# 3D Virtual Tracing and Depth Perception Problem in Mobile AR



**Leo Gombač**
**Klen Čopič Pucihar**
**Matjaž Kljun**
University of Primorska, FAMNIT
6000 Koper, Slovenia
leo.gombac@student.famnit.upr.si
klen.copic@famnit.upr.si
matjaz.kljun@upr.si

**Paul Coulton**
Imagination, Lancaster University
Lancaster, UK
p.coulton@lancaster.ac.uk

**Jan Grbac**
University of Primorska, FAMNIT
6000 Koper, Slovenia
jan.grbac@student.famnit.upr.si

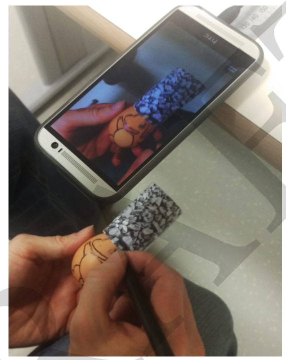

**Figure 1:** Virtual tracing on an egg.


CHI'16 Extended Abstracts, May 07-12, 2016, San Jose, CA, USA
ACM 978-1-4503-4082-3/16/05.
http://dx.doi.org/10.1145/2851581.2892412.

## Abstract
Mobile Augmented Reality is most commonly implemented using a camera and a flat screen. Such implementation removes binocular disparity from users' observation. To compensate, people use alternative depth cues (e.g. depth ordering). However, these cues may also get distorted in certain AR implementations, creating depth distortion which is problematic in situations where precise hand interaction within AR workspace is required such as when transcribing augmented instructions to physical objects (e.g. virtual tracing — creating a physical sketch on a 2D or 3D object given a virtual image on a mobile device). In this paper we explore how depth distortion affects 3D virtual tracing by implementing a first of its kind 3D virtual tracing prototype and run an observational study. Drawing performance exceeded our expectations suggesting that the lack of visual depth cues, whilst holding the object in hand, is not as problematic as initially predicted. However, when placing the object on the stand and drawing with only one hand (the other is used for holding the phone) their performance drastically decreased.

## Author Keywords
Depth perception; depth ordering; virtual tracing; 3D virtual tracing;

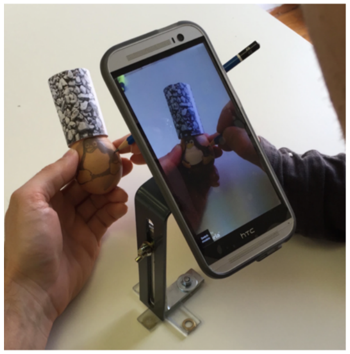

**Figure 2:** Setup A – participants sit at the table and hold the Easter egg in hand whilst the mobile device is fixed on a stand.

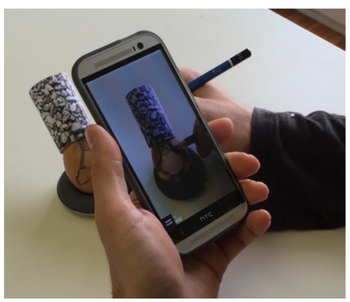

**Figure 3:** Setup B – participants sit at the table holding a mobile device in hand while Easter egg is placed on an egg stand.


## ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;


## Introduction
Mobile Augmented Reality (AR) is most commonly implemented using one camera and a flat screen – e.g. on a tablet or a smartphone. Although, viewing a three dimensional world on such devices removes binocular disparity from user's observations. To compensate, people use alternative depth cues (e.g. motion parallax, depth ordering, etc.). However, when implementing AR systems these cues may get distorted and may result in depth distortion – difficulty of perceiving spatial relationships between real and virtual objects in the observed AR scene. As such, the depth distortion is considered one of the most common perceptual problems of AR [14].

