Huerta, Schade, Granell (Eds): Connecting a Digital Europe through Location and Place. Proceedings of the AGILE'2014 International Conference on Geographic Information Science, Castellón, June, 3-6, 2014. ISBN: 978-90-816960-4-3

Augmented Reality and GIS: On the Possibilities and Limits of Markerless AR

Falko Schmid and Daniel Langerenken
Cognitive Systems
University of Bremen
Germany
schmid@informatik.uni-bremen.de, daniel.langerenken@gmail.com

Abstract

The application of Augmented Reality (AR) in the geo-spatial domain offers huge potentials: AR can visualize invisible properties of spatial entities, can display historic data for them, or can help in finding places. Whatever the application is, AR in the geo-spatial domain will often be purely sensor based, thus without the help of visual or sensory markers. In this paper we analyse the achievable accuracy of AR projections under everyday conditions with consumer hardware. We can show that AR can be applied in applications in smaller geographic scale, but is not sufficient if it comes to the preciseness required when inspecting infrastructural data of small scale.

Keywords: AR, GIS

1 Introduction

Enriching the direct perceivable environment with complementary information bears great potential in the geo-spatial domain. We can make the invisible visible, we can browse through history and future of a place, we can learn about legal issues, we can assist during navigation, advertise properties, etc. With augmented reality (AR) we can visualize the road to take, underground pipe and cable installations, the type of soil below us, its quality, and contamination with toxic substances. We can learn about archaeological discoveries of filled up digging sites, see the places that have been flooded or will be at a certain water level, or how buildings will look like when they are built.

The possibilities are endless and with the broad availability of sensor-packed devices like smartphones and the advent of data glasses in the end-user market, augmented reality (AR) will be the tool of choice for many of these applications. AR applications can help to make informed decisions, reduce costs, entertain, and assist during spatial tasks. However, this is only possible if the applications can support the required level of accuracy. I.e., accurate projections are required to ensure that projected data corresponds with the entities of the camera image. The level of required accuracy depends on the domain: some applications will be usable even if the results are displaced by 10 meters, others will require a high degree of precision.

Projecting data at the correct camera image technically requires accurate positioning, clear sensory data, and ideally some visual or sensory makers for precise alignment of data in the environment. State-of-the-art techniques ensure accuracies down to millimetre precision, this level of accuracy will be out of reach for the majority of geo-spatial applications for the next years. High precision can be achieved in constrained domains and controlled settings where the system knows about clear markers, visual properties of environments and entities, or has access to precise sensors. Although precision and availability of technology constantly increase positioning and 3D orientation sensing will have limited accuracy in everyday settings and away from lab conditions.

GPS-based positioning with non-survey grade devices is known to be inaccurate, Wi-Fi is and will not be available everywhere in the world, and the environment is constantly changing due to evolution, seasonal features, or events. Landmarks, buildings, signs, trees, and parks appear and disappear. Thus, the available data, which is the potential source for sensory or visual registration methods can differ significantly from reality: the building an algorithm is looking for can be replaced, the street can be covered with snow, and the tree is currently without leafs.

AR literature and its evaluations suggest that that markerless, pure sensor-based AR is not sufficient for applications requiring high precision projections. However, this is certainly true for applications requiring a high degree of precision (e.g., surgical applications) - for other classes of applications the limitations might be acceptable. In this paper we analyse the limits of pure sensor-based, markerless AR under everyday conditions and identify classes of applications suitable for the achievable accuracy.

2 Related Work

During the last years the application context of AR-based applications strongly moved to the direction of the broad mass of users. Due to technically very powerful and affordable smartphones and the possibilities of developing your own mobile applications, more and more applications are published that mix real and virtual environment. Liarokapis et al. justify this by the rise of GIS. Therefore they developed a tangible user interface for visualizing geographical data received by shape files [1]. Another source of geodata is shown by Schmid et al in mapIT [2, 16]. They provided a possibility to gather, annotate and send geodata to a GIS by using camera, sensor- and positioning data of smartphones.

Behringer linked sensor and positioning data with the image of a camera and height maps to register horizontal silhouettes in the viewport. This, however, requires good lighting conditions [3]. Stricker and Kettenbach describe an approach based on markerless, optical tracking. Depending of the current field of view of the camera, a collection of reference pictures is pre-sorted. From these images, the best reference image is calculated and then projected onto the camera image. Though, a known environment is needed to pre-sort a collection of reference pictures [4].

