2] Mologogo (http://mologogo.com/)
[3] The Shroud (http://www.shroudgame.com/)
[4] Locomatrix (http://www.futureplatforms.com/fp/clients/locomatrix/gps_gaming/)

interaction technique in mobile multiplayer gaming (Mitchell *et al.* 2003); or in audio-assisted pedestrian navigation (Strachan *et al.* 2005).

Despite years of research, prototypes, and first commercialization activities[5,6], little work can be found in literature that actually goes into the details of the achievable performance of such systems under real-world limitations like GPS inaccuracy and compass error caused by external factors such as electromagnetic interference (Simon *et al.* 2005). In this paper, we present a series of tests conducted with a prototype *Geo-Wand*, based on a mass market mobile phone equipped with custom orientation-sensors. The remainder of this paper is structured as follows: Section 2 describes the setup that was used to conduct the tests. Section 3 introduces a new test methodology we developed to evaluate the functionality of the *Geo-Wand*. The methodology consists of a two stage process: a continuous GPS test over a longer test route, and a series of "pointing samples", where GPS and compass were used in conjunction. Sections 4 and 5 discuss the results of the GPS test and the pointing test, respectively. Section 6 presents some instructive examples for typical error situations that were observed during the tests. Section 7 concludes the paper and proposes directions for future work.

2 Point to Discover

The work presented in this paper is part of the *Point to Discover*[7] research project, which investigates technical and design issues related to the development of orientation aware location based applications. The contributions of the project so far include a server-side framework for directional spatial queries in an environment block model, a novel XML query result format, and a toolkit for prototyping different types of orientation aware mobile user interfaces (Simon and Fröhlich 2006). Due to the novelty of mobile devices that combine integrated GPS and compass, no suitable commercial test device was available over the duration of the project. Custom hardware prototypes were therefore developed for test purposes.

Figure 1. *Point to Discover* device prototype.

Figure 1 shows the prototype that was used for the tests presented in this paper: The setup consists of a custom-made plastic shell that snaps onto the back of a standard, mass market

---

[5] Intelligent Spatial Technologies (http://www.i-spatialtech.com/)

[6] GeoVector (http://www.geovector.com/)

[7] http://p2d.ftw.at

mobile phone. Mounted inside the shell is a digital compass and 2-axis tilt sensor module purchased from a commercial vendor. The module is connected to a standalone power supply and a Bluetooth transceiver chip that transmits the sensor measurements to the phone wirelessly using a simple, text-based protocol. The prototype does not include an integrated GPS receiver; an external Bluetooth GPS is therefore required in addition to the setup shown in Figure 1. A Java 2 Micro Edition (J2ME) library for decoding the sensor data stream, as well as the message stream of a standards-compliant GPS unit is part of the *Point to Discover* prototyping toolkit.

3 Test Design

Obviously, no standardized methodology exists for the evaluation of orientation aware mobile applications. We therefore developed our own test procedure for the *Geo-Wand*. Since we expected that both GPS error and compass error play an equally important role for the performance and accuracy of the system, we designed a two-phase evaluation process: In the first phase, only the quality of the GPS was assessed in isolation. This allowed us to collect continuous measurements over a longer period of time, across a larger test area. The goal of this test phase was to gain insight into the typical GPS errors of our setup that can be expected under different environment conditions. In the second test phase, the pointing functionality was investigated: using GPS and compass in conjunction, samples were collected at a finite number of test locations. The goal of this test phase was to get an approximate value for the "pointing success rate" in different environments (i.e. the percentage of cases where the system would accurately select the object the user physically pointed at in the real world, versus the cases where the system selected a false target due to GPS and/or sensor errors).

By nature, neither GPS nor compass tests are fully reproducible. Factors like non-repeatable GPS satellite constellations, changes of the environment over time (e.g. seasonally changing density of foliage, changes due to removal or addition of buildings, etc.), or other temporary external influences (e.g. temporary electromagnetic interference affecting the compass) prohibit an absolute truth-referencing. The authors of (Aloi *et al.* 2007) discuss several of these challenges with regard to GPS tests. As the simplest possible validation method, they suggest plotting the receiver's position estimates against a map: While this method does not provide quantifiable statistics, it does capture large outliers and provides a qualitative insight in the performance of the tested system. Our evaluation method is therefore based on a comparison between the recorded measurements and a reference map.

