

---

# Finding AR/XR [Dox #1]

## **Sahaj Garg**

Masters of Animation, Games  
and Interactivity (MAGI),  
School of Design,  
RMIT University,  
Melbourne, Australia  
s3738367@student.rmit.edu.au  
sz.sahaj@gmail.com

## **Abstract**

Forked from initial enquiry of creating an XR, audio reactive space to be. This project aims to develop and explore AR technicals possibly supporting other AR-XR experience based projects.

The experiment of resurrecting the sculpture done in the project displays an immersive visual way on how AR can be used in preserving, exploring and studying sculptures, historical monuments and places or objects of interest.

This project mainly started with the enquiry of creating a device agnostic experience - ei- available on all devices, by means of WebAR. With future rise in AR HMD's ,web-cloud based experiences is a needed solution and the synopsis discuss on why and tries to explore the approach by means of research and practical experimentation.

This part of folio revolves around initial webAR questioning and it's current limitation. Also experimentation with AR, and submodules needed in creating AR experiences.

*Exploring the best practices in AR(augmented reality)for a better User and Developer Experience.*

## **Authors Keywords**

AR; WebAR; WebXR; RMIT Hand Sculpture; Vuforia; 3D object tracking; User Experience.

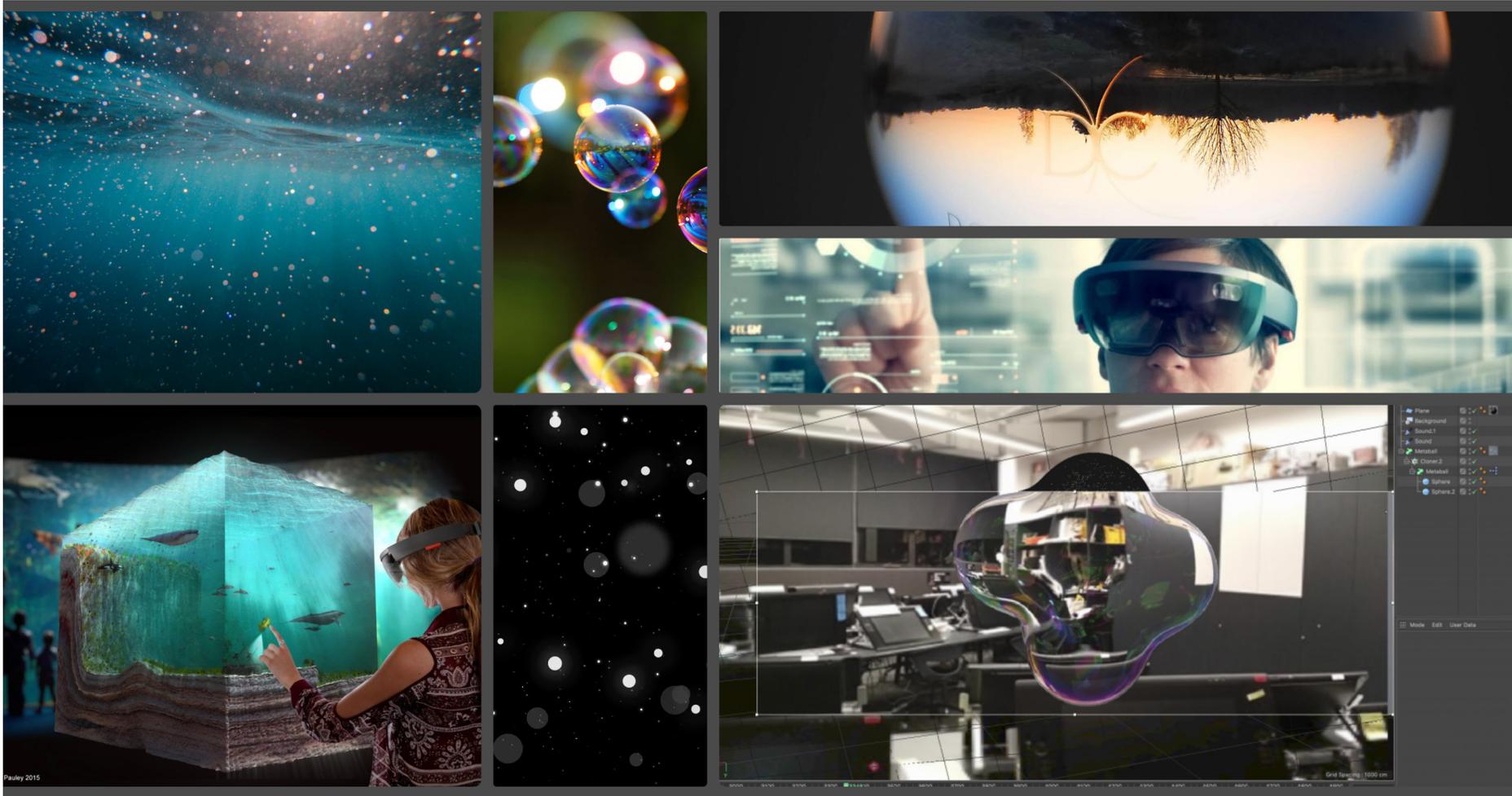
**Initial IDEAS**

Current developments in the feild of AR HMD like hololense/magic leap have fascinated me to jump into to create an experience in the same. Inspired and driven by music, the experience will be an interactive ar experience where virtual bubbles and orbs will be overlaid uppon real, morphing through sound input and interactive via user’s hand input. In an attemp to create a magical space to be in, promoting awarness through interaction, audio and visual feedbacks.

The mind map was created in a reflective way, starting out on real paper and overlaid by virtual drawings, a blend between CG and real, kind of ar.

In attempt to create this experience, there are a lot technicals which needs to be worked out, I’ll be expanding on this experience more after that.

