Image Recognition to Improve Positioning in Smart Urban Environments

Sara Paiva(✉), Pedro Rodrigues, and Benjamim Oliveira

Instituto Politécnico de Viana do Castelo, Viana do Castelo, Portugal
sara.paiva@estg.ipvc.pt, sara.paiva@ieee.org,
{pmiguelrodrigues,benjamimoliveira}@ipvc.pt

Abstract. This paper describes a solution and algorithm to enhance positioning in outdoor environments with high buildings to be used in a mobile application to aid visually impaired people for navigation purposes. We used an image recognition algorithm and adjusted the android app algorithm to decrease the initial error average of 85 m (without any correction from GPS obtained signal) to a 5 m error, in the final version of our solution.

Keywords: Image recognition · Outdoor positioning · Outdoor navigation · Urban mobility · Visually impaired people

1 Introduction

Obtaining a precise position is necessary in several applications we use in our day-to-day life and its improvement gains increasing importance as we witness new research and studies in these areas. The need for precise positioning can happen indoors (inside buildings) or outdoors (outside buildings). Each situation has its own particular and specific approach. We will focus on this paper in the outdoor location that is associated in the vast majority of cases to the use of GPS. In fact, most of the applications we have on our mobile phones need only this technology to guide the user on the road while driving or while walking on foot to a certain location. But when we speak of location, sometimes a margin of error of 5 m is acceptable, in other cases not. We know that the GPS position obtained has a margin of error that can be considerable, especially if we are in urban environments where there are many buildings, sometimes high, that make it impossible to receive a better GPS signal. In fact, achieving a high precision in urban environments is clearly still an issue to solve [11]. If we consider the traditional applications developed for drivers on the road, it can be considered that the current state of these applications is already advanced, considering that the effectiveness and real usefulness to end users, in this case drivers, is already quite good. However, if we look at other types of applications, such as assisting visually impaired people (VIP) people to make footpaths in the historical centers of cities, the accuracy achieved with the GPS of a commercial smartphone is still clearly insufficient. World Health Organization

---

S. Paiva—IEEE Senior Member.

P. Pereira et al. (Eds.): SC4Life 2019, LNICST 318, pp. 59–70, 2020.
https://doi.org/10.1007/978-3-030-45293-3_5

(WHO) [1] stated in 2018 that globally, it is estimated that approximately 1.3 billion people live with some form of vision impairment. The International Classification of Diseases classifies vision impairment into two groups, distance and near presenting vision impairment.

With regards to distance vision, 188.5 million people have mild vision impairment, 217 million have moderate to severe vision impairment, and 36 million people are blind. With regards to near vision, 826 million people live with a near vision impairment. Globally, the leading causes of vision impairment are uncorrected refractive errors and cataracts. According to the study published in 2017 by The Lancet Global Health, compared to 1990, the number of blind people increased by 17.6% and that of individuals with moderate and severe problems about 35.5%. The same study also predicts that by 2020 there will be 237.1 million people with moderate to severe vision impairment and the number of blind people will be 38.5 million. By 2050, these figures are estimated to be about 587.6 million and 114.6 million individuals [14]. In Portugal, according to the 2011 Census, there are 900,000 cases of visual impairment, of which 28,000 are estimated to be entirely blind [2]. Approximately 80% of all vision impairment globally is considered avoidable. The majority of people with vision impairment are over the age of 50 years. A person's experience of vision impairment varies depending upon many different factors. This includes for example, the availability of prevention and treatment interventions, access to vision rehabilitation (including assistive products such as glasses or white canes), and whether the person experiences problems within accessible buildings, transport and information. According to this reality, technological research related to the promotion of accessibility and quality of life of the VIP has been under study for several years. Several solutions have also been developed sustained by the evolution of computational processing and data processing. Technologies have evolved and reached a big level of maturity that allow new technologies, approaches and algorithms to solve problems related to VIP more effectively, always with the goal of making their lives less affected by this deficiency to the greatest extent possible. In this article, we present a practical case study in the city of Viana do Castelo, Portugal, where we worked alongside an association of blind and partially sighted people (with whom previous work has already been done to help them through computer solutions in bus navigation in the center as well as pedestrian navigation [7, 10]). In this paper, we present an algorithm we developed and that is incorporated in a mobile app to make image recognition in order to allow the correction of a received GPS position. Images are captured from the visually impaired people (VIP) smartphone camera which triggers the algorithms to identify that image with a pre-loaded geo-referenced image in a dataset, so the GPS position of the user can be corrected and enhanced compared to the position directly obtained from the GPS.

