Augmenting Microsoft’s HoloLens with vuforia tracking for neuronavigation

Taylor Frantz¹,² ✉, Bart Jansen¹,², Johnny Duerinck³, Jef Vandemeulebroucke¹,²

¹Vrije Universiteit Brussel (VUB), Department of Electronics and Informatics (ETRO), Pleinlaan 2, B-1050 Brussels, Belgium
²imec, Kapeldreef 75, B-3001 Leuven, Belgium
³Vrije Universiteit Brussel (VUB), Department of Neurosurgery, Laarbeeklaan 101, 1090 Brussels, Belgium
✉ E-mail: taylor.frantz@vub.be

Published in Healthcare Technology Letters; Received on 16th August 2018; Accepted on 3rd September 2018

Major hurdles for Microsoft’s HoloLens as a tool in medicine have been accessing tracking data, as well as a relatively high-localisation error of the displayed information; cumulatively resulting in its limited use and minimal quantification. The following work investigates the augmentation of HoloLens with the proprietary image processing SDK Vuforia, allowing integration of data from its front-facing RGB camera to provide more spatially stable holograms for neuronavigational use. Continuous camera tracking was able to maintain hologram registration with a mean perceived drift of 1.41 mm, as well as a mean sub 2-mm surface point localisation accuracy of 53%, all while allowing the researcher to walk about a test area. This represents a 68% improvement for the later and a 34% improvement for the former compared with a typical HoloLens deployment used as a control. Both represent a significant improvement on hologram stability given the current state-of-the-art, and to the best of the authors knowledge are the first example of quantified measurements when augmenting hologram stability using data from the RGB sensor.

1. Introduction: Neurosurgery, a field heavily reliant on medical imaging for preoperative planning and perioperative navigation provides an excellent, though challenging, landscape for the exploration of augmented reality in medicine. Numerous authors have described the disadvantages of current neuronavigation paradigms [1-8]. Of paramount concern is that the attention of the physician is divided between the procedural workspace of the patient and the navigational workspace of the displays on which the data are shown. This requires one to mentally transform and relate the visualised two-dimensional (2D) imaging data into 3D physical space, often without the ability to see both at the same time. Augmented reality (AR), using a head mounted device (HMD), addresses both problems by directly rendering 3D models of anatomy, planning information, or other pertinent data into the physician’s field of view (FOV), aligned with the patient anatomy.

Research into the use of AR for neuronavigation has predominately focused on augmenting information displayed on an external screen, while initial registration and tracking have commonly been addressed using an outside-in approach, i.e. with the aid of external tracking hardware similar to existing commercial navigation systems. This trend is highlighted in the review paper from Guha et al. [1], where a majority of the 33 studies highlighted made use of some form of external tracking, whether for a hand-held camera coupled to an external monitor, or tablet-like device. In contrast, only four studies made use of an HMD, and only one of these was capable of self-pose estimation. This inside-out approach to tracking eliminates the need for external tracking hardware, reducing OR clutter, avoids line-of-sight problems and could prove to be considerably more cost-effective.

Microsoft’s HoloLens, when released in 2016, was one of the first headset devices to implement AR on a commercial scale and has provided the biomedical field with a unified platform for the development of AR in medicine. Since then, it has been a focus of discussion and research in the context of neuronavigation, as it allows for the potential marriage of the patient with their medical imaging data; a long overdue paradigm shift in the art. It contains an inertial measurement unit, one front facing depth camera and four flanking grey-scale cameras used to map the spatial surrounding. Proprietary algorithms perform a fusion of the sensor readings to achieve simultaneous localisation and mapping (SLAM), aimed at providing a robust approximation of scene and pose. As such, it is possible to display computer models in the HoloLenses’ optical system, (‘holograms’ only insofar as Microsoft double-think), and keep these holograms fixed, up to a certain accuracy, with respect to the physical space as the wearer changes their viewpoint and position.

The spatial alignment of data obtained from preoperative imaging to the patient anatomy is an essential component to any navigation system. The process is typically composed of two steps. During an initial registration step, the correspondence between the patient anatomy and computer model needs to be established. Next, that correspondence needs to be maintained, potentially even after the patient anatomy has moved, which we will refer to as tracking.