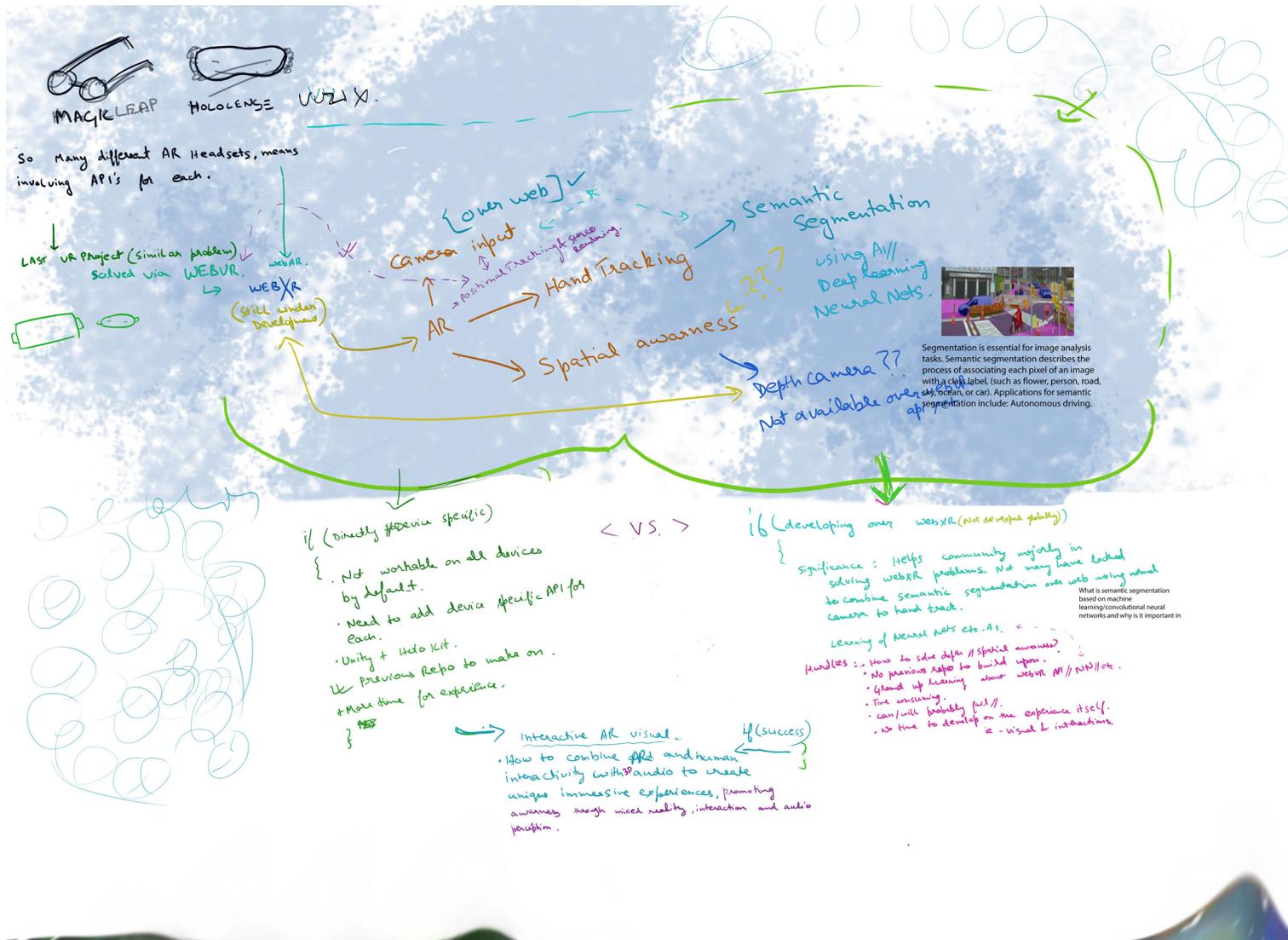




Starting and reflecting on moodboards , a shader , and visual study was done in C4D(botom right image in above picture) of how it could look like (click to go to video).

A period of deep research in this feild, I realised there's a very similar problem I faced with last semester, with so many upcoming AR HMD's releasing everymonth in this just born industry, each has an api's of it's own, which means as a developer you need to add support and study all these closed ecosystems or just enable and work with one device. This is a major hurdle in adoption of technology as a developer, and if developers are unable to adapt an industry, users certainly cannot.

Last semester, while devolping an interactive VR experience, I solved this issue by means of WebVR, (Virtual Reality over browser) which is still new. I started to dive and explore WebAR, after researching on the same and talking to developers, this is far from working (during this project). I started to dive in on 'why' and have realised there's a web of functionalities which needs to be sorted out , using different libraries and browser's and there are a series of issues of incompatilby, permissions and more between them. Enthusiastic developers including those from mozilla, google and microsoft , have created an online community to explore and resolve the same



In an attempt to get around this issue of being the experience device specific, I dug in to the deepest part of internet and doc's to see how webar(webXR) might work on platforms such as hololense, Vuzix, Magic leap. In doing so, I got introduced to even more new HMD's coming along the way . I was looking to solve spatial depth detection and hand tracking over a web browser. My mindset at this moment was not to access any device hardware(such as depth sensors) as it is different to each device, and if I maybe able to just take out all the needed data from a camera feed. (I was overlooking more of functionalities needed by AR at this point).

I came across a lot interesting concepts and methods , such as semantic segmentation, in which an AI NeuralNet is used with machine learning to detect and do hand tracking. From NASA research papers to personal dev chats, but there was a different approach and mindset needed here. ~ new data was a lot, a beautiful chaos. I started to declutter and went to basics of AR, (Augumented Reality), not even (Mixed Reality).

