Poster: Augmented Reality Smartphone Compasses:

Opportunity or Oxymoron?

David S. Bowers
School of Computing & Communications
The Open University
Milton Keynes UK
david.bowers@open.ac.uk

ABSTRACT

The development of augmented reality capabilities on smartphones has led to the emergence of a range of AR apps, including AR compasses. Some of these apps claim to be as good as professional magnetic navigation compasses, and suitable for navigation use. This poster presents detailed measurements of compass deviation (error) curves and offset errors for augmented reality compass apps on 17 mobile devices. The magnitude of the deviation errors measured casts serious doubt on claims the apps are appropriate for navigation purposes. This in turn emphasizes the need for the ubiquitous computing community to help ensure adequate awareness of the limitations of some onboard sensors, including compasses, on devices such as smartphones.

CCS CONCEPTS

• Human-centered computing~Mixed / augmented reality • Human-centered computing~Smartphones • Human-centered computing~Empirical studies in ubiquitous and mobile computing • *Human-centered computing~Mobile devices*

KEYWORDS

Compass; error measurement; calibration

ACM Reference Format:

David Bowers. 2019. Augmented Reality Smartphone Compasses: Opportunity or Oxymoron? In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct)*, September 9--13, 2019, London, United Kingdom. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3341162.3343777

UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom

ACM ISBN 978-1-4503-6869-8/19/09.
https://doi.org/10.1145/3341162.3343777

1 INTRODUCTION

Inspired by the success of urban mobile AR, including games such as Pokemon Go, a range of open field AR navigation apps is emerging. Some claim to offer a high-precision, “professional” grade, hand bearing compass [9], with the additional ability to overlay a desired course, or label navigational markers [1, 6].

The internal magnetometer (compass) in mobile devices is an important component of mobile augmented reality (AR). However, since the earliest days of AR, it has rarely been used on its own to determine the orientation of a device, owing to the inevitability of significant compass errors [2]. Instead, considerable effort has been expended on registration techniques, to align markers in the displayed image with their real-world counterparts [5]. Image processing techniques such as edge detection and 3D mapping are used to snap the AR universe onto the observed image. Where no map of objects exists, more advanced techniques such as Simultaneous Location and Mapping (SLAM) [8] are deployed.

However, registration is not always possible. In an open outdoor setting there may be no distinguishable features nearby. The distances to landmarks may be significant, and the landscape may be observed from a distance rather than moved through, so techniques such as SLAM are not applicable. The challenge is complicated further in a marine context by part of the view itself moving, due to surface waves. Panorama stitching operations are too difficult in these circumstances, let alone locating landmarks. In such contexts, AR displays are completely reliant on the accuracy and precision of internal sensors – particularly the magnetometer.

If the internal compass is being used to tag objects at a distance then its accuracy is critical. Each degree of error over 2km is about 35m. With larger errors, the offset may be comparable with the scale of the objects themselves, and possibly larger than their separation in the landscape.

This experiment measured deviation curves for s mobile devices, including tablets, iPhones and Android phones. The full results are presented in the poster.

2 RELATED WORK

Given the use of registration techniques rather than the inbuilt compass for AR, it is unsurprising that there are few studies of the (in)accuracies of embedded compasses.

Blum et al [3] conducted an in-the-wild study using three devices held or carried by a subject walking in an urban environment. They deduce standard compass errors of up to 30 degrees, but they report only limited checks on the orientation of the device during the trial.

Smartphones have attracted attention recently in the geology community, where there is a need for rapid data collection in remote areas [7, 10]. The analyses of the errors observed in comparison with professional magnetic compasses are rigorous, but appear to assume that the errors are random, rather than having any sort of pattern.

The form and causes of the deviation (error) curve for a magnetic compass on a ship has been understood since the 19th century. In [4], Doerfler characterizes the shape of a deviation curve as having

- A constant term, due to misalignment of the compass
- One or more semicircular components, with a period of 360 degrees, and
- Quadrantal components, with a period of 180 degrees,

The various components may not be in phase, so the resulting overall deviation curve can be complex, as with some of those reported in this poster.

However, the important point is that a deviation curve represents reproducible errors for different orientations of a magnetic sensor: the errors are not simply random.

In the context of a smartphone compass, an additional, variable, component is relevant. The currents powering major components of a smartphone, such as the processor or screen, and even the current through the battery, will generate magnetic fields comparable to that of the earth.

