Nat. Hazards Earth Syst. Sci. Discuss., 3, C2723–C2734, 2015
www.nat-hazards-earth-syst-sci-discuss.net/3/C2723/2015/

Natural Hazards and Earth System Sciences
Open Access
Discussions

Interactive comment on "Mobile Augmented Reality in support of building damage and safety assessment" *by* W. Kim et al.

W. Kim et al.

jinsilto@gmail.com

Received and published: 20 December 2015

Dear anonymous referee, thank you very much for your interesting comments and feedback. Below we answer your questions one by one.

p. 2602, line 10 Authors state that they retrieve location and navigation information from the mobile device (Smartphone). However, for the LOCs mentioned later in the paper, this information coming from current Smartphones will probably not precise enough.

Authors response: We agree that mobile device like smartphone and tablet PC have accuracies that may be limited for some LOC, especially in densely built-up areas. However, one of the research purposes was to assess accuracy and uncertainty for each LOCs, and we already describe accuracy limitations of mobile devices in each

Table of LOCs, and in the "Accuracy", "Uncertainty" part of 3.1.1, 3.1.2. p. 2602, line 12-14 Authors state that AR superimposes computer-generated information on real world images, which is correct. However, as an example for computer generated images they give a picture taken by the camera. What is computer-generated here?

Authors response: It refers to all data in digital format: digital photos, movie clips, 3-D building models, maps, sounds.

p. 2604, line 17 In their concept, authors state that they will use marker-based tracking inside a building. I think this is unrealistic. Who put the markers there and when? Were the markers already there before the event happened? It is very unlikely that every building will be applied with markers just for an emergency situation. And even if so: As soon as a building's structure is damages, the position of the markers is also wrong and thus the tracking doesn't work anymore.

Authors response: We agree that maker-based AR is currently not realistic, or at least it is limited in its utility. That is why we also did not actually use a marker-based approach. In the paragraph we mention examples of previous research that used not only marker- but also vision-based AR, which recognize building shapes without markers but using image processing. Looking ahead, though, we consider it entirely feasible that with improvements in indoor mapping and navigation (e.g., via projects such as Google Tango), there will be a growing number of applications for mAR. Once can thing here of using floor plans for a given design type (to be recorded, say, in the entrance area), with mAR projecting additional useful, floor-specific information as search- and rescue forces move through the building.

p. 2605, line 15 Authors mention Google Glass in the paper. This product is discontinued.

Authors response: Google only stopped project Glass, which means they stopped this specific development program. Furthermore, we only mentioned Glass as an example of wearable devices – there are still many other wearable VR products being developed, such as Sony's SmartEysGlass, Microsoft's HoloLens and so on. With VR, including mAR having such potential business value we consider it a matter of time before better performing devices are readily available.

p. 2605, line 16 Authors state that using some kind of glasses already allows for a hand-free operation. This is not true, since the interaction still has to be done on the Smartphone.

Authors response: Wearable devices such as Google Glass can be run independently, without smartphone connection. In addition, wearable devices usually work by voice control, which means the user does not need interaction with the smartphone.

p. 2606, line 10-11 I am wondering, if LOC 1 is really needed. Buildings usually do not completely disappear, but typically a fundament remains. This would already be sufficient to detect the position of a house. The only thing I can think about would be the height of a building, but this already belongs to LOC 2.

Authors response: It is not easy to identify a building's existence in a severely damaged area. For instance, many buildings on vulnerable areas are built without proper foundations, meaning a user cannot recognize even the existence of building. Even if the damages building does have a foundation, orientation is extremely challenging with road signs and similar visual markers gone. We suggest that the reviewer look at some images of a tornado site, the Japan tsunami-affected area, coastal areas affected by Katrina or any such site. We find the concept image in LOC1 (Figure 1) quite appropriate in this respect – with utter destruction it is nearly impossible to identify individuals buildings, especially when scanning the surrounding area from a ground position.

p. 2606, line 15 For LOC 2, authors stated that the system could also provide information on the material of a building. Why is this important?