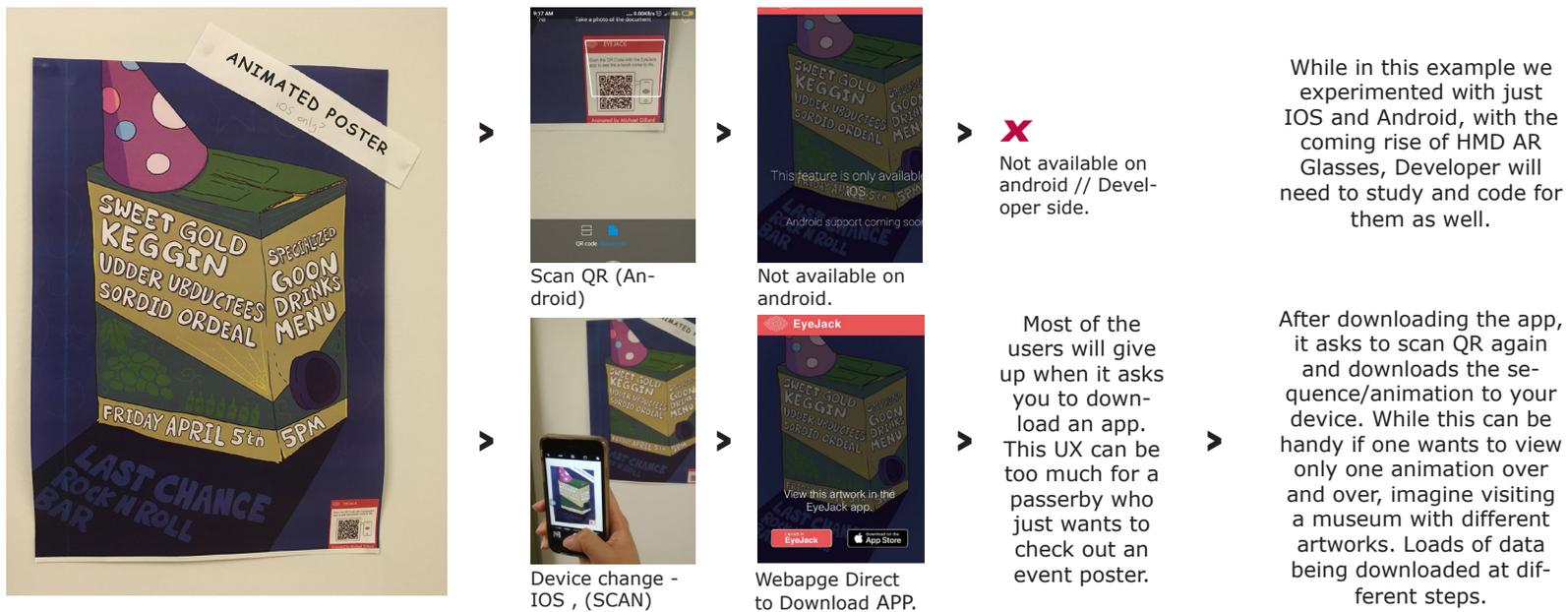
**Why?**

Web based platforms open a whole new user experience, imagine having an AR app for every shop, mall or maybe a museum you are visiting, a user will end up with many apps even if its for one time use. Most of the users would give up on an experience when it asks one to download an app, on the other end, web based ar can just provide a url redirect and the experience is ready to go. Web based ar/vr enables a better adoption amongst users and thus businesses.

**An experience**

While platform like EyeJack(AR App) provide a very easy solution to create an AR animated experience, and have been somewhat successfull in having an AR app for each artwork, that model still relies on mass adoption of one app, which again has a user experience such as :

*While walking down the corridor , we came across this poster which read animated poster, being animated creatures what we are , everyone gets excited and tries to fiddle around with their phones in order to see this static come alive.*



**How webar(XR) refines this:**

In webxr our whole experience is transferred over a browser.

**WebAR : Ux as a user:** User Only does the first step in above process ie- Just one scan or QR or a url redirect, no downloads , no device storage.

**WebAR : As a developer:** No need to implement for ios/ android separately. will work with anything with an internet and browser, even hmds like hololense etc.

My initial enquiry started with AR HMD's such as Hololense/Magic Leap/Vuzix. We are trying to make it more and more compact, which means ideally less storage. With upcoming rise of this technology, and with upcoming aid of 5G network, we would essentially not want any content on our devices and everything on a cloud over internet. As a developer web based solutions automatically makes an app device agnostic, be it android, ios or even HMD,s/AR Glasses.

**AR foundations :** Camera access, Spatial understanding(positional tracking), object rendering, marker/markerless.

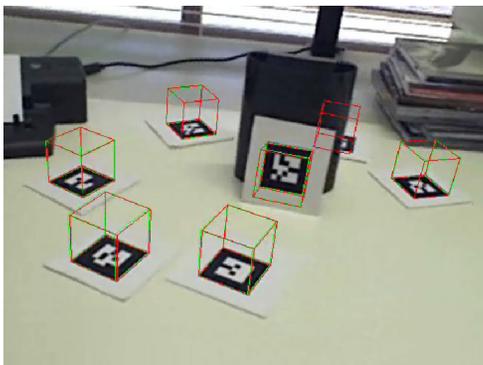
*Marker vs Markerless :*

*Markerless AR:* the user moves the virtual objects, Imagine an AR application that can place virtual furniture inside your living room. This allows you to try different combinations of objects, styles and location. For this application, the user needs to decide where to place the virtual object. This is called "markerless AR". Sometimes, all we want is to place a virtual video-game character right in front of the user. The interaction here is the game-play, and the user does not need to worry about the location of the virtual objects. Use markerless AR for those applications that do not require an "anchor" to the real world. Typically this means virtual objects will appear to "float" in mid-air. But it is also possible to automatically place a 3D augmented reality object onto a flat surface using plane detection etc.

*Location Based AR:* Location based AR can be defined as where a user's geographical location is needed to display AR, it can be useful in maps. ARCity- Ar navigation is one such application. Though it relies on phone's GPS and locomotion systems and can be off if not accurate, which can be further refined by hybrid models.