The rest of the paper is structured as follows: the next section presents a literature review on positioning technologies. Section 3 explains the methodology followed for this project and the next chapter has all details about its implementation. Section 5 presents results and discussion and Sect. 6 presents main conclusions.

2 Positioning Technologies

GPS (Global Positioning system is probably the major outdoor positioning technology we have at the current moment, assisting several location-based services. It is used in industries such as aviation, nautical navigation, and land surveying, to personal and commercial applications such as driving navigation, people tracking, location sharing [21]; also, it can be of great help and fundamental to make a difference in the lives of people with disabilities facilitating day-to-day tasks. We have referred already to the imprecision of GPS position, namely in places with high-building, that can go from 1 to 20 m. Considering this, one study to overcome this situation is to use GPS to implement a high-accuracy global positioning solution and human mobility captured by mobile phones [22]. Authors developed a prototype system, named GloCal, and conduct comprehensive experiments in both crowded urban and spacious suburban areas. The evaluation results show that GloCal can achieve 30% improvement on average error with respect to GPS. Another study combines GPS with smartphone-enabled dead reckoning to produce a high-accuracy global positioning solution based on GPS and human mobility captured by mobile phones [21]. The performance comparison of four self-position fixing methods is made in [8] where authors analyze RSS Statistics, Least squares, Weighted LS (WLS) and Constrained WLS. Other authors provide a solution that makes use of a handheld device [3]. Their ideas to estimate the 3D location and 3D orientation of the phone camera based on the knowledge of the street objects. The use of vision process to generate a3D model with inertial and magnetic measurements is yet another proposal to update a Pedestrian Dead-Reckoning process and to improve the positioning ac-curacy [4]. Other contribution investigates the spatial correlation of multi-path error to estimate the multipath error at a pedestrian by using a regression model and leveraging the multipath errors at nearby points [12]. In [20] authors propose a new positioning algorithm based on signals only with pseudo range error modeling in association with an adapted filtering process. Another approach corrects signals using the European Geostationary Navigation Overlay Service (EGNOS) [19] and also presents an overview of the accuracies that can be achieved with different modes of GPS positioning. In [18] a method to mitigate multiple path in GNSS applications is proposed. Their principle is based on the addition, to the geographic map of a city or a neighborhood, of a supplementary information that consists of the correction of error caused by the multiple paths. Issues and problems associated with GNSS are presented and discussed in [13]. GPS and GLONASS were chosen for a comparison on accuracy and precision. The tightly-coupled single-frequency multi-GNSS RTK/MEMS-IMU integration is proposed in [9] to provide precise and continuous positioning solutions in urban environments. In a similar way, authors in [23] describes and applies a DGNSS-correction projection method to a commercial smartphone. An example of the use of images to enhance positioning is used in [15] where authors use the camera and a GPS-tagged image database to identify with precision the position the user stands. In [6] authors propose to improve the localization of scene items based on state-of-the-art map data, combined with a coarse and cheap position estimation as provided by standard GNSS. In [16] authors propose a method to geo-locate a mobile device by recognizing what is captured by its camera. A visual recognition algorithm in the cloud is used to identify geolocated reference images that match the cameras view. Some case studies are also presented and found in the literature. A study on the performance of BeiDou Navigation Satellite System in urban environments have been made [5] in Wuhan, central China. The compared data (visibility, position dilution of precision and the positioning accuracy) showed that the single point positioning accuracy of the GPS/Beidou dual system is much better and more stable than single BeiDou system. The POI Explorer mobile application aims to help VIP in their spatial orientation and urban navigation, using a smartphone, its sensors and data connectivity. Interaction is made through text-to-speech and notifications [17].