One example of an AR system where depth distortion is problematic are x-ray visualizations where users attempt to visualize virtual objects that are positioned behind the real world objects. Cases include AR supported maintenance tasks and AR technical manuals for which precise depth perception is vital [20]. Another important instance are dynamic AR workspaces where augmented instructions need to be transcribed to physical objects and require precise hand and/or tool interaction within AR workspace. The potential uses of such AR applications are vast and include: (i) remote expert assistance where expert guides user's actions by remotely augmenting a physical object (e.g. AR supported maintenance and first aid), (ii) AR assisted sculpting or modelling where augmented instructions guide users in shaping models from clay like materials, and (iii) virtual tracing (Figure 1) where the user attempts to sketch the observed virtual contour framed on a 2D surface (e.g. as in [5]) or a 3D object (e.g. pottery).

In all of the above cases, when the users' hands and augmented content are introduced to the AR scene, the rendering of augmented content with the correct depth order becomes very difficult, as it requires real time scene reconstruction. While the problem of depth distortion is expected to be notable in the case of transcribing augmented instructions to a 2D surface, such as paper (where users' hands are navigating on a 2D plane), it is our belief that the problem only exuberates when transcribing information to a 3D objects (where user's hands are navigating on a 3D surface).

It is thus worthwhile investigating the depth perception problem in the context of transcribing augmented instructions to physical objects in detail. Particularly, due to the number of aforementioned examples with large community bases (e.g. crafting communities) where AR could provide benefits. Based on the Google image search for "pottery painting" or "Easter egg painting" a large number of people are interested in such activities. Virtual tracing could support expert users with complex contours and improve drawing abilities of novice users. Therefore, we decided to explore the issue of depth perception on a case of virtual tracing interaction on 3D objects. For this purpose, we built a 3D virtual tracing prototype. The main contribution of this paper is a feasibility study for a 3D virtual tracing prototype. The study demonstrated drawing performance, which exceeded our expectations, suggesting that the lack of visual depth cues is not as problematic as initially expected.

**Figure 4**: A a printed marker was rolled up and placed inside the cup. This simple approach works as long as the walls of the cup are cylindrical in shape. Markings are a result of drawing “blocks” virtual contour (see Figure 6).

**Figure 5**: A printed marker was wrapped around cylindrically shaped object and glued on top of the egg. Markings on the egg are a result of drawing “animal” virtual contour (see Figure 7).

However, it is still an issue before the pen touches the object being drawn on, and it is particularly noticeable when object being drawn on is not held in hand.

## Related work

*Perception of depth*

A large body of research has focused on head-mounted displays in the area of depth perception for AR [8, 22, 15, 12]. Contrary to mobile devices, such systems are not readily available to consumers, thus mobile AR is currently the platform of choice for consumer oriented AR applications. Even on mobile devices the depth perception problem has already been identified [6, 7]. To solve this problem, alternative depth cues have been studied such as motion-parallax [11, 1] and object depth ordering [14]. Nevertheless, when mixing AR and moving physical objects the depth ordering may get distorted, and when the observed scene (through a mobile device screen) is in close proximity to the observer (0-2m), these alternatives are not accurate enough and binocular disparity still provides the most accurate depth judgment [4].

In order to reintroduce binocular disparity to mobile AR researchers utilized stereoscopic display technology and dual camera capture [13, 16, 23]. The results are promising; however, devices with such abilities are not readily available even though they were introduced to the mass market in 2011 (e.g. Nintendo 3DS1, HTC EVO 3D mobile device). Additionally, 3D displays technology on mobile devices mainly utilize parallax-barrier technology which limits observer’s point-of-view (POV) to small angle and distance variations [6].

*Augmented Reality Sketching*

Virtual tracing is an alternative to traditional methods for supporting sketching such as (i) the template approach incorporating a transparent drawing surface (e.g. tracing paper on top of the desired contour) or a stencil cut (e.g. following the defined contour lines of a shape), or (ii) carbon paper placed under the image and on-top of the drawing surface such that when the users exerts pressure on the image it is transferred to underlying surface.

Compared with traditional methods, virtual tracing has a clear advantage in that it does not require the physical production of sketching aids. Producing such aids can pose a difficult task when one desires to draw on large 2D formats (e.g. wall paintings) or when one desires to trace draw on 3D objects (e.g. pottery, Easter eggs, etc.). In the case of virtual tracing, the drawing size is only limited by camera’s pose tracking capabilities, and its ability to reconstruct a 3D world.