Research on the use of the HoloLens for medical interventions has suffered from the lack of access to the raw sensor readings of infrared (IR) and grey-scale cameras. This has resulted in the use of manual registration methods based on direct visual surface matching in lieu of more traditional automatic IR-based schemes [9, 10]. Furthermore, the perceived spatial stability of any placed hologram can only be as good as the headset localisation, and the inability to tune headset tracking for the close range, high-accuracy setting of surgery has resulted in tracking errors in the neighbourhood of ±6 mm [11-13].

Registration and tracking using HoloLenses’ centred spatial mesh and SLAM system, without access to exact sensor information by Xie et al. [14] indicated that the low-vertex density and surface bias of the mesh, as well as the uncertainty of the SLAM, make this approach unsuitable. Garon et al. [15] went as far as to attach an external depth camera to a HoloLens, thereby bypassing the data limitation. However, neither of them reported on accuracy.

The notion of using the HoloLenses’ RGB camera and image processing techniques for tracking has been suggested as a work-around, as this is an already established technique in machine vision research. Towards the tail-end of papers, the proprietary image processing SDK Vuforia has often highlighted a solution in this regard, as it contains object tracking libraries compatible with HoloLens development [9, 16, 17].

Vuforia’s image processing library allows HoloLens to track a known target image from the front facing camera’s coordinate system. Therefore the apparent world position of any object whose position is made dependent on that tracked target is then
Researchers from Duke University appeared to have used Vuforia for the tracking during a simulated external ventricular drain placement, however, no results or conclusion have been reported [18]. The use of RGB data from the front-facing camera in image processing has also been reported on by McDuff et al. [19] and Eckert et al. [20] who applied it towards ballistocardiography and object detection, respectively. Yet, to the best of our knowledge, no group has yet published quantitatively the advantages of augmenting HoloLens with object tracking information through the use of Vuforia, or other similar means.

In contrast to the media front by Microsoft concerning HoloLens as a tool within the medical field, the current body of published literature concerning its use in medicine is severely under quantified, with many publications (and media demonstrations) being of anecdotal quality. The following text expands on the current body of literature by quantifying the use of Vuforia’s SDK in a simulated neuronavigational workflow.

2. Methods

2.1. Segmentation and 3D modelling: Imaging data were provided by physicians from UZ Brussels. Ultimately a 3D printed skull phantom, for which computed tomography data was available, was chosen. Obtained digital imaging and communications in medicine (DICOM) images were segmented using 3D Slicer to provide masks for bone and neoplastic structures. Segmentation was accomplished by a mixture of thresholding and cut-grow techniques using the plug-in Fast Cut Grow. In total three segmentations were made; the skull, and two mass phantoms (Fig. 1).

Following segmentation, 3D slicer was used to construct computer models of each structure. Model geometry was then smoothed using a Laplacian filter over 25 iterations and then exported in STL format. Due to the rendering hardware limitations of the HoloLens, the models were subsampled significantly using the open source program MeshLab. Subsampling was accomplished using a quadric edge collapse decimation filter. The complete computer model, as well as 3D printed phantom, was named Sara (Fig. 2).

2.2. AR neuronavigation application: The personal distribution of the game development engine Unity (version 2017.2.0f3) was used to construct the main application which was to run on HoloLens. The decision to use an outdated Unity build, rather than the current 2018.1 build was due to known compatibility issues with the latest release (2017) of Microsoft’s Mixed Reality Toolkit SDK (MRTK).

All holographic content was placed in a single parent object, whose persistent spatial anchor could be repositioned by the researcher. The models were not given mass or collision boundaries, as this would make it impossible to place them inside a phantom target or person. A single overhead directional light was used in the scene. In addition to this, the detailed phantom models were illuminated with global ambient lighting, ensuring a realistic interventional hardware load. The application was deployed to HoloLens after which it was ensured that performance never fell below 59 frames per second.

2.3. Object tracking: The Vuforia SDK (Version 6.5.22) was used for target tracking. To attenuate fluctuations in Vuforia’s tracking output, the positional and rotational averages of 16 frames were taken using a sliding window; no appreciable delay or ill-effect was observed. During tracking, the model’s transform was updated as the product of the camera–target transform and target–phantom transform.