Authors response: Building material is a one of the critical types of information. There are three reasons: First, according to the building material, a user can judge the build-

ing vulnerability, and say something about the type of collapse (e.g., in terms of the presence of cavities someone could survive in). This is especially important for building safety assessment. Second, evaluation forms like ATC-20 or EMS-98 require the building material to be recorded in the assessment. Third, during the survey, users indicated that the most important information for the assessment is building information including building materials (Please refer to Table 3 that shows value of different data type).

p. 2607, line 15 Authors state, that a lot of information for the LOCs has to be collected by interacting with the Smartphone using touch. This contradicts the prior statement of a hands-free operation. Moreover, touchscreen interaction might not be suitable at all if users wear protective gloves.

Authors response: This research focuses on mobile device like smartphone that rely on touch handling. The reason we mentioned wearable devices is to show that additional approaches may be possible in the near future. There are many limitations to current ground-based building assessment; first, regardless of wearing gloves, users bring many items such as maps, cameras, pens, notepads, or compass, and they just fill the evaluation form with a pen. A smartphone can replace all these items, making the process more efficient. In addition the result in Table 2 shows that touch handling can improve operation performance in the field by about 34%. Lastly, nowadays there are gloves that work on a smartphone screen.

p. 2607, line 19 Authors say that "text information is a form of point data". What do they mean by this?

Authors response: This means that one text information matches one locational coordinate (x, y), which is registered on the real building on the screen as a point.

p. 2608, line 7 For many LOCs, authors presume that there is a stable internet connection. However, in case of catastrophes there is typically no internet at all. Thus, all LOCs proposed cannot be used at all. It would be good if authors could take this situation into account and also offer a solution for this, e.g. by a remote station (command station as it is usual in ERM), which still has internet connection and could transfer the information e.g. via satellite radio.

Authors response: Except LOC3 that extracts information from internet services such as YouTube, google street view, each LOCs can copy data from a server into a mobile device without internet connection before dispatching to the field. In fact the prototype also saved all data in a smartphone before the outdoor test. However, we agree that we need to reflect about internet connection problem and possible solution on our paper, and we will expand this discussion in the revision.

p. 2608, line 15 Many LOCs need intense user interaction for selecting data. For this, the interactive surface of a Smartphone is simply too small to do all this in a reasonable time and under possible stress situations.

Author response: Basically, mAR works without any user interaction. It is a reference system that does not any require any type of user input and handling, and the system just shows reference data around user automatically. However, if a user wants to see a given building from a different point of view, the user just needs to touch the display in order to get additional information. In addition, mAR is limited to smartphone, but can be equally used on a tablet PC or smartphones with a large screen. Lastly, the survey results taken from experience structural engineer say that mAR can improve not only operational performance in the field but also location awareness.

p. 2609, line 9 When talking about accuracy, authors mainly think about the accuracy of the 3D data and the geospatial position. However, besides this the actually much more important problem is the information mapping. Due to the movements of a user when holding the device in his hand, the stabilized and precise alignment of the additional virtual information is much more important and cannot be solved by GPS information, but needs some image processing in the mobile device. Thus, it is not really the accuracy of the 3D data, but more a stabile spatial synchronization of the virtual data in the

visible image on the Smartphone.

Authors response: The mobile device does not only use a GPS but also the compass and the gyro sensors, which increase the accuracy and the stabilize of user's movement and alignment. However, we agree that in some cases, these kinds of locational information are not enough, and this kind of accuracy and uncertainty problem of each LOCs are already described throughout this research. In addition, this is why we recommended vison-based AR that uses image processing for future research.

p. 2610, line 9-10 I think it is not the 3D data that challenges the mobile device, but is more the required accuracy of all the sensors needed for the proposed system.

Authors response: We agree. We also already mentioned the accuracy problem in "Accuracy" in 3.1.2

p. 2610, line23-26 Using markers (fiducials) makes the system useless. One cannot expect that a building has markers inside. Thus, someone has to attach them forehead and precisely measure them, before the first responders come :-). This is unrealistic and would also require a lot of time and makes the question e.g. regarding whether a building could collapse obsolete. And even more: who could guarantee that inside a building WiFi is still available?

Authors response: In the paragraph the marker does not mean fiducial marker. It can be a simple map or satellite image of the area. Moreover, it is not an outdoor system that is used in field, but it is an indoor system that is used in indoor environments. Please see also the earlier answer regarding the use of floor plans for marker-based mAR. We see many possibilities to identify features (markers) within a building and associating them with auxiliary information in a mAR setup.

p. 2611, line 2 Authors strongly rely on touch interaction on the mobile device. However, I think this is the wrong way of interacting, since it would require that user have both hands free. This is typically not the case.