The magic lens paradigm has proved popular for supporting virtual sketching [9, 21, 17, 23] where the lens acts as a transparent glass plane revealing an augmented scene behind the pane [3]. However, its research predominately focused on complementing physical sketches and not on supporting in-situ sketching through virtual tracing. The main challenge for this and other sketching tasks is to provide accurate and robust pose tracking without disrupting the sketching process.

An alternative to camera pose tracking is to remove the need for it by placing the device and target object at a fixed position, as in the case of a virtual mirror or camera sketcher. Both are seen as inappropriate for 3D virtual tracing as they prevent users to move and rotate the object to a position that would be best suited

**Figure 6**: Example of cup drawing content trailed on the study.[1]

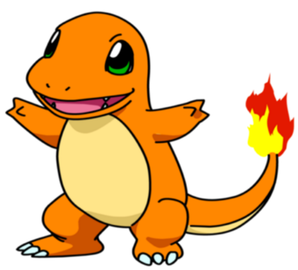

**Figure 7**: Example of Easter egg drawing content trailed on the study.[2]

[1] Image Courtesy of CharManDeer at http://charmanderrumble.wikia.com/wiki/File:004_Charmander_OS1.png

[2] Image courtesy of Toby at http://www.tandhblog.co.uk/2013/09/how-to-add-3d-blocks-to-your-graffiti-designs/

for drawing (due to the change in geometry, as this position continuously changes). Additionally, the virtual mirror captures the back side of the object which is not easily accessible with a pen.

In our previous work on virtual tracing we focused at mitigating the effect of camera pose tracking on a 2D virtual tracing experience by implementing a dual camera magic lens where the front-facing camera was used for pose tracking and the back facing camera for augmentation. Observational user evaluation showed that virtual tracing on 2D objects is possible and that the dual camera approach can significantly contribute to the user experience [5].

To our knowledge, there has been no research done on the 3D virtual tracing and depth perception problem. In this paper we present the first AR prototype for virtual tracing on 3D objects and present an observational study of its feasibility.

## 3D Virtual Tracing Prototype design rationale

The prototype presented is designed to support physical sketches on 3D objects using virtual tracing (see Figure 1). The core challenge of implementing this prototype was camera pose tracking. Geometry and feature based tracking [19, 10] of the object are not possible in virtual tracing due to hand interaction, which constantly occludes the object being tracked.

Another possibility is to use the marker. In our 2D study, the results showed that sketching is easier when user does not have a possibility to occlude the marker, hence we used both the front-facing camera for tracking the marker placed above the user and the back-facing camera for rendering the virtual contour [5]. To overcome the complexity of pose tracking 3D objects we decided on cylindrical markers placed on top of selected objects so users will not be able to occlude them. To achieve this, we limited ourselves to small 3D objects (such as coffee cups and Easter eggs) in order to keep both the object and the marker in camera's field of view when sketching (see Figure 4 and 5).

In the case of the teacup, marker placement is easily achieved. The user only needs to roll up the printed target and place it into the teacup as seen in Figure 4 (assuming that the cup is of cylindrical shape). In the case of the Easter egg the printed target needs to be wrapped around arbitrary cylindrical item and attached on the egg using a suction device or mild adhesive as seen in Figure 5.

## Observational study

Observational evaluation of the prototype was conducted in an informal setting. We invited 3 participants from our computer science department to use the prototype and produce a 3D graffiti contour (see Figure 6) on a teacup and a cartoon character (see Figure 7) on an Easter egg. During initial experimentation with the prototype we realised that immediate creation of the final product whilst looking through the phone's camera is difficult. Thus, we decided to split the drawing process for this observational study into two steps – sketching with pencil first and painting with permanent markers afterwards – in order to improve drawing performance.