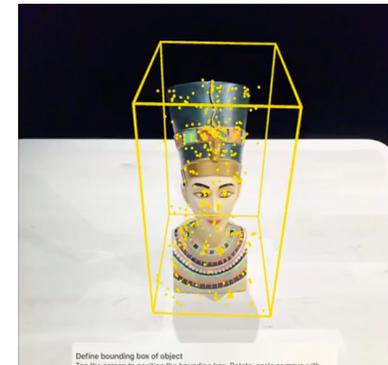
*Marker-based AR:* The digital world is anchored to the real world , the application knows what it's looking for specifically , an example could be to display an educational animation right onto the page of a book.

This can further be divided into 2D marker and 3D marker :



2D marker needs a plane based marker, and could lead to some design aesthetic challenges but are more useful in creating custom spaces, and animations over things such as posters, books etc.

3D marker does not need any kind of plane markers but are useful while tracking real life three dimensional objects such as sculptures, toys etc. It uses a trained Dataset often created via exact 3D model of the object.



To create a strong foundation with AR, i'll be doing small experiments with marker, markerless, 3D object tracking, first via native app, and then will try to implement and experiment merging and developing upon SDK's to create a web based solution ideally with most of these features.

## Initial WebAR questioning~

In order to find out why WebAR is not yet majorly adopted , I started to dig in SDK's, starting from Vuforia(the biggest in AR competition right now) , After doing some tinkering and a failed test to export it as a webbased (webgl) , I realised the base of Vuforia is built for native devices and uses the file types .dll , which are not supported in web browsers.

[I was also comparing arkit and arcore](#) (and [this link](#)) at this moment, coming across a web implementation of AR ToolKit - JsArToolkit(ar.js)it is very limited and doesn't have unity support. Also [Arcore Unity support: NO](#) , **Why UNITY?** Having access to an editor like Unity enables to create advanced experiences, and profile optimisation better which is important for web based content.

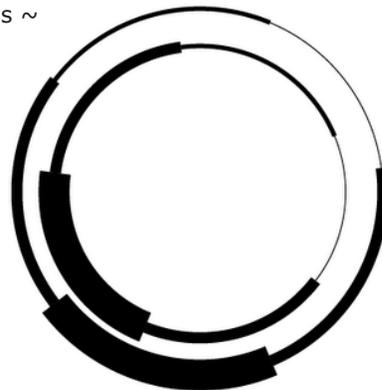
*"The bad news is that none of the current libraries built for doing so offer a Unity plugin.. Meaning you'll either have to create a wrapper, do some complicated RPC call to talk to the JS library via Unity, or completely scrap Unity altogether and use only the library. To my knowledge, the best browser-based AR library is AR.js"* ^ ^

comment at <https://stackoverflow.com/questions/48562493/how-to-create-augmented-reality-web-app-using-unity-vuforia>

While AR.js (library) is able to create web based AR, it seems very limited(one being like unable to do proper markerless AR [\[open issue link\]](#)) and also doesn't have support to an advanced editor like unity ("*Watch out @unity3d JS is not ECMA script! There is so much incompatibility.*" [@Comment on twitter to developer of AR.js](#)". Also, based on the open source ARToolkit tracking library, JSARToolkit uses WebGL & Three.js to render 3D models over real world objects, it has not yet got support for an editor like Unity. While these libraries are there, they are very limited in terms of tracking(3d object tracking) plane detection, rendering etc, however further research is needed on the topic and if unable to get a workaround through unity, I'll visit these individual libraries and will test their capability as individuals(outside unity).

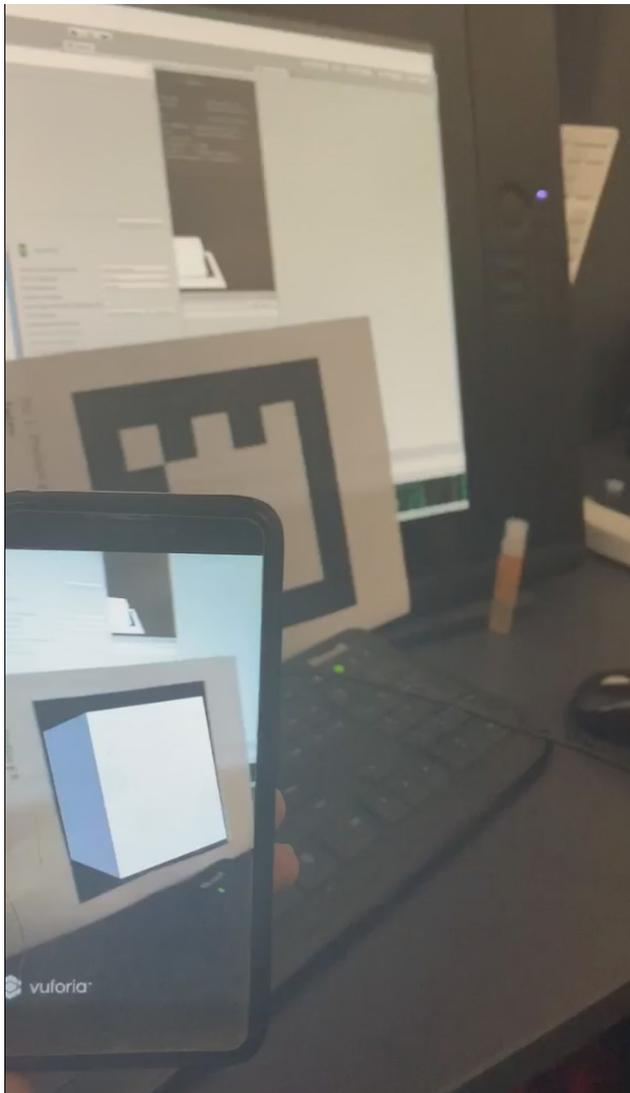
On further research, I came across [100's\(many\)](#) of AR SDK's , though almost none was able to offer all what I'm looking for, : Unity support, Web compatible, 3D object tracking feature. Though few paid versions say to have these feature, it means going exclusively through their team to find or develop anything, (eg one being xzimg SDK' , I'll be visiting this SDK's trial, later - as through support chat they said, it can export to webgl via unity, though is not able to do any 3D object tracking over webgl. *100's of tabs, dev chats, reddit posts, forums , to the deepest of internets*