3.1 Test Environments

In the automotive telematics industry, three types of terrain have evolved as commonly accepted test environments for GPS tests: *open highway*, *foliage* and *urban canyon* (Aloi *et al.* 2007). Pedestrian use of GPS is, however, different from GPS use in car-navigation: users normally move on sidewalks close to the building line; they may move in and out of buildings frequently; and they cannot necessarily be expected to carry their GPS unit or GPS-enabled mobile device in a place that ensures ideal reception; rather they might carry the unit stowed away in a backpack most of the time, or in a pocket of their shirt, trousers or jacket. Due to these differences, we considered it necessary to refine the terrain type definitions for the pedestrian use case, towards a more fine grained classification scheme. Based on the automotive test environments, we therefore define the following five environment types: *open environment*, *low density urban* (*suburban*) *environment*, *park environment*, *urban environment*, and *urban canyon environment*. Due to the fact that no two outdoor environments are the same, it is difficult to specify hard, quantitative thresholds for the properties of each environment. The following paragraphs attempt to describe each environment type in a qualitative manner, through representative examples.

Open Environment. The *open highway* driving environment known in the automotive industry refers to terrain with no or minimum obstruction of the clear view of the sky and only short, momentary outages, e.g. caused by driving underneath overpasses, in short tunnels or passing by large trucks (Aloi *et al.* 2007). In accordance to this definition, we define the *open environment* type as an environment with analogous properties. Examples of this terrain type are rural areas, with no or few buildings and foliage, spaced far enough away from the user to not cause any considerable obstructions of the sky.

Low Density Urban or Suburban Environment. This environment is characterized by a low percentage of obstructed sky. Examples are suburban areas with low (2 to 3 floor) buildings, or open spaces in inner city areas, surrounded by medium-sized buildings (up to 6 floors) but spaced sufficiently far away from the user to obstruct only a small portion of the sky. Typical examples are suburban residential zones, or medium to large inner city squares.

Park Environment. In reference to the *foliage* driving environment, we define the *park environment* as an environment that shares the properties of the *low density urban environment*, but with the addition of close-by trees and overhead foliage. Wet foliage, in particular, is known to cause more severe attenuation effects on GPS satellite signals than dry foliage. Furthermore, since the density of the foliage affects the amount of signal attenuation that occurs, attenuation in this environment can be expected to be a seasonal effect.

Urban Environment. This type of environment is more densely developed than *low density urban* environment, with medium-sized buildings (up to 6 floors) spaced closely together, narrow streets, pedestrian zones and alleyways. Examples are downtown areas and historical city centers found in typical Western- and Central-European cities like Paris, Vienna or Prague.

Urban Canyon Environment. The *urban canyon* driving environment type known in the automotive industry is modelled according to the densely developed downtown environments of North American cities such as Chicago, Los Angeles or New York. Characterized by high-rise buildings that obstruct a high percentage of the sky, this type of terrain is the most challenging environment for GPS receivers: multipath propagation and frequent loss of line-of-sight limits the number of satellites that can be tracked, leading to reduced accuracy of the location estimate and outages where the computation of a position is not possible at all.

Figure 2 shows photographs of some of the areas where tests were carried out. Each photo depicts one representative example of the following environment types: *low-density urban environment*, *park environment*, *urban environment* and *urban canyon environment.*

Figure 2. Representative areas for *low-density urban environment* (top left), *park environment* (top right), *urban environment* (bottom left) and *urban canyon environment* (bottom right).

We did not conduct any tests in *open environment*: since this type of terrain provides optimum conditions that do not stress a GPS receiver's tracking capabilities, evaluating GPS performance here is least insightful (Aloi *et al.* 2007). It can also be expected that open environment is the least problematic environment for the compass, as external electromagnetic interference will be mostly absent. It is therefore reasonable to expect that the *Geo-Wand* will operate at least at the same level of performance as in *low-density urban environment*.