The pottery painting is performed either on a pottery placed on a stand (e.g. a rotary stand for round objects or a canvas like stand for plates) or by holding pottery

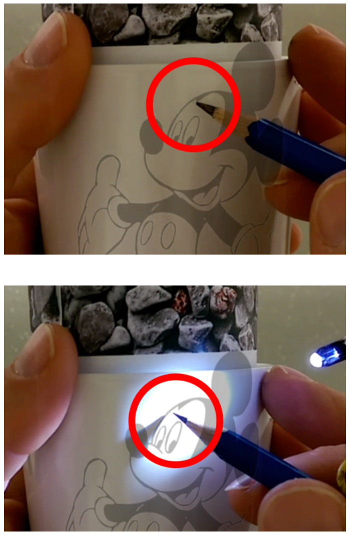

**Figure 8**: From visual depth cues on top image it is not possible to understand if the pencil has touched the surface or how far away it is from it. After introducing natural shadows by turning on LED light mounted on a pencil, it becomes obvious that the pencil hovers above the surface. When pencil comes closer to the surface the distance between the pencil tip and its shadow becomes shorter.

in one hand. These two setups are also very likely to occur in aforementioned cases of transcribing augmented instructions to physical objects (remote expert assistance, assisted sculpting or modelling) and other similar activities, because it is not always possible to hold the object in one hand and draw with the other (e.g. because of object size, setup, accidental smear of wet paint, and availability of the other hand for interaction). For these reasons, we decided to try two setups. In setup A participants sat at the table and held the object in hand whilst the mobile device was fixed on a stand — either mounted on the edge of the desk (see video) or placed on the table (Figure 2). After completing both designs (egg and teacup) in setup A, participants were asked to repeat the same using setup B where participants held the mobile device in hand whilst the object was placed on the stand (Figure 3).

In setup A all participants successfully completed both designs within 5 minutes we set as maximum time limit for completing the task. They were also satisfied with their results (Figure 4 and 5) and expressed the wish to use the system in the future. The two stage drawing — the sketch planning stage (drawing with a graphite 2B pencil) and the finalizing stage (drawing with permanent colours) — was also welcomed. During the finalizing stage, the participants checked their drawing against the virtual contour even though they already had the sketch on the object. When asked what the most difficult part of the task in setup A was, the participants mentioned the slippery surface of the teacup and the difficulty of drawing on curved surface.

In setup B, participants highlighted a problem when the contact with surface was lost (pencil moved away) and they needed to land the pencil on the object surface again. This was also observed by researchers when participants had more difficulties landing the pencil on the surface in setup B than in setup A. Even in the latter, users needed some initial adjustment to the system to continue drawing normally. However, in the former adjusting took longer and each time the contact with the surface was lost the problem was reintroduced.

## Discussion & conclusion

Our hypothesis that the depth perception problem would significantly affect virtual sketching on 3D objects has arisen from previous studies. During a 2D virtual tracing session we observed users experiencing some depth perception problem while sketching on a 2D plane [5]; although, we have not reported it in the paper since users have still been able to complete all the sketching tasks successfully and just needed some initial adjustment. Thus, the aim of the observational feasibility study presented in this paper was to find out how problematic the depth distortion really is when drawing on 3D objects.

The results of the drawing sessions with participants have proved more positive than expected in setup A (holding the object in hand). Despite our expectations, when users used both hands (one for holding the object and the other for drawing) they managed to finish the contours without significant problems, which was not the case when only one hand was in use (see Figure 3).

There are several possible reasons why this was the case. One potential reason for this was reducing the depth perception problem by proprioception sense in case of holding the object in hand. Among other things, proprioception helps one understand the relative

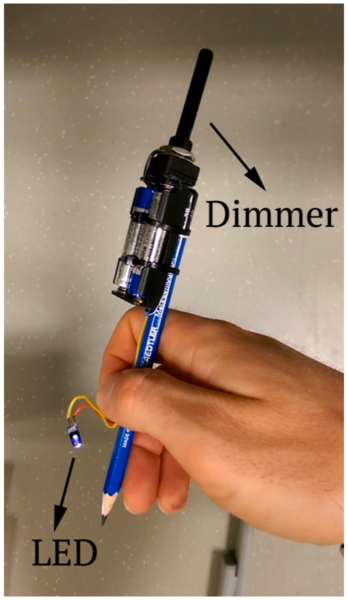


**Figure 9**: Pencil with LED light.

position of neighbouring body parts to each other [18]. Our brain compensates for visual information of incorrect depth ordering and binocular disparity with the input from proprioception, which provides sensory input about where one hand (and its fingers) is positioned in relation to the other hand (and its fingers) and can improve depth perception.