Phantom tracking was accomplished through Vuforia’s proprietary feature detection algorithms and a known RGB cylindrical target. Through comparison of extracted visual features on the tracking target and those of the computer model in memory, the 3D rigid transformation of the target can be estimated using a single camera. A priori knowledge of a transform between a well manually registered hologram and the phantom allowed for the consistent automatic manual registration of the hologram to its phantom target when using Vuforia. Continual tracking of the target allowed for this perceived registration to be maintained from any perspective.

Early on, plane image targets were found to be impracticable due to the line of sight limitations, and reduced tracking accuracy at high incidence angles between the target and camera. The implementation of a cylindrical target solved this by allowing the target to remain normal to the headset’s camera over 360°. To that end, the well performing provided stock Vuforia image tarnac was printed onto semi-mat photo paper at 1200 dots per inch. Measuring 165.7 mm in height and 235.6 mm in width, it was then wrapped around a cylinder whose diameter measured 75 mm and secured by thin double sided tape in such a manner where it would not be visible (Fig. 3b).

2.4. Phantom setup: To measure hologram accuracy, an A4 sheet of millimetre paper whose minor lines were of 2 mm spacing, and whose major lines were of 10 mm spacing was affixed to a coated expanded polystyrene support board. To this was then affixed through a non-permanent bonding agent the 3D printed Sara skull phantom. This assembly was then secured using common camera mounting hardware to the working end of a tripod. The pitch, yaw, and a roll of the tripod’s head, and by extension the phantom, were confirmed to be <1°. The distance between the floor and the XY plane of the millimetre paper was 130 cm.

To assure that the obtained data were comparable, fixed locations were chosen for repeatability. On the floor, marked points were placed at −90°, −45°, 0°, +45°, and +90° at the distances of 40 and 80 cm from the centre of the phantom, Fig. 3a.
These measurement positions were chosen to replicate the distance and range of movement a surgeon may experience during the planning phase of an intervention.

The phantom had been scanned and modelled with a series of radio-opaque fiducials placed at various surface points (Fig. 2). These were removed from the physical model, however, their exact placement was marked on the phantom itself.

2.5. Experimental conditions: Two test conditions were designed to test the efficacy of augmenting HoloLens with Vuforia. The control condition represents a typical device configuration, while the Vuforia test condition included target tracking. As the test lab is static, there was no need to update the spatial map during measurements. A detailed spatial mesh was made of the lab and saved to memory for retrieval.

2.6. Experimental metrics: Registration time. Quantified by recording the time required by the researcher to align the hologram with the phantom using programmed translational and rotational hand gestures. The perceived alignment of surface features was used as a criterion for proper registration.

Hologram drift. Quantified from the perceived translation of the hologram relative to the phantom in the coordinate system of the underlying millimetre paper. Translation in the millimetre paper’s Z direction was measured using a ruler (Fig. 4b).

Localisation accuracy. Quantified by the ability of the researcher to locate a surface point on the phantom based on holographic markers. The centre fiducial points of the three front markers, still visible on the hologram but not the phantom, are identified by a researcher using a fine tipped stylus. The radial distance from the known control point and the perceived centre is recorded (Fig. 4a).

2.7. Experimental methodology: Registration was performed once at the beginning of each trial with the researcher standing at 0° and 80 cm distance; this was either manual for the control condition or automatic in Vuforia condition. Only for control trials was registration time noted, as the automatic method was near instantaneous. From this position, both perceived hologram translation and surface point localisation accuracy were noted. The researcher then moved up to the 40 cm distance and repeated the measurements. Following this pattern of 80–40 cm distances, data was collected, in order, from the −45°, −90°, +45°, and +90° positions. Measurement data obtained from both distances were then aggregated to a single angle.

Multiple iterations of each test condition were performed by an engineering researcher intimately familiar with the program and controls, resetting the program between use. By iterating each test condition it was possible to smooth variance in data, and maintaining a single researcher helped to ensure a consistent bias. This is important as the exact registration of holograms manually is never guaranteed, and translational drift errors rely on the subjective opinion of the researcher. It should be noted that this subjectiveness results from a difficulty to accurately judge matter-hologram boundaries, discussed more in depth later. The decision to reset the program between trial runs was to assure a cold-start, clearing out any data which may have affected future results, and better simulating a clinical use case. Cumulatively data was collected and averaged from 19 trials into the two test conditions.