3.2 GPS Test Procedure

The GPS tests were carried out with a customary Bluetooth GPS unit, based on a state of the art receiver chipset[8]. The unit was used as is, without an external antenna, and carried in the front pocket of the test person's shirt. A mobile phone was used to record position fixes to a log file at two second intervals, while the test person walked along a predefined test path. In addition to longitude and latitude values, the logging application also recorded the number of satellites the receiver used to compute the position fix, as well as the horizontal dilution of precision (HDOP) value (a quality measure reported by the GPS unit that is derived from the geometry of the current satellite constellation, and which indicates whether the constellation is favourable for computing good quality position fixes). The test was repeated three times, at a four hour interval. Since the duration of a GPS satellite orbit is slightly below twelve hours, three tests spaced at four hours can be expected to ensure different satellite constellations for each test.
As a first test path, a circular route with a length of approx. 5.2 km was chosen in the inner city

---

[8] GlobalSat BT-338 SiRF Star III (http://www.globalsat.com.tw/eng/product_detail_00000039.htm)

district of Vienna, Austria (see Figure 3 left). The route was selected such that it lead through the three environment types of *low density urban*, *park*, and *urban environment* at about equal shares. The route took the test person between 47 and 52 minutes to complete (49 min, 47 min, and 52 min, respectively). Due to the absence of distinctive *urban canyon* environment in the inner city of Vienna, a second test path of about 1.4 km length was defined in a business district in the North of Vienna (see Figure 3 right). This route took the test person 12 to 13 minutes to complete (12, 13 and 13 min respectively).

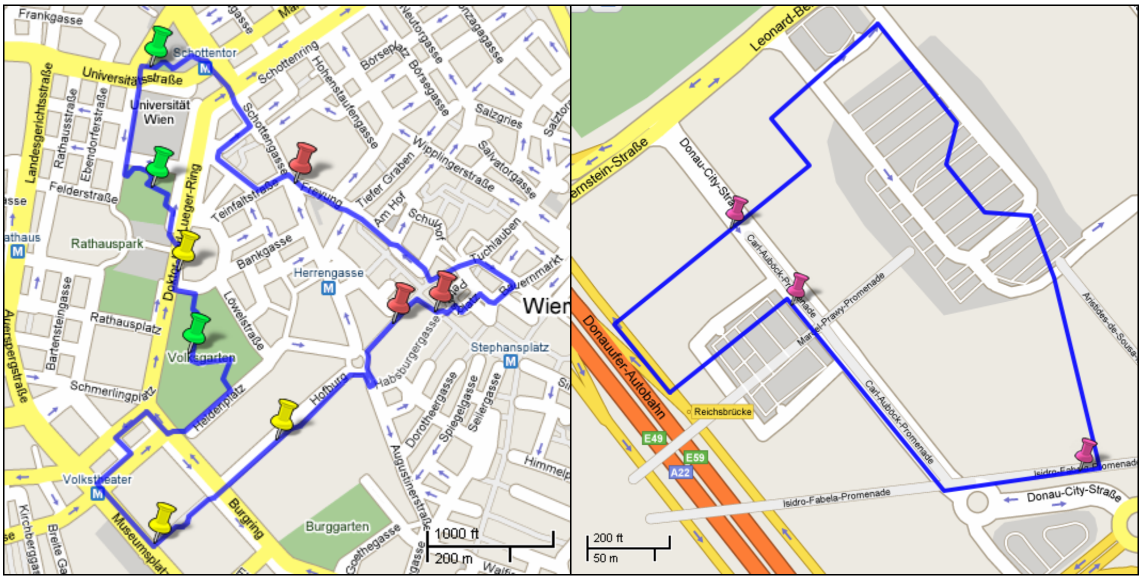

Figure 3. GPS test routes and pointing test locations (left: inner city, right: urban canyon)

3.3 Pointing Test Procedure

The pointing test was conducted with the same Bluetooth GPS unit (again carried in the test user's shirt pocket) and the device prototype described in Section 2. A total of twelve locations, situated along the test routes from the GPS tests, were chosen as representative test areas, with three locations in each type of environment; i.e. 9 of the test points were located on the inner city test route, and 3 test points were located along the urban canyon test route (see Figure 3).