Besides depth perception problem, other factors could have influenced the drawing performance. For example, the haptic feedback and affordances of the object held in one hand as well as haptic feedback of the pressure of the pen touching the surface of the object (which is felt by both hands) could have assisted users in sketching. In the condition when users had to hold a phone in one hand while trying to draw on an object with the other hand the haptic feedback of the pressure of the pen pressed on the object was not as strong or effective and affordances of the object were lost. Holding the phone in one hand while focusing on the drawing on another object could have also contributed to performance drop. Moreover, fixing the object in place limited users' options of optimal positioning forcing them to draw in uncomfortable poses.

However, the role of depth perception might have been limited once the pencil has landed on the object because sketching (as done in our study) required small movements and did not require drastic changes possibly affecting depth perception. How important depth is for both of study conditions remains to be answered in future studies. To better understand the role of depth perception and the role of holding the object or phone in hand while drawing or transcribing augmented instructions to physical objects we plan to run three future studies. For understanding the role of holding the object on tracing we plan to study real tracing on 3D objects whilst holding the object in hand or not; for understanding the role of holding the phone on virtual tracing we plan to compare setups of holding a phone or having it on a stand (in both cases not holding the object). In the third study we plan to explore depth perception and how natural depth cues such as natural shadows shown in Figure 8 could be used to improve depth perception — even if just in the initial phase of landing the pen on the object. We already built a "natural shadow" prototype (see Figure 9) after observing problems of depth distortion with one hand interaction in the context of 3D virtual tracing. When the pencil tip comes closer to the surface of the object, the distance between the tip and its shadow produced by LED mounted on the pencil shortens. It thus provides precise real-time feedback to the user about the distance from object (see Figure 8).



## References

1. Kevin W. Arthur, Kellogg S. Booth, and Colin Ware. 1993. 3d task performance for fish tank virtual worlds. ACM Transactions on Information Systems (TOIS), 11, 3, pp. 239-265.
2. Matthias Berning, Daniel Kleinert, Till Riedel, and Michael Beigl. 2014. A study of depth perception in hand-held augmented reality using autostereoscopic displays. In Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '14), 93-98. http://dx.doi.org/10.1109/ISMAR.2014.6948413
3. Eric A. Bier, Maureen C. Stone, Ken Pier, William Buxton, and Tony D. DeRose. 1993. Toolglass and magic lenses: the see-through interface. In Proceedings of SIGRAPH, 73–80. http://dx.doi.org/10.1145/166117.166126