Authors response: same as previous answers. First, user can use this system without any interaction. It is a just an additional function. Second, our survey shows that touch interaction increase operational performance in field. Third, currently user just brings paper map, form, pencil to assess building damage which is less efficient and even more dangerous than using just one mobile device (and they also need both hands to operate most of the above tools and devices). Fourth, wearable devices like Google Glass or any other product can improve this problem.

p. 2611, line 7 What do authors mean with a "radar display"?

Authors response: Please see Figure 3, user interface. “Radar” shows user’s current location and the relative location of target buildings, the presence of and distance to hazardous substances, etc..

p. 2611, line 17 For the concept and system evaluation, authors stated that they use an online survey. Did the persons only fill out a questionnaire? Or did they have the chance to really test the prototype the authors talk about? I think, an online survey does not give a true answer to the usability of the system, but leaves a lot to the phantasy and imagination of the asked persons.

Authors response: We showed them video clips of every functions of the system before the survey. Also, the responders were highly experienced structural engineers and first responders, mostly with >5 years experience, and located in a range of countries in Asia, Europe and the North America. Having them physically test the prototype was thus not an option. We agree, though, that an online survey has limitation, which is already mentioned “5. Discussion” part.

p. 2612, line 19 Authors state that they are conceptualizing and testing the system. However, I do not see a real test of the system in the paper. There is only a questionnaire based on simple illustrations.

Authors response: testing result are shown in two video clips, which shows every func-

tion of the system.

p. 2612, line 23 Authors state that prototypes were developed and evaluated through online survey taken by experienced users. This is impossible!! Without having the real users’ hands on the system, this can’t be stated as an evaluation. Also a development without the involvement of the final users is already very critical.

Authors response: We agree that an online survey has limitations. Please recall that the principal aim of the research was to conceptualize what information could be meaningful in a disaster context, what LOC it would constitute, how or where data could be sources, etc. In addition we wanted to get detailed feedback from experienced users of the value they attached to the type of system setup we devised, and the different LOC. Now that we understand from the users what information is truly useful, and how it needs to be provided, we can take the next step and think of field tests.

p. 2613, line 5-8 I think the main reason for the negative feedback regarding indoor AR was due to the fact that people realized that they cannot presume markers inside a building, and - even if so - they might be damaged or imprecise due to the house’s structural changes

Authors response: Perhaps the term “indoor AR” is not entirely clear, and we will revise the text in the revision. “Indoor” does not mean inside of a damaged building. Indoor AR can also be used by stakeholders working “inside”, e.g. of an emergency headquarter or organization.

p. 2613, line 12 Authors state that mAR gains in value when combined with a map system (GIS). This is true, but is then the difference to Google street view or Google maps? And why not using this instead?

Author response: The major difference is that AR supplements reality with virtual data, while traditional GIS replaces reality. This means that in a GIS virtual data need to be translated into a specific shape or form, while in AR a user can directly interpret data comparing virtual data with real object.

p. 2613, line 17-19 Authors state that "users rated highly the usability of mAR based on touch gestures (e.g. zoom-in/out or rotation of 3-D building with two fingers) that let users manipulate contents on the screen interactively". How could they do this? In an online survey, they could have never experienced the system. Otherwise they would have noticed that the screen might be too small to retrieve all the required information interactively.

Authors response: We showed video clips to the users, which was enough to experience every function of the system. We think that most digitally-versed users know the difference between a smart phone and tablet, and what either can do in terms of showing complex information . We developed the prototype on a smartphone, but a tablet can be used as well (which was clearly communicated to the users)

p. 2614, line 2 Why is it necessary to have a street level imagery taken at different time periods? Wouldn't the pre-disaster view be sufficient?