3. Results: Manual registration times yielded a mean time of 95 s, with a minimum and maximum time of 51 and 165 s, respectively, summarised in Table 1. The mean perceived holographic drift of the control condition was 4.39 mm. The Vuforia test condition yielded 68% improved results, with a mean perceived holographic drift of 1.41 mm; summarised in Table 2. The reduction in perceived drift was statistically significant at all measurement angles, except for 0°, per Student’s t-tests at 5% confidence.

The standard error of the mean (SEM) among all trials for the control condition was 1.29 mm. The Vuforia test condition yield 48% more consistent results with a SEM of 0.67 mm.

Surface point localisation in the control condition showed a mean error of 5.43 mm. The Vuforia test condition yields a 65% reduction, with a mean error of 1.92 mm. This resulted in a 34% improvement in sub 2-mm accuracy versus the control; summarised in Table 3. All three points showed a statistically significant improvement versus the control per Welch’s t-test at 5% confidence.

4. Discussion

4.1. Registration time: The obtained data highlights the extreme variability in the ability of the user to register, as best they may, a hologram based on surface features alone. This observation is consistent with the ranges of registration time reported in the literature, summarised in Table 1. When comparing the obtained manual registration times with those of Rae et al. [9], when matched like-for-like for an experienced user, our results

Table 1 Comparison of published manual hologram registration times

| Study            | Range, s | Mean, s |
|------------------|----------|---------|
| Proposed method  | 51–165   | 95      |
| Pratt et al. [10]| 60–120   | NR      |
| Rae et al. [9]   | 135–219  | 172     |

NR: not reported.
A statistical improvement at 0° is not surprising as the hologram perceived drift between experimental conditions. The lack of any difference in registration times can be attributed to a number of factors, particularly user experience and control schemes.

4.2. Perceived drift: Table 2 summarises the difference in mean perceived drift for each measurement angle.

Table 2 Change in mean perceived drift for each measurement angle

| Condition | −90° | −45° | 0° | +45° | +90° | Mean | σ | SEM |
|-----------|------|------|----|------|------|------|---|-----|
| control, mm | 6.27 | 3.59 | 0.62 | 4.60 | 6.90 | 4.39 | 3.34 | 1.29 |
| vuforia, mm | 0.83 | 1.46 | 1.24 | 0.08 | 3.42 | 1.41 | 1.08 | 0.67 |
| Δ | −87% | −59% | 0% | −98% | −50% | −68% | −68% | −48% |

*aShown to be insignificant.
*σ, standard deviation.
*SEM, standard error of the mean.

Table 3 Comparison of surface localisation results

| Study | <2 mm | 2–5 mm | 5–10 mm | >10 mm |
|-------|-------|--------|----------|--------|
| Rae et al. [9] | 50% | 0% | 50% | 0% |
| Proposed method | Matched* | 87% | 13% | 0% | 0% |
| Control | 19% | 34% | 40% | 7% |
| Vuforia | 53% | 40% | 7% | 0% |

*aResults matched to reflect measurements taken from a starting position, without introducing movement.

The difference in registration times can be attributed to a number of factors, particularly user experience and control schemes. Dissatisfied by those provided in the MRTK, and taking inspiration from the model controls inside of Unity, the iterative collaboration between the engineering team and neurosurgical staff resulted in the consistent transform used for registration. In contrast to this, the less consistent results from the control are most likely due to the difficulty in perceiving exact hologram–matter interactions during manual registration, resulting in poor depth judgment between the two. This is most pronounced when a hologram is registered from a single perspective. Fig. 6 highlights this effect.

4.3. Point localisation: Rae et al. reported on similar metrics used in their experiments, and a comparison can be found in Table 3. It is worth noting that due to the known holographic drift, their group measured results from only a single position after performing manual registration of the hologram. In consideration of this, we have included a selection of matched results which were taken following registration from the 0° position. It can be seen that before introducing movement associated holographic drift errors, there is a marked improvement versus literature.

Fig. 5 Control schemes used for manual transforms of models
a Translation: allows the user to axially move the model in either the model or world coordinate system using the ‘pinch and drag’ command
b Rotation: Allows the user to highlight any axis of the model coordinate system with their gaze and ‘pinch and drag’ to rotate along with it

Fig. 6 Perception of hologram–phantom relationship from a single view
a Similarity of an observed hologram relative to its phantom from a single perspective
b Two manual registrations highlighting manual registration error. Top row: view from standing; bottom row: view from head on. Variance in vertical height between each registration may be seen
This trend continued when analysing all the collected data. Implementing Vuforia tracking yielded similar or better results accuracy across all threshold categories when compared with the published literature, while still allowing unabated movements within the test parameters.