At each test location, four to nine buildings were designated as 'pointing targets': During the test, the test person would stand stationary on the test location, and point the device at the designated target buildings in a pre-defined order. Each pointing gesture was confirmed by a button click, causing the system to log the current geo-coordinates measured by the GPS and the heading and pitch values measured by the orientation sensors to a file for later analysis. In total, 80 pointing targets were defined: 17 in *low-density urban* environment, 17 in *park environment*, 23 in *urban environment*, and 23 in *urban canyon environment*. As in the GPS test phase, the pointing tests were repeated three times, spaced at four hour intervals.

4 GPS Test Results

The recorded GPS tracks were analyzed according to their approximate deviation from the reference path on the map. As a simple measure for the deviation, the normal distance to the reference path was computed for each position fix. Since this distance is not necessarily exactly the same as the distance between the test person's real location and the position fix, and since the reference path itself can not be assumed to be 100% accurate, the deviation measure can not be taken as an absolute statistic, but rather a qualitative indication of the GPS fix quality.

Figure 4 and Figure 5 show the recordings from each of the three GPS tests conducted along the two test routes, along with the HDOP value reported by the GPS receiver. A marker in the track image indicates the start location; the inner city route was traversed in clockwise direction, the urban canyon route was traversed in counter-clockwise direction. Additional markers on the image of the inner city track, which relate to the different types of environment, allow a rough visual mapping between track and deviation plot. Throughout all tests, 6 to 9 satellites were tracked by the receiver at all times (with the exception of an underpass on the inner city test path, where the number of tracked satellites would temporarily fall to 3 or less, as discussed below).

4.1 Inner City Route

As shown in Figure 4, the recordings on the inner city route show largely corresponding results: The route starts in densely developed *urban environment*. The narrowest alleyways are located on the right-most portion of the path (see Figure 4, marker 'URBAN'), with deviations in this area in the range of at least 20 meters or more. After traversing through more urban terrain for approx. 200 meters, the route passes through an underpass. This can be best observed in the third test, where a singular spike of approx. 300 meters was recorded due to temporary loss of GPS line of sight. The route then enters a large square (i.e. *low-density urban environment*; see Figure 4, marker 'L.-D. URBAN'). Deviation from the reference path drops considerably, below 10 meters or less. After a short portion through more densely developed terrain, the route enters *park environment*. Deviation from the path is in the range of 10 meters in the first park (see Figure 4, marker 'PARK 1'), and slightly higher in the second park (around 20 meters, see Figure 4, marker 'PARK 2'). (We attribute this difference to the fact that in the second park, the path passed underneath denser foliage than in the first park.)

Summarizing, we conclude that in *urban environment*, it is reasonable to expect an error in the range of at least 30 to 40 meters with a state of the art handheld GPS receiver. In *low-density urban environment*, our tests showed consistently good performance with maximum deviations from the path in the range of 10 meters or less in all tests. In the *park environment* portions, path deviation was shown to be similar or slightly worse than in the *low-density urban environment* case, depending on the density of foliage.

A further effect that can be observed in Figure 4 is that GPS inaccuracy was most pronounced in test 3: Most likely, this results from a less favourable GPS constellation at the time of the test. As can be seen in Figure 4, however, there is no noticeable correlation between the HDOP value and the approximate deviation during the tests; the HDOP is only slightly higher than in the first two tests. We therefore assume that multipath effects were a primary source of error in the *urban environment* parts of the route, rather than loss of satellite signals. Hence, the HDOP alone can not necessarily be assumed to be a reliable quality indicator in this type of terrain.