3 EXPERIMENT DESIGN

The principle of swinging a compass is well-established for ships. The compass is sited in a known position, away from possible magnetic influences, surrounded by eight or more markers in known relative positions. Compass bearings are taken from the central position to each of the markers and compared with the known bearings; the difference between measured and known bearings is the compass error, or deviation, in the direction (heading) of the known bearing.

We emulated this setting on a cricket field. The central point was a plastic stool, with twelve fluorescent markers positioned at the edge of the field at roughly equal angular separations, and distances of 35 to 60 metres from the stool. For each experiment, the reference bearings were measured multiple times with a professional hand-bearing magnetic compass. The angles between the markers, from the central point, were checked by sextant, confirming that the magnetic compass bearings were not subject to measurable deviation. Calibrating the hand bearing compass against two other magnetic compasses confirmed an alignment error of approximately 2 degrees in the hand bearing compass. This is not unexpected in a fairly old device (~30 years) and was taken into account before calculating the deviation of the smartphone compasses.

Participants were asked to remove potentially magnetic objects from their person – such as keys, coins, spectacles, etc. – and to sit on the stool. Using their device and their chosen AR compass app, they measured the bearing to each marker in turn. Each participant completed four circuits of the markers: the first without calibrating their device's compass, and the next three following calibration.

Finally, following calibration, four additional apps were compared on a single device (a 2018 iPad) against the app used in the main experiment.

4 EXPERIMENTAL RESULTS

The three results for a single device were very close in all cases, differing by no more than 2 degrees for each marker, and usually less. Given that the precision of the compasses was 1 degree, this is a small error, and was eliminated by simple averaging.

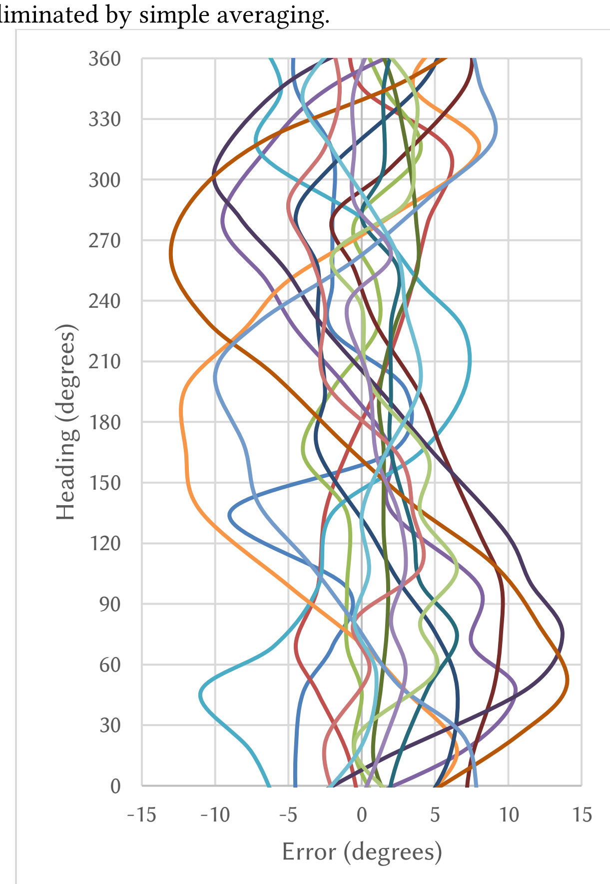

Figure 1: Deviation curves measured for 17 devices: 14 smartphones + 3 tablets. Curve for each device shows measured error against direction of compass (heading).

The difference between the observed bearings and those measured by the reference hand bearing compass, is the deviation of the compass in the device, as represented by the App used. The plot of deviation against the known bearings (to each marker) is the deviation curve for the device. The deviation curves for all 17 devices, after calibration, are shown in figure 1.

Even given the density of information in figure 1, some observations are possible:

First, as expected for magnetic sensors, the majority of deviation curves have a distinctive sinusoidal character. Some also have higher frequency components, if the devices may be subject to several sources of deviation.

Second, the amplitude of the deviation curves, although varying between devices, is at least a few degrees, and is over 10 degrees in several cases.

The data are summarized in table 1. The devices were grouped as tablets (all iPads of various ages), iPhones (again, a range of ages) and Android phones. Data are shown for both after and before calibration.

The relevant parameters are the misalignment, or offset error; the amplitude of the deviation curve; and the route mean square error (RMSE) after the curve has been re-centred to correct any misalignment error. Average figures are shown for each group of devices. Rather than the standard deviations, the maximum values give a clear indication of the severity of the errors observed.