4.4. Remarks: One of the challenges when using the active RGB camera tracking was the absolute need to maintain a line of sight between the target and the phantom. This, while not prohibitively difficult at greater distances from the phantom did become so at an arm’s length distance, where the target, though within visual line of sight, was outside the frustum of the camera. This may be remedied by better placement of the tracking target, use of multiple targets at various key points, and/or rethinking its form. For example, using the RGB camera to track 3D printed objects rather than RGB targets, a feature Vuforia already supports. As for the practicality of such a large marker, Vuforia, or alternative image processing algorithm(s), may effectively be used to maintain an established registration through a secondary reference, the initial registration possibly obtained through the tracking of a smaller stylus and surface matching. This is indeed how workflows are often structured in commercial markerless neuronavigation systems; however, unlike these bulkier systems, HoloLens could easily be deployed beyond the walls of an OR.

Limiting manual registration to a single perspective was arguable to the detriment of accuracy in the control, however, it does highlight the necessity to adopt automatic registration into AR platforms. This is particularly true in situations where the wearer has neither the space nor time to manually adapt the registration from numerous perspectives, especially if the registration from one perspective is disturbed due to the tracking inaccuracies of the headset as he/she changes perspectives.

In the scope of medicine, the authors feel that the use of any spatial mapping information is superfluous, as its principal role is to allow holograms to interact with the environment; an unnecessary criterion when one wants to place a hologram inside a head rather than on top of it. A decision to not load, generate, or update any spatial meshing data may therefore further free up already limited hardware resources for future visualisation, tracking, or registration tasks.

With the recent release of Microsoft’s beta Redstone firmware update for HoloLens, the prospect of using data collected from the front facing depth camera becomes reality. Not only does this camera have a greater FOV than the RGB camera, 120° versus 48°, but it also allows for tracking of more familiar, and already in place, instruments in the OR through IR reflectivity.

Having been released in 2016, HoloLens has begun to look outdated versus the competition from the last few years, the AR industry has not stood idly by, and recent HMD releases from ODG, Meta, and Magic Leap represent significant improvements over HoloLens. Of particular interest are those systems whose optics promise greater depth perception. Still, it is expected that Microsoft will release the replacement for HoloLens, codenamed Sydney, sometime in 2019, the delivery of which will keep the workflow and technologies used through this Letter relevant for years to come.

5. Conclusion: It was demonstrated that by augmenting HoloLens with Vuforia’s RGB target recognition for inside out tracking, significantly greater hologram stability could be achieved than without; thus alleviating one of the greatest hurdles for HoloLens as a tool in neuronavigation. Furthermore, the greater variance between trials obtained when performing registration manually, even by those experienced with HoloLens, highlights the need for more effort placed on automatic workflows, and a re-examination of current practices among many professionals.