Running in circles ~



## TESTS

Just starting with Vuforia first, as it has good resources and has better dev experience as compared to other SDK's at the moment, to develop understandings. (2D Plane) Marker based [NATIVE APP] test> [\(click to go to video\)](#)



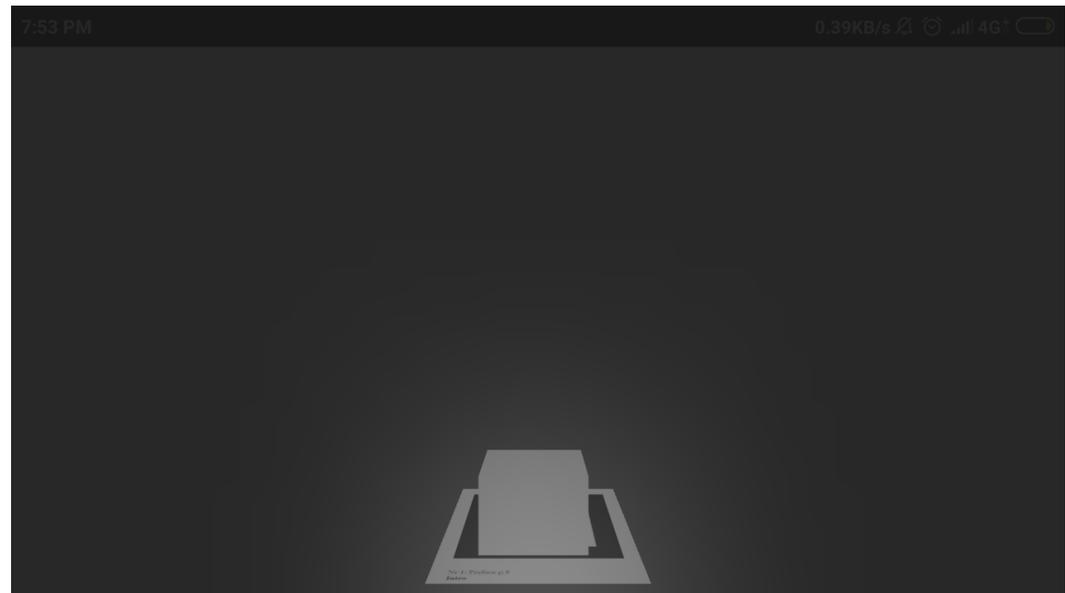
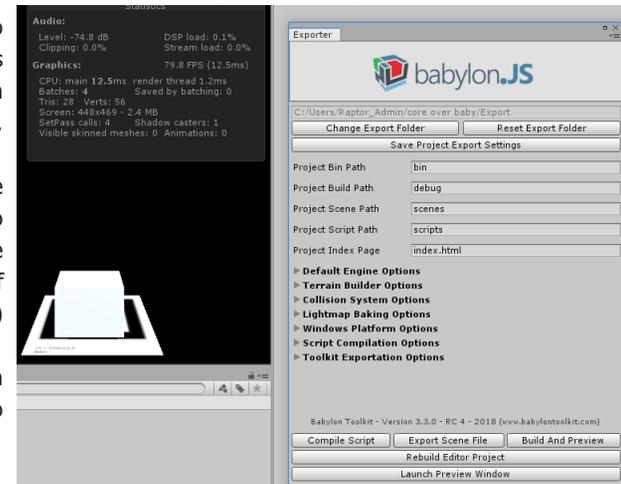
A *webtest* ~ While a normal WebGL export via Unity, won't build the package and throw compile errors while using Vuforia (not supported [v8.1.7]) , I tried a different approach and tried to export it via an external babylonJS exporter. (babylonJS have been active in webXR community) .

: ref : using babylonjs instead of webgl export <https://www.youtube.com/watch?v=Hq-p9EgHaWE?>  
: ref : <http://www.html5gamedevs.com/topic/21631-why-babylonjs-if-unity-export-webgl/>

BabylonJS Unity exporter was able to create a successful build though was unable to function as it should(which was expected and a longshot) .

Though a step closer, bottom image shows how the experience came out to be on a mobile browser, it's unable to access camera and work with any of AR components.(maybe the browser?)

BabylonJS is working extensively on webXR api, and will be revisited to explore.



### Further AR tests // A Side experiment :

To explore and oil up the AR mind, I decided to do a side experiment, where I'll try to ressurct the broken RMIT hand sculpture by means of AR. Ideally, it'll be great to have it through web, so every passerby can see by just scanning a QR but that needs further development work and time. So I decided to build it through native app.



Primate Hands and Feet , Royal Melbourne Institute OfTechnology, Public Sculpture, Bronze, Bowan Lane, Melbourne City by [Lisa Roet](#)

The basic idea being, when someone sees the sculpture through the AR app, they'll be able to see what the original sculpture looked like (on right) before it broke (on left).

To create, detect the location and display the Virtual Hand, I had several ways to approach it, first and most easy being a plane marker based (as experimented with above[Test1^]) though this approach meant putting up a marker, or some-kind of painted X. (marker's work best, when easily distinguishable simple straight shapes on a black and white contrast). Putting a marker on a sculpture is one of the less preferable things to do.

There was a better option to this ~ 3D marker (hand be the marker)// model target. In order to do this, an exact 3D model of the detectable object is required preferably with exact texture.



The sculpture was first observed broken in February 2019.

These experiments will help me evaluate AR workings, best possible and user friendly ways, and app performance ie-how heavy/detailed game objects is a native AR app able to work with on a standard phone, further, to optimise performance ~ which is very crucial when further experimenting with webAR.