Azuma, Hoff et al. took care of the problem of inaccurate data and therefore developed a motion-stabilized outdoor AR-system. This system stabilizes the received sensor data and attempts to avoid delays by predicting. However, it is subject to some limitations due to the needed equipment. A fixed location is required to stabilize the received data. Changes in the location are not supported [5].

Yi Wu et al. studied the possibilities of outdoor AR in cities under consideration of the position, the orientation of the device and the current camera image. They linked sensor-based AR with natural marker-based AR. A database provided the necessary information for the current GPS position. [6].

For maintenance support Roberts et al. presented an AR-application which allowed to project gas, telephone, water and power lines located behind walls into the environment [7]. A similar approach is described by Behzadan et al. in projecting construction graphics into the real world [8]. They developed an AR-application, equipped with a HMD, a GPS receiver and a portable computer. The aim was to combine virtual reality with the construction, while the user is able to move freely in the environment.

Veas et al. investigated possibilities to extend the viewport in AR applications under different circumstances. Therefore they described the multiview-AR and variable-perspective-view. Thus, the user was able to see the field of view from different perspectives without the requirement to move. Moreover it is possible to swap between the first-person-view and a third-person-view to change the perspective variable [9].

Considering planar objects from a distance, thus causing the perspective projection to display objects in very small sizes which causes them to be very difficult to detect. This problem is known as "long flat view", studied by King et al. [10]. One possible solution was to use a second camera, which is twice as high as the user. This doubles the field of view and therefore provides improved data for the depth. In addition to this problem King et al. studied also the problem of unreadable displays due to high solar radiation. This problem could be minimized by the use of dark, semi-transparent plastic on the screen or the use of umbrellas or hats. Also discussed was the issue of transparency of objects that are either not visible at certain color values during sunlight or they mask the reality completely.

In addition to the projection of objects there also exists the possibility to make objects disappear. This approach was described by Avery et al. [11]. In this case a mobile roboter was used to record hidden areas and transferring them directly to the user. Similar approaches to project hidden objects have been investigated by Webster et al. [12].

However, for most approaches it remains unclear which precision can be achieved under nowadays everyday conditions. Most approaches were tested under laboratory conditions, are marker-based, or hardware and software reality have changed drastically during the last years. In this paper we provide a glimpse on achievable accuracy under everyday conditions with standard AR projection techniques and consumer devices.

3 MapAR: An AR Tool for Geo-Data

In this paper we present MapAR, an AR tool for projecting invisible data or properties (e.g. collected by OpenStreetMap) in the camera image of everyday smartphones. With MapAR we are also evaluating the feasibility of markerless AR in context of geographic applications.

3.1 System Design

MapAR provides the possibility to project invisible data or properties in the camera image. Therefore it requires the coordinates of the data to be displayed. Figure 1 shows the projection of a parking lot in the main view of the application.

Figure 1: Arrow pointing to a parking lot. In MapAR.

3.2 Projection

Within MapAR we implemented following projection. To calculate screen coordinates, the position and orientation of the camera is required. Also the object to be projected must be available in Cartesian coordinates. Subsequently this data is used for a camera transformation to move the camera into the origin of the coordinate system. Thus, the coordinate system has to be rotated around the camera orientation $(\theta_x, \theta_y, \theta_z)$. As a result we obtain the point P (d_x, d_y, d_z) in camera coordinates [13, 14]. Figure 2 shows the corresponding matrix operation where (p_x, p_y, p_z) the current point to be projected is illustrated [16]. The first matrix causes the necessary rotation about the x-axis, the second matrix for rotation about the y-axis, and the third matrix of the rotation around the z-axis. Subsequently, the position of the camera from the point to be projected is subtracted to determine the position of the point in the camera system. Due to the perspective projection, we obtain a point in the camera coordinate system $(B(x, y, z))$.