Table 1: Average/maximum values for deviation measurements. All data values in degrees.

Device	*Misalignment*		*Amplitude*		*RMSE*	
	Mean	*Max*	*Mean*	*Max*	*Mean*	*Max*
Tablets (n=3)						
Calibrated	0.8	2	7.0	12	4.4	8.0
Uncalibrated	1.3	4	5.7	6.8	3.8	4.4
iPhones (n=6)						
Calibrated	2.9	5	3.8	6.0	2.2	3.9
Uncalibrated	27	96	24	118	19	99
Android Phones (n=8)						
Calibrated	0.9	-2	8.3	14	5.5	9.3
Uncalibrated	5.8	14	29	80	20	51
t-test: iPhone vs Android (calibrated)	0.0016		0.0081		0.0065	

The final row of table 1 shows the t-value for a two-tailed t-test for significant differences between the mean (calibrated) values for iPhones and Android phones. There were too few tablets in the sample for corresponding comparisons for tablets to be significant.

The second set of data, for different apps on the same device, is summarized in table 2 and figure 2. The "reference" app was used first, then four other apps, and finally the reference app again. The two (different!) deviation curves measured for the reference app are shown as thicker lines (red and blue) in figure 2.

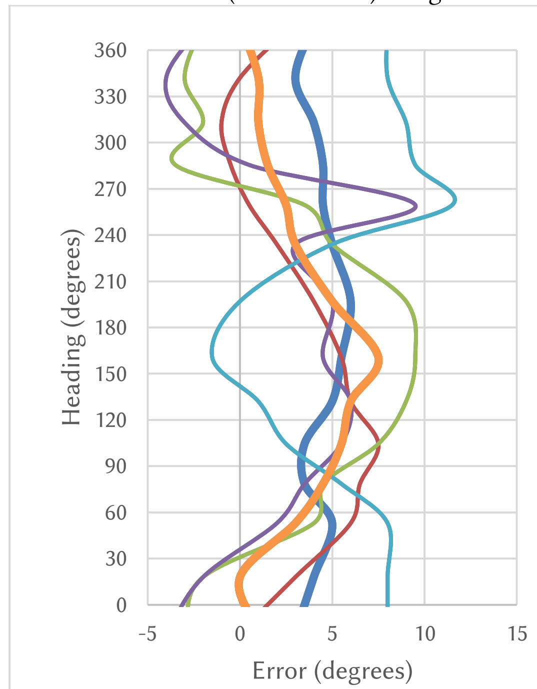

Figure 2: Deviation curves for different apps on a single device, all measured within one hour

Table 2 Comparison data for five apps on the same device. All error values in degrees.

App	*Misalignment*		*Amplitude*		*RMSE*	
	Mean	*Max*	*Mean*	*Max*	*Mean*	*Max*
Reference (1)	4.5		1.5		0.85	
Alternatives	3.7	5.5	6.0	6.75	3.9	4.7
Reference (2)	3.4		3.75		2.23	

5 COMMENTARY

The commentary on this experiment falls under several headings: scale of the deviation data; differences between iPhone and Android phones; the effect of calibration; and differences between apps on the same device.

5.1 The Scale of the Deviation Curves

For a "real" hand-bearing compass, used by mariners around the world, the expected accuracy is about 1 degree. Although compasses may suffer minor drift over time, there should be no observable deviation in the compass itself. The reading precision should be half a degree.

Ignoring for the moment the misalignment errors for the devices in this experiment, "calibrated" deviation curves with an amplitude of 4 or 5 degrees are worrying. At two kilometres, a 5 degree error corresponds to a displacement of more than 150 metres, which could be very serious if the device is being used for navigation.

However, these concerns are insignificant in comparison with those due to poor calibration.

5.2 iPhone versus Android Phones

Other authors have reported a differences between the magnetometers in iPhones and Android phones[7]. This difference is seen also in this experiment, and appears to be significant. However, some iPhones in this experiment were worse than some Android phones. A more extensive study would be needed to determine which brand of phone might be best, and under which circumstances. But "best" might still not mean "good enough for navigation"!

5.3 Calibration

Calibration can lead to dramatic improvements in the observed error curve. Uncalibrated deviation curves with an amplitude of more than 90 degrees, or misalignment offsets of over 90 degrees, do not inspire confidence.

However, applying the standard calibration technique of moving the device smoothly in a horizontal figure of 8 appears not always to be beneficial. For five of the devices – an iPad, 2 iPhones and 2 Android phones – the deviation curve was worse after calibration than before.