4. James E. Cutting. 2003. Reconceiving perceptual space. In Looking into pictures: An interdisciplinary approach to pictorial space, Heiko Hecht, Robert Schwartz, and Margaret Atherton (Eds.) 215–238. MIT Press.
5. Klen Čopič Pucihar, Jens Grubert, and Matjaž Kljun. 2015. Dual Camera Magic Lens for Handheld AR Sketching. In Human-Computer Interaction (INTERACT '15), pp. 523-527. http://dx.doi.org/10.1007/978-3-319-22723-8_53
6. Klen Čopič Pucihar, Paul Coulton, and Jason Alexander. 2013. Creating a stereoscopic magic-lens to improve depth perception in handheld augmented reality. In Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services (MobileHCI '13), page 448. http://dx.doi.org/10.1145/2493190.2494660
7. Anamika Dey, Graeme Jarvis, Christian Sandor, and Gerhard Reitmayr. 2012. Tablet versus phone: Depth perception in handheld augmented reality. In IEEE International Symposium on Mixed and Augmented Reality (ISMAR '12), pages 187–196. http://dx.doi.org/10.1109/ISMAR.2012.6402556
8. Singh, Gurjot, J. Edward Swan II, J. Adam Jones, and Stephen R. Ellis. 2010. Depth judgment measures and occluding surfaces in near-field augmented reality. In Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization (APGV '10), page 149, ACM Press. http://dx.doi.org/10.1145/1836248.1836277
9. Nate Hagbi, Raphaël Grasset, Oriel Bergig, Mark Billinghurst, and Jihad El-Sana. 2010. In-place sketching for content authoring in augmented reality games. In Proceedings of the Virtual Reality Conference (VR '10), 91-94. http://dx.doi.org/10.1109/VR.2010.5444806
10. Pengfei Han, and Gang Zhao. 2015. CAD-Based 3D objects recognition in monocular images for mobile augmented reality. Computers & Graphics. http://dx.doi.org/10.1016/j.cag.2015.05.021
11. Alex Hill, Jacob Schiefer, Jeff Wilson, Brian Davidson, Maribeth Gandy, and Blair MacIntyre, B. 2011. Virtual Transparency: Introducing Parallax View into Video See-through AR, In Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '11), 239-240. http://dx.doi.org/10.1109/ISMAR.2011.6092395
12. Adam Jones, Edward Swan II, Gurjot Singh, Eric Kolstad, and Stephen R. Ellis. 2008. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In Proceedings of the 5th symposium on Applied perception in graphics and visualization (APGV '08), page 9. http://dx.doi.org/10.1145/1394281.1394283
13. Frederic Kerber, Pascal Lessel, Michael Mauderer, Florian Daiber, Antti Oulasvirta, and Antonio Krüger. 2013. Is autostereoscopy useful for handheld AR? In Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia (MUM'13), 1–4. http://dx.doi.org/10.1145/2541831.2541851
14. Ernst Kruijff, J. Edward Swan II, and Steven Feiner. 2010. Perceptual Issues in Augmented Reality Revisited. In Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '10), 3–12. http://dx.doi.org/10.1109/ISMAR.2010.5643530
15. Kytö, Mikko, Aleksi Mäkinen, Timo Tossavainen, and Pirkko Oittine. 2014. Stereoscopic depth perception in video see-through augmented reality within action space. Journal of Electronic Imaging 23, 1: 011006. http://dx.doi.org/10.1117/1.JEI.23.1.011006
16. Maija Mikkola, Atanas Boev, and Atanas Gotchev. 2010. Relative importance of depth cues on portable autostereoscopic display. In Proceedings

of the 3rd workshop on Mobile video delivery, (MoViD '10), 63-68. http://dx.doi.org/10.1145/1878022.1878038

17. Xin, Min, Ehud Sharlin, and Mario Costa Sousa. 2008. Napkin sketch: handheld mixed reality 3D sketching. In Proceedings of the 2008 ACM symposium on Virtual reality software and technology, 223-226. http://dx.doi.org/10.1145/1450579.1450627
18. Mark R. Mine, Frederick P. Brooks Jr, and Carlo H. Sequin. 1997. Moving objects in space: exploiting proprioception in virtual-environment interaction. In Proceedings of the 24th annual conference on Computer graphics and interactive techniques. 19-26. http://dx.doi.org/10.1145/258734.258747
19. Gerhard Reitmayr, and Tom W. Drummond. 2006. Going out: robust model-based tracking for outdoor augmented reality. In Proceedings of the IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '06), 109-118. http://dx.doi.org/10.1109/ISMAR.2006.297801
20. Gerhard Schall, Erick Mendez, Ernst Kruijff, Eduardo Veas, Sebastian Junghanns, Bernhard Reitinger, and Dieter Schmalstieg. 2009. Handheld augmented reality for underground infrastructure visualization. Personal and ubiquitous computing 13, 4: 281-291.
21. Hartmut Seichter. Sketchand+: a collaborative Augmented Reality sketching application. 2003. In Proceedings of CAADRIA, 209-219.
22. Edward Swan II, Adam Jones, Eric Kolstad, Mark A. Livingston, and Harvey S. Smallman. 2007. Egocentric depth judgments in optical, see-through augmented reality. IEEE transactions on visualization and computer graphics 13, 3:429–42.
23. Brandon Yee, Yuan Ning, and Hod Lipson. 2009. Augmented reality in-situ 3D sketching of physical objects. In Intelligent User Interfaces Workshop on Sketch Recognition, 1-6. http://dx.doi.org/10.1109/VR.2010.5444806