3 Methodology

The methodology followed for this project consisted in the concrete identification of the type of signal collected in the historical center of the city of Viana do Castelo, in one of its busiest streets. The goal was to realize the real mistake and therefore the impossibility of using only GPS to develop an application that allows VIP to move around the city and get relevant information about their location that allows them to gain more autonomy. After that, we designed an algorithm that compares captured camera image from the commercial smartphone to compare it to a preloaded geo-referenced image in a dataset so position can be enhanced compared to GPS received position. We made first round of tests that showed necessary improvements to be incorporated in the algorithm which we did to create a second version of it. Final results will be presented based on this second version of the algorithm and tests made in the field.

4 Implementation

Stage 1 - GPS Precision Analysis
As aforementioned, the first step was to pick a strategic street in the historical center and gather some data on the reliability and precision obtained. The street is shown in

Fig. 1. Street chosen for the tests made

Fig. 1 and it is possible to see it has high buildings which naturally affects received GPS signal.

The tests made included going down from one side and then going up from the other side, covering both sides of the beginning of the street. 167 positions were chosen, the exact position was marked in a map using a mobile app and the received position was compared to it. The average error received, in meters, was of 85.5. The minimum error obtained was 6.3 m and the maximum 205.1 m. Figure 2 shows the exact location of tester going down the road (green markers) and the obtained position through GPS (red markers). It is possible to see that most locations were received has been on the other side of the road, which makes it completely impossible to use only GPS to create a location-based app to aid VIP in their navigation needs.

Fig. 2. Correct position vs. received GPS position without any correction (Color figure online)

Stage 2 – First Algorithm

This stage consisted in developing the algorithm that would allow for an image from the VIP cellphone camera to be compared with a dataset of geo-referenced images, so a better position than the GPS could be returned to a mobile app to be used for navigational purposes to aid VIP. For building the dataset we captured some relevant objects throughout the street that could serve as a reference point and allow for the image recognition to be successful. At this point, we were aware that the success of the entire algorithm was mostly related to the performance of the image recognition algorithm.

Choosing the Image Recognition Algorithm

OpenCV (Open Source Computer Vision Library) is a well-known open source computer vision and machine learning software library, with different comparison methods. The Histogram Comparison Method is the fastest and simplest to adapt regarding image comparison. After we have analyzed this algorithm, we concluded it would not be appropriate to use in this case, because images are compared mostly by color, and as there are several buildings with the same color, it could lead to a huge margin of error, since two images containing buildings of the same color would be considered the same. The Template Matching method basis its comparison mainly depending of its size, which would also not work in this project, since all the images would be considered equal, assuming they would be captured through the smartphone camera. A more thorough search has allowed us to choose the Feature Matching Algorithm. Initially the key points of a given image (in this case the image already stored in the database) are chosen and later compared with the key points of another image (the image captured by the smartphone camera). The algorithm keeps track of the points so that they can be detected and recognized regardless of how the image is, either re-scaled or turned.

The Android App Algorithm

After choosing the algorithm for image recognition, it was time to incorporate it in the global algorithm to be executed in the Android mobile app to make the tests to conclude about enhancement on precision. The behavior of the first version of the algorithm, shown in Fig. 3, is as follows:

- We receive the GPS coordinates from the mobile app. These are the coordinates we will try to enhance since, as we know, they will most likely be wrong by a few tens of meters;
- Check if the quality of the signal is good enough, using the “accuracy” parameter received alongside the GPS coordinates;
- Depending on the GPS quality: (a) If the GPS signal is considered to be good enough, we use the coordinates provided by the GPS. The algorithm ends. (b) If the GPS signal is considered “weak” we computed a search area (based on the “accuracy” parameter) and retrieve all the images geo-referenced inside that same area;
- Two parallel actions are made at this point: (a) Obtain images on the range of the search area. Each image on our dataset is geo-referenced. Using the area previously calculated we retrieve all the images that are inside it. (b) Obtain image from the smartphone camera;
- Compare captured image with images from the dataset. To find out which image matches best the image retrieved by the smartphone we have to loop through all of the images we have filtered;
- Retrieve position of the most similar image. Once the previous loop ends, we will retrieve the image with more similarities to the image being captured by the smartphone camera;

- Place a marker on the map in the new corrected position and use it to inform the user of his location or nearby obstacles.