**Modelling the hand:** In order to create a 3D representation/model of the hand, photogrammetry was used, (Click first image to go to video). Visiting the sculpture several times to cover up the portions which were not tracked properly, right timing and day was necessary as well as it was outside, we don't want to have hard shadows and highlights, a part cloudy day works best. Computer also went through several restarts because of stuck portions and redo's. Open Source software Meshroom by AliceVision was used which is a Photogrammetric Computer Vision framework for 3D Reconstruction and Camera Tracking.



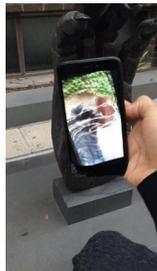
Being extra onto getting every-detail together and taking a lot of pictures. I was able to generate a highly detailed model of the hand.



The above image shows the mesh form of the 3D sculpted model outputted by photogrammetry. The number of vertices and mesh is very heavy, it'll need major optimisation to be able to use in any kind of interactive experience.



Above is 3D output by Photogrammetry, the final output and quality is much better than expected, though which is a result of a lot of pictures and a heavy mesh underneath, The texture output is also divided into seven materials and polygon subdivisions which is not handy to be used in any interactive/engines or AR.



**Creating the 3D model target :** Model Target is a feature (Vuforia) where AR is able to recognise a 3D object as a marker. In order to feed this data in the our app, we need to create a dataset from our 3D model (using Vuforia’s model target generator).

Feeding the 3D output via Gammaterly needs a lot of tweaking and issue fixes/adjustments. First being ~cutting the model, and getting rid of excess, surroundings.

Combining the multiple textures, subdivided by polygons as outputted by Grammatery.

Several hours of renders and texture baking only to realise it’s still wrong, and redoing unless model target accepts it.

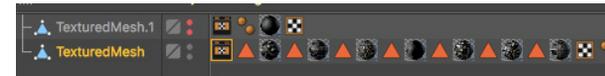
First (many) try’s were fail, 3D model tracker was unable to detect the object. Trying refining the model Cut/Cleanup and baking/combining the texture with higher resolution and settings.

Creating a *trained Model Target dataset* : (Using this above created ModelTarget, we create a Dataset) : uses the latest Deep Learning technology to train a neural network from the selected models (targets of hand sculpture model from different angles in this case), so that the model can be recognised from multiple sides. Creating and training this dataset needs a lot of considerations and a carefull step by step approach, Vuforia has done good elaborated documentation on using their Target Generator to do so.

<< clickable

**Warnings**

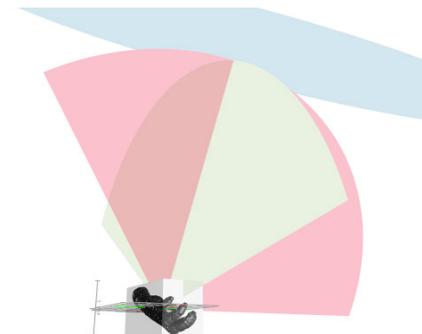
Training succeeded with problems:  
 Overall accuracy of the training is low, please try to retrain using models and views that are more different from each other  
 There is at least one pair of targets that are highly confusable, please check the marked models and views'



**CAD Model too complex**  
 The CAD model has too many parts (42) and/or too many triangles (2293150).

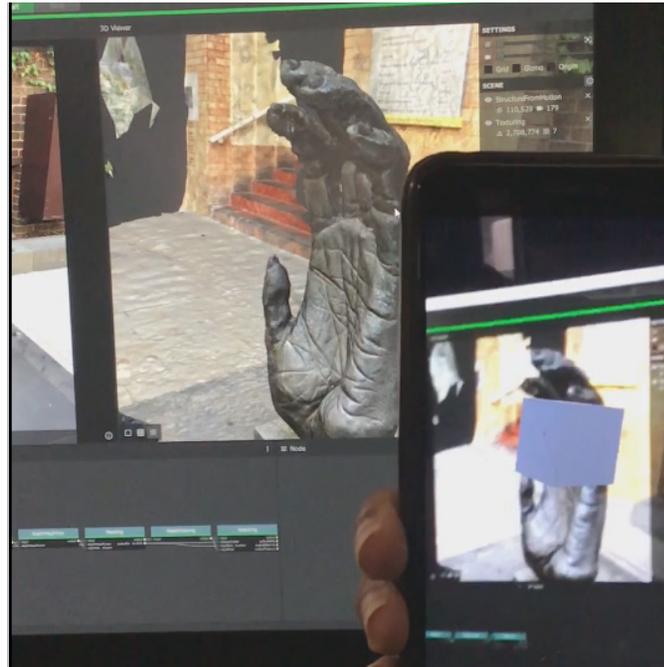
**Too much texture memory usage**  
 The uncompressed texture memory footprint of the model is too large. Please reduce the number of textures or their sizes to fix this.

X	-0.11	-3.295m
Y	-165.4°	0.771m
Z	-89.7°	-10.558m



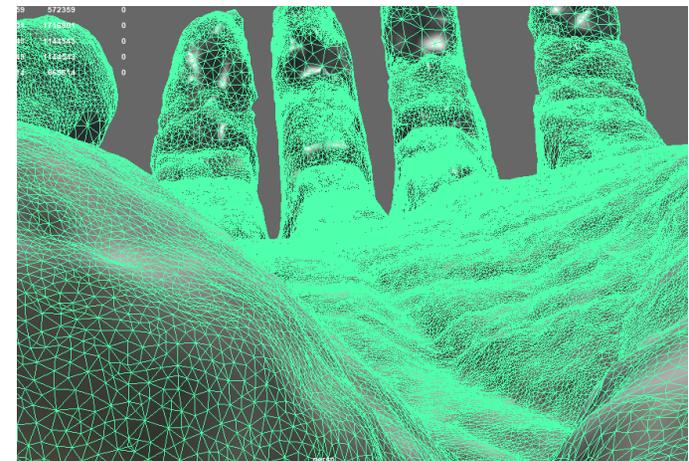
Several Model Datasets were trained refining and learning from errors of previous ones.