Authors response: Relevant, location-specific changes (construction, demolition, modification) may have taken place at some point in the past, hence allowing the user to call on street data from different dates if available is useful.

p. 2614, line 7-9 Both scenarios LOC 5 and LOC 6 require a precise and repeatable positioning of the mobile device. This is impossible with current technology integrated in current Smartphones

Authors response: We agree that current technology has limitations. That is why we mentioned it as an additional concept that has not been proven yet. However, projects such as “Spike” or “Google Tango” already show such data integration possibility with smart devices.

p. 2614, line 17-18 Authors state, that "Using SfM in combination with dense image matching (DIM) 3-D models of post-disaster buildings can be reconstructed". The only

question is whether the Smartphone's sensor accuracy (including the camera) is suitable for doing so. As mentioned earlier, this also requires a stabile positioning of the Smartphone in all 6 DOF

Authors response: For LOC5 and LOC6, as we described, the system does not rely on GPS or gyro of mobile device because of accuracy. That is why we suggest technologies such as SfM or DIM which do vision recognition processing instead of using GPS.

p. 2614, line 24-25 Mounting an active scanner on the device will significantly reduce the usability of the system mainly due to handling problems in such harsh environments.

Author response: Active scanners (such as in the “Spike” system) are very small and integrate quite seamlessly with the smartphone platform. In systems such as Tango the scanner is not even a separated component, but rather it is built-in.

p. 2614, line 27-28 In "TANGO", they use a tablet, but not a Smartphone

Authors response: Develop Kit is only provided in form of tablet PC which does not mean Tango is targeting on tablet PC, and we are not considering only smartphones. The research targets mobile devices including tablet PCs.

p. 2615, line 1-2 I think the authors overestimate the outcomes of TANGO. They cannot detect such small changes. Moreover, in the beginning of the paper, the authors stated that the remote investigation of a building is not sufficient, and now they offer this as an option? This is not clear.

Authors response: We are suggesting Tango as one possible solution for vision-based AR that gives better registration accuracy between real and virtual objects. It gives an accurate user location and 6 DOF without relying on GPS or gyro. In addition, it can also give accurate information of real building as it scans in real time, e.g., building shape, height, inclination and so on. We will clarify this point further in the revision.

p. 2615, line 10-15 A large portion of the overall concept is based on assumptions and on future technical capabilities of mobile devices that might or might not be integrated in future mobile devices

Authors response: We already made a prototype using current mobile device for LOC1 to LOC4. LOC5 and LOC6 the only concept rely on future device, which is already described.

p. 2616, line 19 I think, some important references are missing, e.g. Leebmann, J.: "AN AUGMENTED REALITY SYSTEM FOR EARTHQUAKE DISASTER RESPONSE"; in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 34, Part XXX or: Aameer R. Wani, Sofi Shabir, Roohie Naaz: : Augmented Reality for Fire & Emergency Services; in: Proc. of Int. Conf. on Recent Trends in Communication and Computer Networks; 201

Authors response: Thank you for the suggestion. We will mention this paper as a very early attempt at AVR.

p. 2625 I doubt that a smartphone is a suitable interface to display information to first responders. This is because of the following reasons: Working in hazardous and partly destroyed environments requires both hands for performing actions. Thus, a hands-free operation of information devices would be necessary. There is simply no hand free to hold the smartphone. Moreover, first responders often wear gloves out of several reasons: to protect their hand from injuries, because of epidemic reasons, to avoid electrical shock, and so on. In any case, they won't be able to handle the smartphone to retrieve new information. Authors might also have a look at Billinghurst's "Earthquake VR" http://www.hitlabnz.org/index.php/research/augmentedreality?view=project&task=show&id=24 or: Ming-Kuan Tsai , Yung-Ching Lee, Chung- Hsin Lu, Mei-Hsin Chen, Tien-Yin Chou,Nie-Jia Yau: Integrating geographical information and augmented reality techniques for mobile escape guidelines on nuclear accident sites; Journal of Environmental Radioactivity 109 (2012) 36-44

Authors response: As we mentioned before, in any situation, first responders or structural engineer carry many items for the assessment, such as paper forms, paper maps, compass, pens, digital cameras, etc., which is already more dangerous than carrying one mobile device. Therefore, by replacing these items with one mobile device we can reduce the risk. The above items also typically require 2 hands and gloves off. However, we agree that there is still the risk of carrying something, hence we proposed wearable devices with voice recognition as a future solution, to reduce these kinds of risk. In addition, nowadays touch interaction is possible even with gloves, and also phone screens are getting large enough to allow decent interaction.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 3, 2599, 2015.