6. Funding and declaration of interests: None declared.

7 References
[1] Guha D., Alotaibi N.M., Nguyen N., et al.: ‘Augmented reality in neurosurgery: a review of current concepts and emerging applications’, Can. J. Neurol. Sci., 2017, 44 (3), pp. 235–245. ISSN 03171671. doi: 10.1017/cjn.2016.443
[2] Meola A., Cutolo F., Carbone M., et al.: ‘Augmented reality in neurosurgery: a systematic review’, Neurosurv. Rev., 2017, 40 (4), pp. 537–548. doi: 10.1007/s10143-016-0732-9
[3] Inoue D., Cho B., Mori M., et al.: ‘Preliminary study on the clinical application of augmented reality neuronavigation’, J. Neurol. Surg. A. Cent. Eur. Neurosurg., 2013, 74 (2), pp. 71–76. doi: 10.1055/s-0032-1334145
[4] Cutolo F., Meola A., Carbone M., et al.: ‘A new head-mounted display-based augmented reality system in neurosurgical oncology: a study on phantom’, Comput. Assist. Surg., 2017, 40 (4), pp. 537–548. doi: 10.1080/20499322.2017.1358490
[5] Mascielli J.R., Schlachter L., Charron A.G., et al.: ‘Navigation-linked heads-up display in intracranial surgery: early experience’, Oper. Neurosurg., 2017, 15 (2), pp. 184–193. doi: 10.1093/ons/opx205.
[6] Tepper O.M., Rudy H.L., Lefkowitz E.A., et al.: ‘Mixed reality with HoloLens: where virtual reality meets augmented reality in the operating room’, Plast. Reconstr. Surg., 2017, 140 (5), p. 1066–1070. doi: 10.1097/PRS.00000000000003802
[7] Chauvet P., Collins T., Dégibz C., et al.: ‘Augmented reality in a tumor resection model’, Surg. Endosc.: Other Inter. Tech., 2018, 32 (3), pp. 1192–1201. doi: 10.1007/s00464-017-5791-7
[8] Orriinger D.A., Golby A., Jolesz F.: ‘Neuronavigation in the surgical management of brain tumors: current and future trends’, Expert. Rev. Med. Devices, 2012, 9 (5), pp. 491–500. doi: 10.1586/err.12.48
[9] Rae E., Lasso A., Holden M.S., et al.: ‘Neurosurgical burr hole placement using the Microsoft HoloLens’, March 2018. doi: 10.1117/12.2293680
[10] Pratt P., Ives M., Lawton G., et al.: ‘Through the HoloLens looking glass: augmented reality for extremity reconstruction surgery using 3D vascular models with perforating vessels’, Eur. Radiol. Exp., 2018, 2 (1), pp. 2. doi: 10.1007/s41747-017-0033-2
[11] Auvinet E., Galna B., Aframian A., et al.: ‘O100: validation of the precision of the Microsoft HoloLens augmented reality headset and hand motion measurement’, Gait Posture, 2017, 57, pp. 175–176. doi: 10.1016/j.gaitpost.2017.06.353
[12] Vassallo R., Rankin A., Chen E.C.S., et al.: ‘Hologram stability evaluation for Microsoft HoloLens’, Proc SPIE, 2017, 10136, p. 1013614. doi: 10.1117/2017.101365
[13] Liu Y., Dong H., Zhou L., et al.: ‘Technical evaluation of HoloLens for multimedia: a first look’. IEEE Multimdia, 2018, pp. 1–7
[14] Xie T., Islam M.M., Lumsden A.B., et al.: ‘Holographic iRay: exploring augmentation for medical applications’, Adjunct Proc. ISMAR-Adjunct 2017, 2017, pp. 220–222. doi: 10.1109/ISMARAdjunct.2017.73
[15] Garon M., Boulet P-O., Dorion J-P., et al.: ‘Real-time high resolution 3D data on the HoloLens’, doi: 10.1109/ISMAR-Adjunct.2016.0073
[16] Kuzhtagaliyev T., Clancy N.T., Janatka M., et al.: ‘Augmented reality needle ablation guidance tool for irreversible electroporation in the pancreas’, 2018. doi: 10.1117/12.2293671. Available at http://arxiv.org/abs/1802.03274
[17] Graham P.: ‘Scopis announces holographic navigation platform using HoloLens for open and minimally-invasive spine surgery’, 2017. Available at https://www.vrfocus.com/2017/05/scopis-announces-holographic-navigationplatform-using-hololens-for-open-and-minimally-invasive-spine-surgery/
[18] Manke K.: ‘Brain surgery may get a bit easier, with augmented reality’, 2016. Available at https://today.duke.edu/2016/10/brain-surgery-may-get-bit-easieraugmented-reality
[19] McDuff D., Hunter C., Gonzalez-Franco M.: ‘Pulse and vital sign measurement in mixed reality using a HoloLens’. Proc 23rd ACM Symp. on Virtual Reality Software and Technology – VRST ’17, Gothenburg, Sweden, 2017, pp. 1–9. doi: 10.1145/3139131.3139134. Available at http://dl.acm.org/citation.cfm?doid=3139131.3139134
[20] Eckert M., Blex M., Friedrich C.M., et al.: ‘Object detection featuring 3D audio localization for Microsoft HoloLens’. Proc. 11th Int. Joint Conf. on Biomedical Engineering Systems and Technologies, 2018, pp. 555–561. doi: 10.2200/S00665505550561