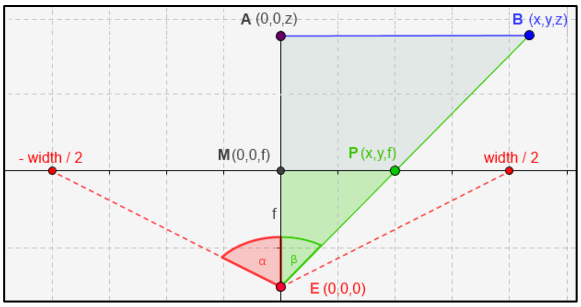
The next step translates the obtained point B in screen coordinates. Therefore the viewport of the camera as well as the size of the screen ($width * height$) is required. The focal length, so the distance from the camera center to the projection area, can be calculated through trigonometric calculations. In the figure (2) the focal length is displayed by f. For the calculation of f, the horizontal view angle α and the width of the screen ($width$) is required:

$$\tan\alpha = \frac{\frac{(width)}{2}}{f} \qquad (1)$$

Equation (1) can be resolved to f:

$$f = \frac{\frac{(width)}{2}}{\tan\alpha} \qquad (2)$$

Figure 2: Calculation of screen point x



By using the side-splitter-theorem the corresponding screen position can be calculated. The side-splitter-theorem states that a line that is parallel to a side of a triangle and intersects the other two sides of the triangle, divides the area of the triangle proportional. Figure (2) shows the triangle ABE. This triangle is divided by $\overline{MP}$. In addition to that the line $\overline{MP}$ is parallel to the line $\overline{AB}$. The following applies:

$$\frac{E * P}{P * B} = \frac{E * M}{M * A} \qquad (3)$$

By substituting the values of Figure (2) in Equation (3), we get:

$$\begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_z & -\sin\theta_z \\ 0 & \sin\theta_z & \cos\theta_z \end{bmatrix} \begin{bmatrix} \cos\theta_z & 0 & \sin\theta_z \\ 0 & 1 & 0 \\ -\sin\theta_z & 0 & \cos\theta_z \end{bmatrix} \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \left(\begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} - \begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix} \right)$$

Figure 2: Calculation of screen point x

$$\frac{x}{B(x)} = \frac{f}{B(z)} \qquad (4)$$

By substituting f with the calculated focal length in Equation (4) we get:

$$\frac{x}{B(x)} = \frac{\frac{\frac{width}{2}}{\tan\alpha}}{B(z)} \qquad (5)$$

And finally, we can solve for x :

$$x = \frac{B(x) * width}{2 * \tan\alpha * B(z)} \qquad (6)$$

The calculation of y is done equivalently. Instead of the width, the height is used and the horizontal view angle is replaced by the vertical view angle.

We repeat this procedure for every point in the object's outline and connect the points in the projection following the input sequence. Hence, the polygon can be displayed on the screen.

3.3 Sensor Fusion

Determining geographical locations requires sensor data received from GPS and orientation sensors of current smartphones. As orientation and GPS sensors don't provide very accurate data due to hardware and environmental factors (e.g., reflections) the information needs to be filtered. We implement different methods for sensor fusion and noise elimination. E.g., we weight the incoming GPS readings according to their timestamp, as typically more recent information provides more accurate information. We smooth the positioning information by calculating the average of this weighted value and previous weighted values.

4 Evaluation

With MapAR we want to explore the possibilities and limits of AR in geographic application scenarios. We designed different test cases under different conditions. We have chosen areas in the real-world under controlled and varying conditions and evaluated the projected areas with respect to accuracy of area, angles, perimeter and distance.

4.1 Evaluation Setup

For testing the precision of projecting objects under markerless everyday conditions with consumer devices, we decided to project parking lots as reference objects, as they have a defined rectangular shape of the size 5 x 2.35 meters and are visible on satellite imagery. With this simplistic shape we also can easily assess the properties of the projection with respect to the real-world object. The used device was a Samsung S3.

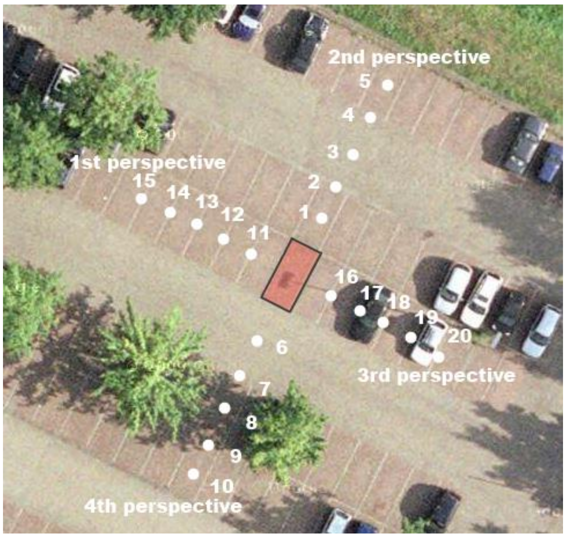
We recorded screenshots from projected parking lots. On a desktop computer with a 24" screen we manually selected the corner points of the projected rectangle with very high precision (we used a 27" screen with a resolution of 2560x1440 pixels, images where zoomed in to identify the correct position as precise as possible). We then translated the projection into geographical coordinates and reversely calculate the deviations from the correct parking lot, see Figures 4 and 5 for an illustration of the work flow.