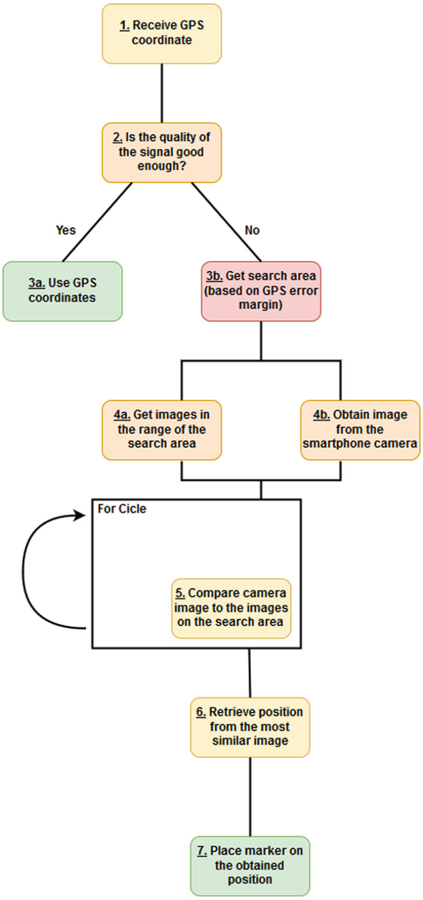

Fig. 3. Algorithm - version 1

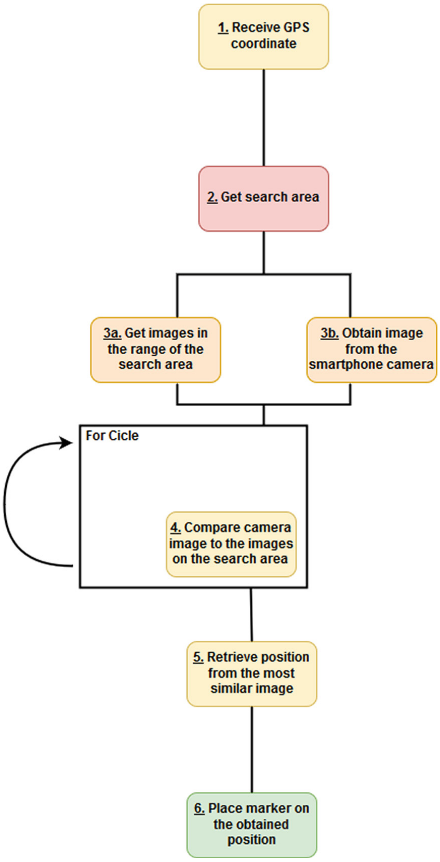

Fig. 4. Algorithm - version 2

When testing on the field this first version, two main problems were noticed. The first has to do with the “accuracy” parameter returned by the GPS. It cannot be used to decide about precision of the signal as plenty of signals are marked as having a good accuracy with tens of meters of errors. The second change has to do with how the dataset was being created. Figure 5 shows some images in the dataset, with which captured images on the smartphone were to be compared with. The image recognition algorithm did not have a good behavior as its strength is to detect recognizable symbols which are small in these first images and therefore, we had few matches.

Fig. 5. Examples of images from the first dataset

So, we developed a second version of the algorithm shown in Fig. 4. The main changes were the removal of step 2 regarding accuracy, for the reasons already mentioned. In this last version, we also narrowed dataset number of images directly after receiving GPS coordinates. This is an important and necessary step as image recognition takes some milliseconds and, depending on the number of images on the dataset, and considering GPS position can be gathered from the smartphone every 2 or 3 s, the algorithm had to be ran in a fast way. To achieve this, images in the dataset had to have a considerable small size, otherwise several seconds would be needed to run the algorithm which would not suit the mobile application itself.