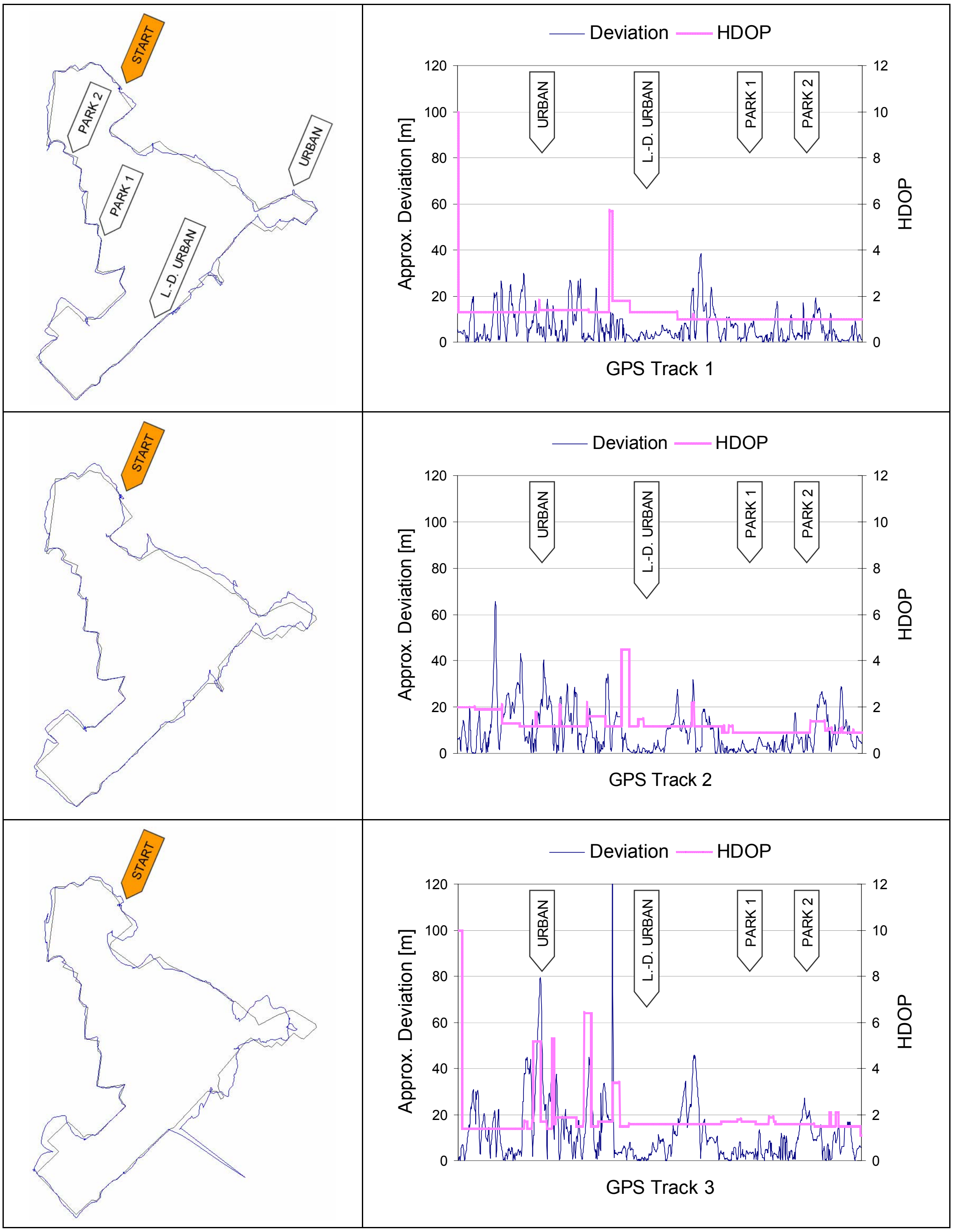

Figure 4. Approx. normal distance from map reference path and HDOP (inner city tests)

4.2 Urban Canyon Route

Due to the more homogenous environment on the *urban canyon* route, the results recorded on this test path were more uniform: Figure 5 shows the results from the three tests. The approximate deviation from the reference path was measured to be in the range of at least 20 to 40 meters and more.