In order to evaluate MapAR under realistic conditions we evaluated the result with four different variations (see Fig. 3):

- ***Differing perspectives***: we recorded 4 different perspectives for each parking spot in varying distances between 3 and 8 meters in order to rule out influences on perspective adaptation of the method.
- ***Differing distances***: we recorded each parking lot from 5 different distances (2.5, 5.0, 7.5, 10.0, 12.5 meters).
- ***Multiple recordings***: due to varying accuracy of GPS positioning we recorded two pictures for every position to rule out obvious outliers.
- ***Differing entities***: we used two different parking lots.

The conditions in our evaluation setup resulted in 80 individual measurements of the projection.

Figure 3: 5 x 2.35 meter parking spot in 4 different perspectives and different distances



.

Figure 4: The correct parking lot is outlined on the gound.

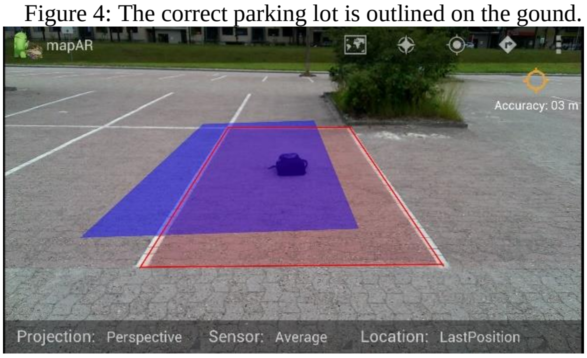
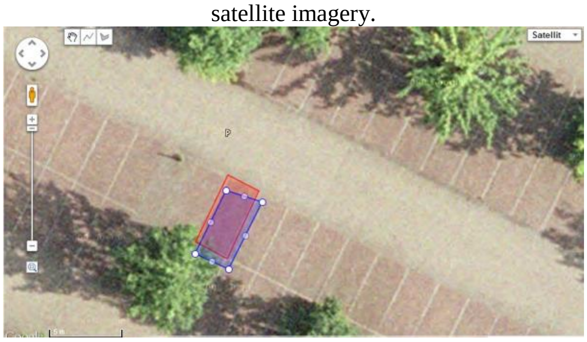


Figure 5: Translating the projection back to (geographic) world coordinates and projecting them back on the used satellite imagery.



We then compared the resulting 80 projected polygons with the original source polygon with respect to following properties:

- ***Center point distance***: the distance from the center of the projected polygon to the correct polygon (positioning accuracy).
- **Area**: we compared the area of the projected polygon with the correct polygon.

- ***Interior angles***: since the parking space is a rectangle, each interior angle has to be 90 °. We measure the deviation of the interior angles of the projected polygon.
- ***Perimeter***: Each parking lot has a perimeter of 14. 70 m. The perimeter of the projected polygon is compared to this value.

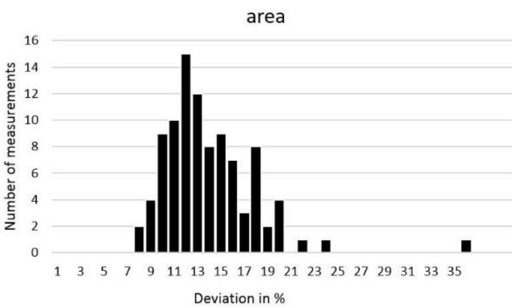
4.2 Results

The deviation of the distance to the center point of the parking lots has two peaks. While the first peak (16% of measurements) expresses a comparable small deviation of below 2m, the second peak (80% of measurements) clearly shows a relatively high deviation of up to 6 meters. This is due to the current positioning accuracy achievable with consumer grade GPS sensors (see Figure 6).