Detecting the 3D hand and displaying a normal cube was the first step. Parameters and hierarchies can be experimented a lot to get a more elaborated result. I was testing via the 3D model generated by Photogrammetry to test the function as it was not possible to test on the real sculpture every time during the development. While few camera angles worked, a lot of the times it didn't and the process of tweaking, refining and created were visited again and again to in an attempt and hope to detect with high accuracy. (all images clickable links to test videos)



After 'some' success with the 3D cube, I imported and played with displaying the broken part of the sculpture (other hand), which went through a process of mesh cutting and a bit of optimisation and also playing with lights .

>> The white hand is the model target(3D marker) , the textured hand is the virtual object placed once marker is detected



In detecting and displaying a virtual AR object using a 3D model, from many angles have two steps(in Vuforia's model). First being detecting from which side of the model we are looking at. It uses and identifies an approximate shape of the structure, and then throws that data to the next step, which is detecting and pinning the exact position of the model target from that viewpoint. Our trained Nueral Net dataset(model recognition object via Vuforia's Unity Bundle) was able to detect the first step well ie- from which side user is approaching the model, but step two of pinning that to final location worked only 1 out of 5 tries in a very angle limited way.

A more regressed state in training a modeltarget , where only one angle(camera) is trained per dataset and having those multiple modeltargets then feed into training the NeuralData set can be approach and can be further explored in comparison to now, where only one modeltarget is configured to target all camera angles.

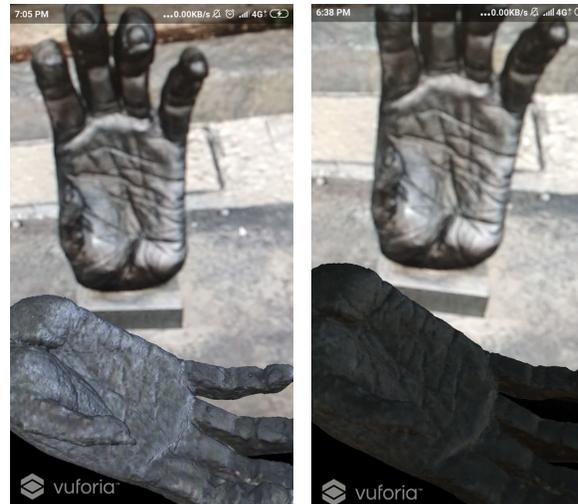
In respect to time, in this (working) model while the user can approach from any side, they need to align it from the front to make it detect. Though once detected the model is free to be explored from all sides. With above (para's) model, this could be experimented upon to have the app detect from more sides, with minimal user movement.

Also to note, it was observed that first time detection took some alingment and time, but after the first detection, it's able to detect fairly fast and almost instantly when pointed towards the object.

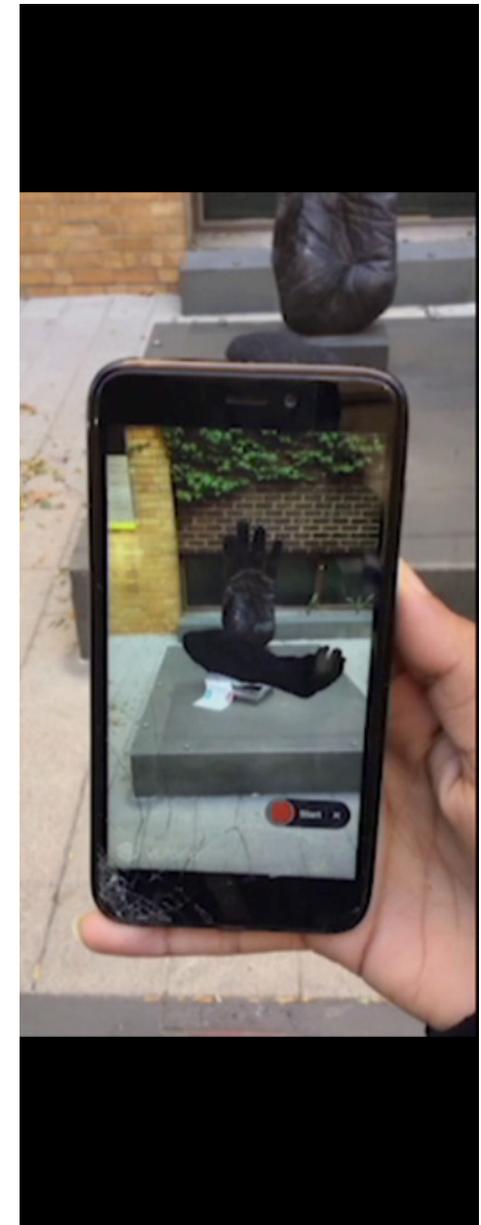
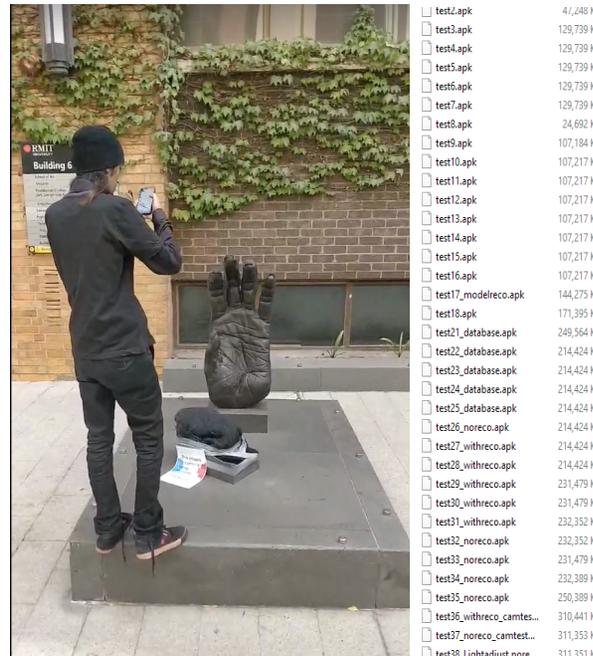
Light was also adjusted again after tests on the original sculpture.