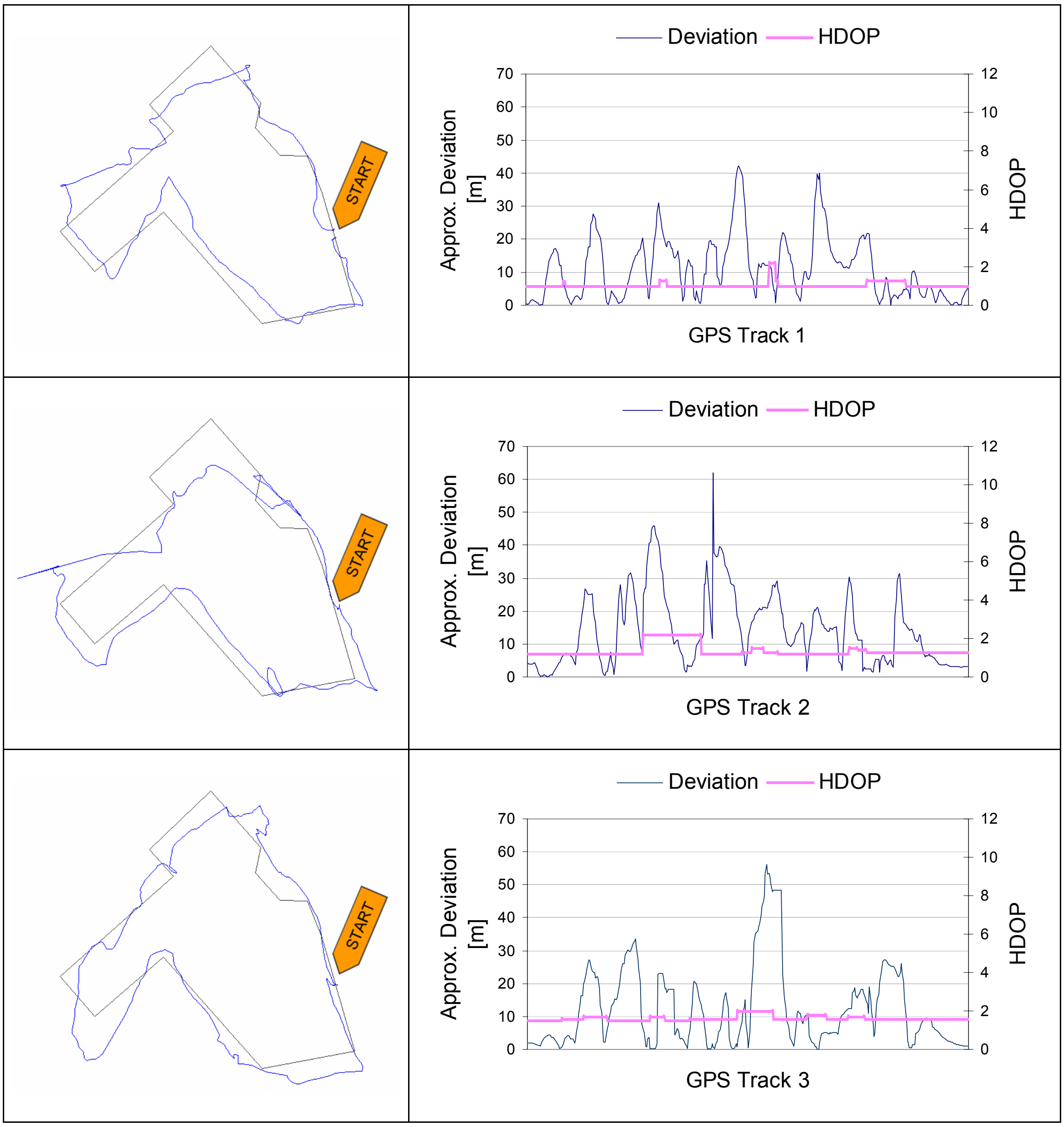

Figure 5. Approx. normal distance from map reference path and HDOP (urban canyon tests)

5 Pointing Test Results

The results of the pointing tests were analyzed according to their ‘pointing success rate’, i.e. the cases where the server-side query engine would select the correct target building from the GPS and compass measurements vs. the situations where the measurements lead the query engine to select a wrong building as the pointing target, or no target at all. In a second analysis, we also allowed for *partial hits*: pointing queries that missed the target by not more than 5 degrees were also counted as successful hits (in this case, the result returned by the query engine is typically a set of targets, rather than a single target). Figure 6 shows typical examples for a pointing success and a *partial* pointing success, as recorded in urban environment.

In total, the test yielded an overall pointing success rate of 71.25% (3 tests with 80 targets each, and a total of 171 successful pointing queries). In the case where *partial hits* were also counted as hits, the overall pointing success rate increased to 83.75% (201 successful pointing queries). Table 1 shows the detailed results for each of the three tests, with results subdivided by environment type.