5 Results and Discussion

As previously explained, this project consisted of three main phases. The first one analyzed the accuracy of the signal obtained without any correction. In a second phase, a dataset of images was defined and developed a first algorithm. Finally, in phase 3, an improved version of the algorithm was developed, according to tests made in phase 2. In each phase, our purpose was to measure the accuracy of the signal received compared to the real location of the user. In the first phase we obtained an average of 85 m of error comparing the real position with the obtained GPS position. In the second phase, we used the first version of the algorithm and the average error was of 54 m. At this point, we expected to have better results and understood the result was very much originated by the incorrect match/recognition of the image the smartphone was capturing, comparing to the ones on the dataset. At this point, we adjusted the algorithm and also changed the dataset of images, so they are mainly composed of brands/logos or very relevant and significant sub-parts of the image being captured by a camera. Figure 6 shows some images used in the second dataset which are significantly different from the ones used in the first dataset shown in Fig. 5. This allowed the image recognition algorithm to have a much better behavior and recognize images more successfully.

Fig. 6. Examples of images from the second dataset

With the changes made in the second algorithm, which corresponds to phase 3 of our process, the obtained average error went down to 5 m, as can be seen in Fig. 7.

Similar to Fig. 2, we have measured correct position versus position obtained with version 2 of the algorithm and results are, as can be seen in Fig. 8, much more accurate.

The improvement was significant (reduction of approximately 80 m in average compared to phase 1 and reduction of approximately 50 m compared to phase 2), but we know it can be even better and we need it to be we this algorithm can be used in mobile solutions for visually impaired people. Images are correctly detected, and position is corrected but 2 or 3 m ahead, the recognition of the image is still done, and corrected position keeps being the same of the coordinates in the image from the dataset which contributes to some errors. To further decrease the obtained error number and solve this problem, we need also to use smartphone sensors, so we detect how much the user has walked since the first correction considering a given image and introduce an additional correction, that not only is based on the coordinate of the image that made a match but also on the user's movement.

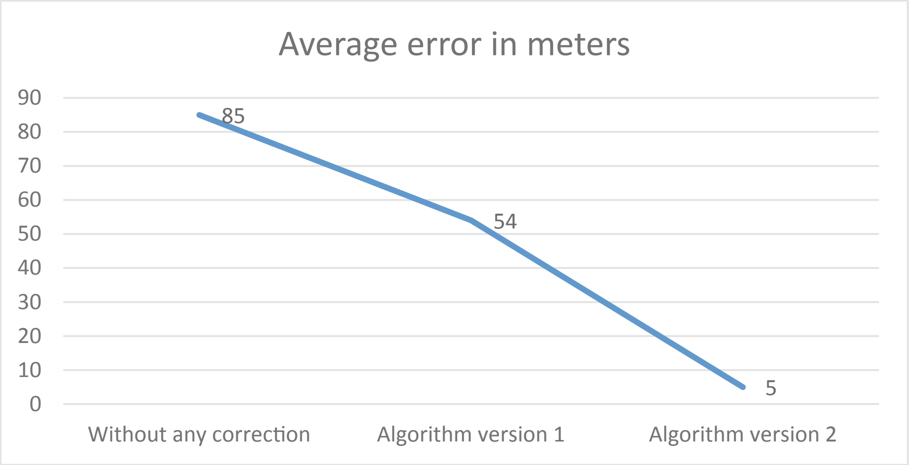

Fig. 7. Comparison of average meters obtained in each of the 3 phases

Fig. 8. Correct position vs. received GPS position using algorithm version 2 (Color figure online)

6 Conclusions

In this paper we have presented an on-going work that has the main goal to develop an algorithm to achieve better precision in positioning that can be used in a mobile application to aid visually impaired people to walk in a more autonomous way within cities. For this, we used image recognition to correct the GPS position received by the satellite that has, as we know and if used alone, still a big error. Our approach started by choosing a street with much usage in the city of Viana do Castelo and understand precision of GPS received signal. We found out average error is approximately 85 m, considering that are high buildings. So, we implemented the first version of an android app algorithm that makes use of an image recognition algorithm so the captured image from the smartphone camera is compared to a preloaded dataset of geo-referenced images. This first approach allowed us approximately to a 54 average error meter. It was still a big number mainly because of the images in the dataset were not the most appropriate as they captured a large part of the buildings. We had to change to significant markers and symbols only so recognition could be more successful. The second version of the algorithm achieved a 5 m error average which was a significant increase. To obtain further results we will, in a next stage, combine work done so far with smartphone sensors to allow to narrow down the error m obtained.