Figure 6: Deviation of the distance of the parking lots

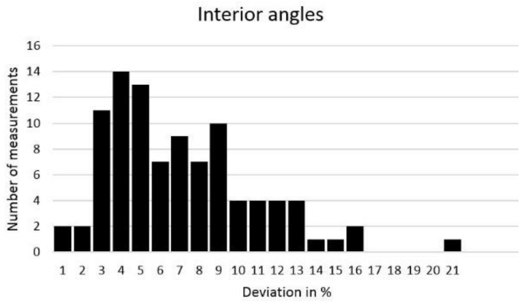
Our results show different deviations for our measurements. The deviation of the area of the projected parking lot is between 7 to 20 percent (Figure 7).

Figure 7: Deviation of the area of the parking lots



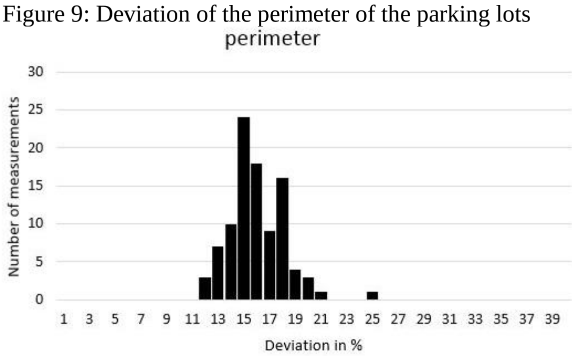
The deviation of interior angles has a peak between 4%- 9% indicating that the rectangle shape of the parking lot is well sustained in the projection (Figure 8).

Figure 8: Deviation of interior angles of the parking lots



The deviation of perimeter has a peak between 13% - 19% with almost 90% of all measurements inside of it (Figure 9). This also indicates a good maintenance of shape and size.

Figure 9: Deviation of the perimeter of the parking lots



5 Discussion

When interpreting the obtained results of our evaluation by means of geographic entities, we can identify the fields of application of markerless AR within geographic applications.

- ***Center point distance***: A large number of measurements (80,2%) showed a distance deviation of 5-6m, due to GPS inaccuracy. Typical entities of this dimension are smaller streets, smaller buildings, larger cars, parking lots, footprints of individual trees, etc. Any object of these or similar classes, depending on the configuration, might not be precisely addressable: if a similar entity is located directly next to the one to be augmented, in many cases the wrong entity will be augmented. I.e., if the entities are of a size in the range of the deviation augmentation is advisable only if the distance is large enough to guarantee disambiguation.
- **Area**: Although areal deviation is also in a perceivable range, most applications will still make sense, as large deviations in distance and shape might in many cases be more problematic. Many projected entities will have a certain counterpart in the real world and will be possible to correctly identify this entity even if the correct size is not preserved. As not arbitrarily large

entities can be projected to full extent, the achieved accuracy will often be below the distance error.

- ***Interior angles***: Our evaluation shows that geometry is preserved to a very high degree, indicating that information of sensors of the device itself already precise allows precise projections (within geographic application context).
- ***Perimeter***: 87% of measurements are between 13% - 18% deviation. This result is similar to the area deviation.

In the current state of technology (which is mainly limited by positioning accuracy), AR applications are applicable for entities of the size of the positional deviation or above. If the entity is perceivable without the help of augmentation and is a rather unique entity with respect to its surrounding, it can be also smaller.

I.e., in scenarios where precision (of currently) <5m is not required or entities can be perceived and matched due to their physical properties, it is feasible to use AR techniques in conjunction with consumer technology. However, in many cases this excludes scenarios without visually perceivable entities: examples are underground infrastructural elements like pipes, cables, or small scale excavation sites; identifying the correct entity can cause large efforts and costs.

The more alternative positioning systems (e.g., GLONASS, BeiDou, Galileo) and precision enhancing techniques are on the rise in the consumer market, the more can markerless AR be applied in geospatial high precision contexts with out-of-the-box consumer technology.

6 Conclusions

In this work we evaluated the applicability of AR techniques within the context of geographic applications.

As our evaluation shows, the application scenarios are mainly limited by the accuracy of the current predominant GPS positioning. This excludes a number of application scenarios from using AR as suitable method for identifying invisible properties or specific entities. Nevertheless there are numerous possibilities in which the application can be used with fewer requirements in terms of precision.