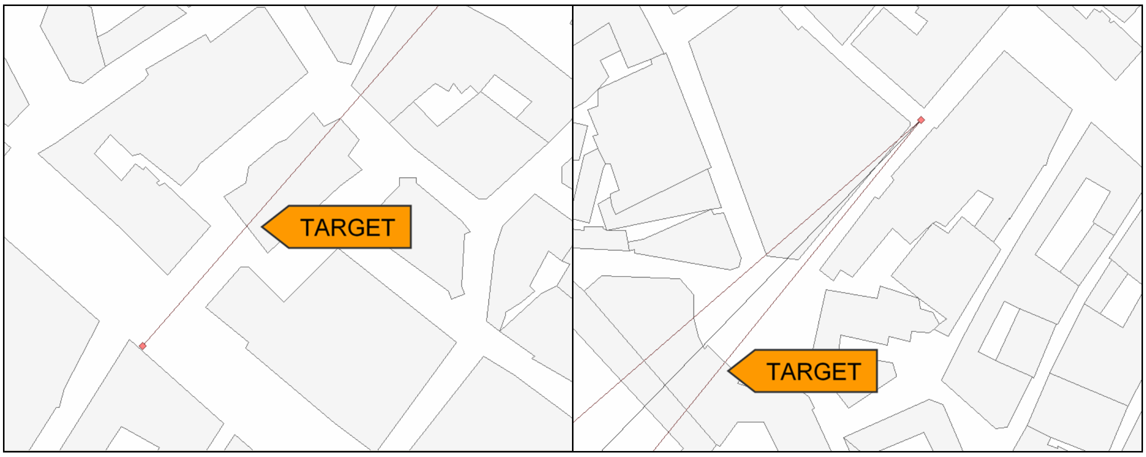

Figure 6. Pointing query examples: successful query (left) and partially successful query (right)

As can be seen, the measured pointing success rate approaches 100% in *low-density urban environment*, and is slightly decreased in the *park environment*. Two more results are worth pointing out: First, there is a high variation between the results from three tests in urban environment (60.87% in test 1 vs. 30.43% in test 3). The increase of the pointing error rate with each test coincides with the decrease of GPS accuracy that was measured in the GPS tests. In fact, we argue that GPS performance is more limiting to the overall performance of the system than compass performance, as will be discussed by several examples in the following section.

Table 1. Pointing success rate by test repetition and environment categories

	Test 1		Test 2		Test 3		TOTAL	
		with partials		with partials		with partials		with partials
Low-Density Urban	88.24%	94.12%	100%	100%	100%	100%	96.08%	98.04%
Park	88.24%	100%	82.35%	94.12%	76.47%	88.24%	82.35%	94.12%
Urban	60.87%	65.22%	56.52%	91.30%	30.43%	34.78%	49.28%	63.77%
Urban Canyon	69.57%	86.96%	65.22%	82.61%	65.22%	86.96%	66.67%	85.51%

The second noteworthy effect is that our system performed considerably better in the *urban canyon environment* than in the (seemingly less problematic) *urban environment*. We attribute this result to two possible causes: First, the area chosen for the test (despite being the only suitable area available to us) was still relatively open, with wide spaces in between high rise buildings. Therefore, it might not fully qualify for a typical *urban canyon* environment. In fact, our GPS tests confirmed that GPS accuracy was not much below the *urban environment* case. As the second cause for the good performance, a closer inspection of the results revealed the following, simple reason: Buildings in the *urban canyon environment* were larger than those in the *urban environment* (e.g. office towers vs. smaller inner city buildings); i.e. the size of the pointing targets was increased, while GPS and compass conditions remained largely the same.

6 Discussion

In order to gain a more detailed insight into the types of errors that lead the query engine to decide incorrectly, we conducted a visual inspection of the pointing test results. The visual inspection was performed using two tools: first, the results were visualized in a custom viewing tool which we developed as part of the *Point to Discover* project, and which produces a top-down view of the pointing logs (compare Figure 6). Secondly, the result data was converted to 3D geometry and imported to *Google Earth*[9]. Figure 7 presents three commonly observed error characteristics. Top and bottom images thereby represent the same situation, as displayed in the custom viewing tool (top) and in *Google Earth* (bottom), respectively.