References

1. Blindness and vision impairment (2018). https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment. Accessed 13 June 2019
2. Associação dos cegos e amblíopes de Portugal (2019). http://www.acapo.pt/deficiencia-visual/perguntas-e-respostas/deficiencia-visual. Accessed 13 June 2019
3. Antigny, N., Servires, M., Renaudin, V.: Hybrid visual and inertial position and orientation estimation based on known urban 3D models. In: 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8 (2016)
4. Antigny, N., Servires, M., Renaudin, V.: Pedestrian track estimation with hand-held monocular camera and inertial-magnetic sensor for urban augmented reality. In: 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8 (2017)
5. Xiang, X., Bian, S.: Research on BeiDou positioning performance in urban environments. In: 2013 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC 2013), pp. 1–5 (2013)
6. Cao, G., Damerow, F., Flade, B., Helmling, M., Eggert, J.: Camera to map alignment for accurate low-cost lane-level scene interpretation (2016)
7. Faria, P., Curralo, A.F., Paiva, S.: A case study for the promotion of urban mobility for visually impaired people. In: Paiva, S. (ed.) Mobile Solutions and Their Usefulness in Everyday Life. EICC, pp. 65–82. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-93491-4_4
8. Krishna Reddy, D., Achanta, D., Satya, V.: Mobile position estimation with RSS based techniques in an urban city with multiple multi-storied structures, pp. 234–234 (2014)
9. Li, T., Zhang, H., Gao, Z., Chen, Q., Niu, X.: High-accuracy positioning in urban environments using single-frequency multi-GNSS RTK/MEMS-IMU integration. Remote Sensing **2018**, 205 (2018)
10. Lima, A., Mendes, D., Paiva, S.: Mobile solutions for visually impaired people: case study in Viana do Castelo Historical Center. In: 2017 12th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6 (2017)
11. Lin, H., Ye, L., Wang, Y.: UWB, multi-sensors and wifi-mesh based precision positioning for urban rail traffic. In: 2010 Ubiquitous Positioning Indoor Navigation and Location Based Service, pp. 1–8 (2010)
12. Patou, Y., Obana, S., Tang, S.: Improvement of pedestrian positioning precision by using spatial correlation of multipath error. In: 2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES), pp. 1–8 (2018)

13. Przestrzelski, P., Bakua, M.: Study of differential code GPS/GLONASS positioning. Ann. Navig. **21**, 117–132 (2014)
14. Bourne, R.A., et al.: Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. The Lancet Global Health **5**, e888–e897 (2017)
15. Salarian, M., Manavella, A., Ansari, R.: Accurate localization in dense urban area using google street view image, pp. 485–490 (2015)
16. Sanjuan, D.M., Adamek, T., Bonnin, A., Trzciski, T.: Enhancing global positioning by image recognition (2011)
17. Skulimowski, P., Korbel, P., Wawrzyniak, P.: POI explorer a sonified mobile application aiding the visually impaired in urban navigation, pp. 969–976 (2014)
18. Titouni, S., Rouabah, K., Attia, S., Flissi, M., Khababa, O.: GNSS multipath reduction using GPS and DGPS in the real case. Positioning **08**, 47–56 (2017)
19. Verbree, E., Tiberius, C., Vosselman, G.: Combined GPS-galileo positioning for location-based services in urban environment. LBS TeleCartography **66**, 99–107 (2003)
20. Viandier, N., Nahimana, D.F., Marais, J., Duflos, E.: GNSS performance enhancement in urban environment based on pseudo-range error model. In: 2008IEEE/ION Position, Location and Navigation Symposium, pp. 377–382 (2008)
21. Wu, C., Yang, Z., Xu, Y., Zhao, Y., Liu, Y.: Human mobility enhances global positioning accuracy for mobile phone localization. IEEE Trans. Parallel Distrib. Syst. **26**(1), 131–141 (2015)
22. Wu, C., Yang, Z., Zhao, Y., Liu, Y.: Footprints elicit the truth: improving global positioning accuracy via local mobility. In: 2013 Proceedings IEEE INFOCOM, pp. 490–494 (2013)
23. Yoon, D., Kee, C., Seo, J., Park, B.: Position accuracy improvement by implementing the DGNSS-CP algorithm in smartphones. Sensors **16**, 910 (2016)