Acknowledgements

We gratefully acknowledge funding granted by the German Research Foundation (DFG) via the Transregional Collaborative Research Center SFB/TR8 Spatial Cognition, as well as funding granted by the European Union via mSAFE, grant agreement no. FP7-PEOPLE-2011-IRSES 295269.

References

[1] F. Liarokapis, I. Greatbatch, D. Mountain, A. Gunesh, V. Brujic-Okretic, J. Raper. »Mobile Augmented Reality Techniques for GeoVisualisation«. In: Proceedings of the Ninth International Conference on Information Visualisation. IV ‘05. Washington, DC, USA: IEEE Computer Society, 2005, S.745-751.

[2] F. Schmid, L. Frommberger, C. Chunyuan, C. Freksa. »What You See is What You Map: Geometry-Preserving Micro-Mapping for Smaller Geographic Objects with mapIT«. In: Geographic Information Science at the Heart of Europe. Eds. Danny Vandenbroucke, Bénédicte Bucher, Joep Crompvoets. Lecture Notes in Geoinformation and Cartography. Springer International Publishing, 2013, S. 3–19

[3] R.Behringer. »Registration for outdoor augmented reality applications using computer vision techniques and hybrid sensors«. In: Virtual Reality, 1999. Proceedings., IEEE. 1999, S. 244–251.

[4] D. Stricker, T. Kettenbach. »Real-time and markerless vision-based tracking for outdoor augmented reality applications«. In: Augmented Reality, 20001. Proceedings. IEEE and ACM International Symposium on. 2001, S. 189–190.

[5] R. Azuma, B. Hoff, I. Neely H., R. Sarfaty. »A motion-stabilized outdoor augmented reality system«. In: Virtual Reality, 1999. Proceedings., IEEE. 1999, S. 252–259.

[6] Y. Wu, M. E. Choubassi, I. Kozintsev. »Augmenting 3D urban environment using mobile devices«. In: Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on. 2011, S. 241–242.

[7] G. W Roberts, A. Evans, A. Dodson, B. Denby, S. Cooper, R. Hollands u. a. »The use of augmented reality, GPS and INS for subsurface data visualization«. In: FIG XXII International Congress. 2002, S. 1–12.

[8] A. H. Behzadan, V. R. Kamat. »Visualization of construction graphics in outdoor augmented reality«. In: Proceedings of the 37th conference on Winter simulation. WSC ’05. Orlando, Florida: Winter Simulation Conference, 2005, S. 1914–1920.

[9] E. Veas, R. Grasset, E. Kruijff, D. Schmalstieg. »Extended Overview Techniques for Outdoor Augmented Reality«. In: Visualization and Computer Graphics, IEEE Transactions on 18.4 (2012), S. 565–572.

[10] G. R. King, W. Piekarski, B. H. Thomas. »ARVino – Outdoor Augmented Reality visualisation of viticulture GIS data«. In: Proceedings of the forth IEEE and ACM International conference on Mixed and Augmented Reality (ISMAR), oct 5-8, Wienna. 2005, S. 52–55

[11] B. Avery, W. Piekarski, B. H. Thomas. »Visualizing Occluded Physical Objects in Unfamiliar Outdoor Augmented Reality Environments«. In: In 6th Int’l Symposium on Mixed and Augmented Reality. p. 2007, S. 285–286

[12] Webster, S. Feiner, B. Macintyre, W. Massie, T. Krueger. »Augmented reality in architectural construction, inspection and renovation«. In: Proc. ASCE Third Congress on Computing in Civil Engineering. 1996, 913–919

[13] H. Goldstein. »Classical Mechanics«. 2nd. Reading, MA: Addison-Wesley Publishing Company, 1980, 146–148.

[14] K.K.F. Riley, M.P. Hobson, S.S.J. Bence. »Mathematical methods for physics and engineering«. Cambridge University Press, 2006, S. 931-942.

[16] F. Schmid, L. Frommberger, C. Cai, and F. Dylla. 2013. »Lowering the barrier: how the what-you-see-is-what-you-map paradigm enables people to contribute volunteered geographic information«. In Proceedings of the 4th Annual Symposium on Computing for Development (ACM DEV-4 '13). ACM, NY, USA