Figure 7. *Geo-Wand* error examples, from left to right: (a) user placed behind object due to GPS error, (b) missed alleyway due to GPS error, (c) temporary compass offset

In Figure 7a, a GPS error places the user wrongly behind a building by about 20-30m. This leads to a blocked line of sight to most of the pointing targets, even though there is clear line of sight in reality. Figure 7b shows a comparable case, where a GPS error places the user away from a street crossing by approx. 30 meters. This leads to an error when the user points towards a building down the street (see pointing line parallel to the street in S-W direction): the query engine wrongly assumes the building South of the alley to be the target, rather than a building down the alley. These frequently occurring examples show that in densely developed terrain, a relatively typical GPS error (compare Section 4) can, unfortunately, have severe consequences on the pointing success rate and, hence, on the perceived system performance.

Figure 7c shows the case of errors caused by a temporary compass offset: samples from two tests are compared against each other, with samples from test 2 shown in green and samples from test 3 in red. While the samples from test 2 are accurate, the two Northern samples from test 3 are offset by their true direction by about 20 degrees. These effects either occurred locally (due to nearby ferromagnetic material) or temporarily, as in the example above. The origin of

[9] http://earth.google.com

the temporary effects is not yet fully clarified; more research is needed to identify their sources (such as electrical equipment operating nearby) and their range reliably. The example shown above, however, was the only one during the test where compass error was observed with such distinctiveness; in all other samples, GPS error was by far the most dominant error cause.

7 Conclusion

The trials we conducted with an early, improvised prototype clearly show that orientation aware interaction is becoming feasible; though with a few restrictions imposed by the accuracy limits of today's positioning methods and orientation sensors. Some of these limitations will become less severe as technology matures (e.g. as magnetometer compasses get replaced by gyro-compensated compasses that are less susceptible to electromagnetic interference). Other limitations, however, will remain for several years to come: seemingly small positioning errors can have large consequences in densely developed terrain, e.g. when they place the user behind a corner, or away from a street crossing. This will most likely remain a challenge for future systems using differential techniques and next generation satellite positioning systems (such as the European *Galileo* system). As a next step, it will therefore be crucial to develop strategies for how errors situations can be handled: On the one hand, smart "pointing-equivalents" to map matching algorithms are needed that might be able identify more plausible pointing targets than those indicated by the query parameters. On the other hand, user interfaces must be designed so that they efficiently communicate the inherent uncertainty, rather than concealing it.

8 Acknowledgements

The work presented in this paper was funded by mobilkom austria, Siemens Austria and the **K***plus* competence centre programme.

9 References

D. N. Aloi, M. Alsliety, and D. M. Akos (2007) A Methodology for the Evaluation of a GPS Receiver Performance in Telematics Applications, *IEEE Transactions on Instrumentations and Measurement*, vol. 56, issue 1, IEEE.

M. Egenhofer (1999) Spatial Information Appliances: A Next Generation of Geographic Information Systems, *1st Brazilian Workshop on GeoInformatics*, Caminas, Brazil.

K. Mitchell, D. McCaffery, G. Metaxas, J. Finney, S. Schmid, and A. Scott (2003) Six in the City: Introducing Real Tournament – A Mobile Context-Aware Multiplayer Game, *2nd Workshop on Network and System Support of Games*, Redwood City, CA. ACM Press.

R. Simon, H. Kunczier, and H. Anegg (2005) Towards Orientation-Aware Location Based Mobile Services, *3rd Symposium on LBS and TeleCartography*, Vienna, Austria.

R. Simon and P. Fröhlich (2006) A Mobile Application Framework for the Geospatial Web, *16th International World Wide Web Conference (WWW2007)*, Banff, Canada, ACM Press.

S. Strachan, P. Eslambolchilar, R. Murray-Smith (2005) gpsTunes – Controlling Navigation via Audio Feedback, *Proceedings of MobileHCI 2005*, Salzburg, Austria, ACM Press.

R. Wasinger, C. Stahl, and A. Krüger (2003) M3I in a Pedestrian Navigation & Exploration System, *Proceedings of MobileHCI 2003*, Udine, Italy, ACM Press.