Limitations ~ Observed limitations were, it sometimes took a long time to detect the model(but only for first time in a session). This also is dependent on the outside lighting conditions. A sculpture in inner environment will work better.

(>Click right image to view video output)



Lighting tests : Light tests were really different when seen on virtual and in Live settings. \\ Apk test exports(below)



## CONCLUSIONS

The project explores the initial of AR technology, its current limitations and overall user experience and touches on current-developments and limitations in the field of WebAR.

The synopsis also suggests how webXR(/webAR) can be used to create a better user and developer experience, and how it may come to aid in for future devices such as AR HMD's / Glasses.

The tests in this projects covers 3D model tracking and 2D marker tracking in a normal native app environment, mainly using Vuforia SDK.

*Future Direction* ~ of this project is to explore these capabilities in a webAR environment, combining and experimental research on different framework's, SDK's etc. Also exploring and touching on optimising the experience to run smoothly in a webbrowser, preferably with an aid of a visual editor.

All answering towards the enquiry of the whole picture:

*'What are the best practices in AR(augmented reality)for a better User and Developer Experience.'*

### *Experiment Significance ~ Resurrecting History*

AR experiment in resurrecting the sculpture display's a very significant use of the technology , which can be used in museums or historical places/monuments, displaying a contrast between now and how it looked in an earlier timeline, also preserving the object from decay (virtually), and can be used to track/ study the ongoing decay. This can pave new ways to experience historical places and educate/learn in an immersive way. The model can be highly useful in tourism, cultural museums, arts and historical studies.



## References

1. @jerome\_etienne. 2017. VQ. Anybody experienced in unity ? could i make ar.js, so a javascript library, available to unity developers ? how would that work ?. Tweet. (19 Jul 2017. Retrieved February 2019, from [https://twitter.com/jerome\\_etienne/status/887651816363741184?lang=en](https://twitter.com/jerome_etienne/status/887651816363741184?lang=en) >
2. Ilya Dudkin , Skywell Software. VUFORIA VS ARKIT VS ARCORE: CHOOSING AN AUGMENTED REALITY SDK . 9 January 2019.< <https://skywell.software/blog/vuforia-vs-arkit-vs-arcore-choosing-an-augmented-reality-sdk/> >
3. Lisa Roet,2004-2008, Primate Hands and Feet (sculpture),< <http://www.lisaroet.com/primate-hands-and-feet-sculpture/> >
4. Vuforia Developer Portal, March 2019, Model Targets < <https://library.vuforia.com/features/objects/model-targets.html> >
5. Vuforia Developer Portal,March 2019, Trained Model Target Datasets < <https://library.vuforia.com/content/vuforia-library/en/articles/Solution/trained-model-target-datasets.html> >
6. Alexey Severin,February 2017 , New in AR: Vuforia's object tracking form 3d models and live annotations in video chats , retrieved March 2019,< <https://www.severin3d.com/blog/2017/6/2/new-in-ar-vuforias-object-tracking-form-3d-models-and-live-annotations-in-video-chats> >
7. Babylon JS Editor < <http://editor.babylonjs.com/> >
8. WebXR Community (Examples) , viewed March 2019< <https://immersive-web.github.io/webxr-samples/> >
9. Jofish Kaye. 15 March 2014. Special issue on science fiction and ubiquitous computing. Viewed February 2019< <http://dx.doi.org/10.1007/s00779-014-0773-4> >
10. BabylonJS Github , WebXR #3899 , Open Issue 8 Mar 2018 < <https://github.com/BabylonJS/Babylon.js/issues/3899> >
11. HTML% Game Devs, why babylon.js if unity export webgl??. Forum topic. April 1, 2016 < <http://www.html5gamedevs.com/topic/21631-why-babylonjs-if-unity-export-webgl/>. >
12. Arcore Unity, Github. Crash on Startup - Unable to find libtango\_3d\_reconstruction\_api and others! #2 , Closed Issue, viewed March2019. < <https://github.com/google-ar/arcore-unity-sdk/issues/20> >
13. Using babylon js instead of webgl export, Babylon.js WebVR in Unity3D - #01 . Youtube < <https://www.youtube.com/watch?v=Hq-p9EqhaWE?> >
14. Use WebXR (WebVR, WebAR, Magic Window. BabylonJS Document. March 2019 < [https://doc.babylonjs.com/how\\_to/webxr](https://doc.babylonjs.com/how_to/webxr) >
15. Joseph Medley, Technical Writer. Developer Portal. Google. Augmented reality for the web. Last updated January 14, 2019. < <https://developers.google.com/web/updates/2018/06/ar-for-the-web> >
16. Joseph Medley, Technical Writer. Developer Portal. Google. Welcome to the immersive web. Last updated January 14, 2019. < <https://developers.google.com/web/updates/2018/05/welcome-to-immersive> >
17. <https://medium.com/inborn-experience/one-more-time-with-ar-b8c93e689e3b>
18. Create Webar. consulted February 2019. < <https://createwebvr.com/webAR.html>
19. Alexandra Etienne. 6 Apr 2017. Augmented Reality in 10 lines of HTML. viewed March 2019< [https://www.youtube.com/watch?v=v\\_Uj0C8sMi4](https://www.youtube.com/watch?v=v_Uj0C8sMi4)>
20. <https://www.youtube.com/watch?v=aLbjwJD92Z4>
21. SocialCompare, Last update 2019 Mar 5 09:26:25, Augmented Reality SDK Comparison,< <http://social-compare.com/en/comparison/augmented-reality-sdks>.
22. Unite LA , 3 Dec 2018, Rendering techniques for augmented reality and a look ahead at AR foundation .Video. , < <https://www.youtube.com/watch?v=DVt26Yk-WqA> >