Also, calibration may not persist. In table 2 and figure 2, for the reference app on a single device, the deviation curve was noticeably different after less than one hour.

So, not only is calibration crucial – if it were possible to find a definitive, effective, calibration mechanism – but it appears also that it needs to be repeated far more often than any navigator would be likely to contemplate.

5.4 Differences between Apps on One Device

Not only are the deviation curves shaped differently, but they also have a rather larger amplitude than the reference app. Also, the reference app had a worse deviation curve when it was used a second time.

There would seem to be two effects: first, and most significantly, different apps interpret the output from the magnetometer differently. The calculations are complex - interpreting the magnetic field measurements on three axes, taking into account the aspect of the device, and also its geographical location, which determines the dip and the variation of the magnetic field. Perhaps algorithms to compute the heading should be standardized.

The second apparent effect is the "wearing off" of compass calibration. Since the calibration is in software, rather than any physical adjustments to the device, it also impacts directly on the calculations alluded to above.

The overall effect is to reduce even further any trust in smartphone compasses.

5.5 Other Observations

Several other observations were made that might be relevant for those developing mobile AR apps:

- For many smartphones, the AR compass app consumed a lot of battery capacity (up to 30%) in an experiment lasting just 15 minutes;
- It was very difficult to use an AR compass within 30 degrees of the azimuth of the sun;
- The availability of "pinch zoom" would have helped locate small objects in the camera view;
- Large, clear readouts are essential, in colours that contrast with the background;

6 CONCLUSIONS

The measurements presented in this poster raise questions about claims for AR smartphone compasses to be of "professional grade" or usable for navigation.

The issue seems to be that the allure of AR is overwhelming, and it is now available to general developers. Apps, ostensibly for navigation, are being developed. And yet the inherent inaccuracies of compasses, and the consequent risks of relying on the built in magnetometer, are not widely understood.

This suggests a need for the AR / Ubicomp community to take a lead in educating users that being able to build something that looks elegant and precise doesn't mean it is actually usable – or, indeed, safe to use.

REFERENCES

[1] Pavel Ahafonau, Spyglass, http://happymagenta.com/spyglass/ (accessed 10.6.19)

[2] Ronald T. Azuma, The Challenge of Making Augmented Reality Work Outdoors, In *Mixed Reality: Merging Real and Virtual Worlds.* Yuichi Ohta and Hideyuki Tamura (ed.), Springer-Verlag, 1999. Chp 21 pp. 379-390. ISBN 3-540-65623-5.

[3] J R Blum, D G Greencorn, J R Cooperstock 2013. Smartphone sensor reliability for augmented reality applications. In Mobile and Ubiquitous Systems: Computing, Networking, and Services (MobiQuitous 2012). p127-138.

[4] Ron Doerfler, Magnetic Deviation: Comprehension, Compensation and Computation (Part I) (2009), https://deadreckonings.com/2009/04/18/magnetic-deviation-comprehension-compensation-and-computation-part-i/ (accessed 12. 6.19)

[5] D. Krevelen and R. Poelman, A Survey of Augmented Reality Technologies, Applications and Limitations, The International Journal of Virtual Reality, Vol. 9, No. 2, 2010, pp. 1-20.

[6] Pocket Mariner, Compass Eye, https://pocketmariner.com/mobile-apps/compass-eye/ (accessed 10.6.19)

[7] Lucie Novakova, Terry L. Pavlis, Assessment of the precision of smart phones and tablets for measurement of planar orientations: A case study, Journal of Structural Geology,**97** (2017) pp 93-103, ISSN 0191-8141, https://doi.org/10.1016/j.jsg.2017.02.015

[8] G. Reitmayr et al., "Simultaneous Localization and Mapping for Augmented Reality," International Symposium on Ubiquitous Virtual Reality, Gwangju, 2010, pp. 5-8. doi: 10.1109/ISUVR.2010.12

[9] Hayato Shimizu, AR Compass Pro, https://appadvice.com/app/ar-compass-pro/1412221726 (accessed 10.6.19)

[10] S.J. Whitmeyer, E.J. Pyle, T.L. Pavlis, W. Swanger, L. Roberts, Modern approaches to field data collection and mapping: Digital methods, crowdsourcing, and the future of statistical analyses, Journal of Structural Geology (2018), ISSN 0191-8141, https://doi.org/10.1016/j.jsg.2018.06.023.