# Mobile AR in Your Pocket with Google Tango†

***Johnny Lee***
**Google, Mountain View CA**

## Abstract

***The ideas of ubiquitous always-on AR glasses that enhance your visual experience of the world is an exciting one. While the industry continues to mature the hardware and software technologies necessary for compelling consumer HMDs, we will share what is possible today in a form factor you carry with you everyday: the mobile phone. We will touch on algorithms and sensors that enable end-user facing applications such as measuring tools, shopping, games, and indoor positioning.***

## Author Keywords
Augmented reality; mobile phones; computer vision; sensors; SLAM; gaming; shopping; indoor positioning.

## 1. Introduction
The future visions of augmented reality stimulate the imagination with scenes from science fiction television and films. The hope and expectations are that consumer products emerging in this space will be able to accurately replicate those fictional experiences in real life. However, the current realities of power, weight, optical systems, displays, computer vision, and object recognition are not yet able to deliver on those ambitions - especially if the goal is a device that you would be willing to wear on a daily basis in public spaces. We've seen selected cross sections of existing technologies brought to market in the form of head worn displays, all having limitations in terms of field of view, tracking capabilities, size, weight, or even application software. While the investments toward developing technologies that enable always-on wearable augmented reality displays will continue to advance over the coming years, the mobile phone will likely remain a dominant form factor for many years to come.

In 2016 alone, over 1.4 billion new smartphones were manufactured and sold. That is roughly one new phone for every 5 people on earth every year. The pressure to reduce the cost and price of mobile phone components, as well as the need to increase value to the user have created the most impressive advancements in processors, radios, display technology, and imaging sensors in human history. It stimulates investment around the world, to push the state of the art closer and closer to known physical limits. This is the type of pressure that is needed to advance the component technologies enabling future form factors. A flagship mobile phone in 2017 would look like alien technology just 15 years prior.

## 2. "Augmented Reality"
We have already seen some experiences on the mobile phones that have been labelled "augmented reality". This term has been used to describe experiences ranging from GPS-location based apps, to camera filters that add colorful elements to selfies, to games that loosely composite creatures on the camera view. It is likely we will continue to see an evolution of these kinds of experiences grow over time and the definition of "augmented reality" to expand in the pop culture.

## 3. Google Tango
Google Tango is an effort to advance the state of hardware and software technology in mobile phones necessary for a device to track its movement through space and recognize their environment. While relevant to a number of different applications, this tracking capability is a corners for experience that attempt to provide the visual illusion that rendered 3D objects react and behave just like real-world physical objects. The location of the user's viewpoint must be tracked in 3D space quickly and precisely to align the with the real-world view of the camera image. Additionally, the surfaces in the environment such as the floor and the walls must be detected so that virtual objects respond to those surfaces in a believable way. Tango-enabled devices lets developers add these kinds of elements to their applications today, enabling new kinds of tools, new ways to shop, new kinds of games, as well as a indoor navigation. Tango hardware and software enables three new major capabilities to the phone: Motion Tracking, Depth Sensing, and Area Learning.

***Motion Tracking*** - If you think about a desktop computer mouse, when you move it left and right, it has a sensor the bottom that tracks its position in 2D space over time. These sensor changes are used by the computer to move the cursor on the screen. Motion tracking is very similar, but instead of just tracking motion along a flat surface, it allows you to move in 3D space detecting x,y, and z position. It also is able to tracking its orientation in yaw, pitch, and roll. This 6-degree of freedom sensing capabilities allows the phone to respond to the physical movements of the user whether it be held in the person's hand or embedded into a head mounted device. By moving the virtual camera in a rendered scene, it possible to have rendered objects composited onto the real-world camera preview with the illusion that the objects are anchored in the physical world.


**Figure 1.** Annotations that move with real world objects.

***Depth Sensing*** - Embedded in Tango-enabled devices is a 3D sensor that detects the distance to surfaces in the environment. This is typically accomplished by projecting infrared light into scene combined with an imager that is sensitive to infrared light. The depth is reconstructed by either measuring time-dependent variations of the light (sometimes referred to as Time-Of-Flight) or by looking at distance dependent shifts in the infrared patterns (sometimes referred to as structured light or assisted stereo). This gives the phone a 3D map of the surfaces in front of the user that is aligned with the color camera view. This allows applications to measure the distances of surfaces for each pixel, which then allows the placement of 3D objects at the correct distance to appear as if they are on the physical floor or table. Game physics engines can be used to make virtual balls bounce off physical walls, or game characters navigate around furniture. Additionally, it can provide the illusion of occlusion, when an object disappears from view because it is behind another object. These distance measurements are needed to determine if surfaces should appear behind or in front of virtual content.

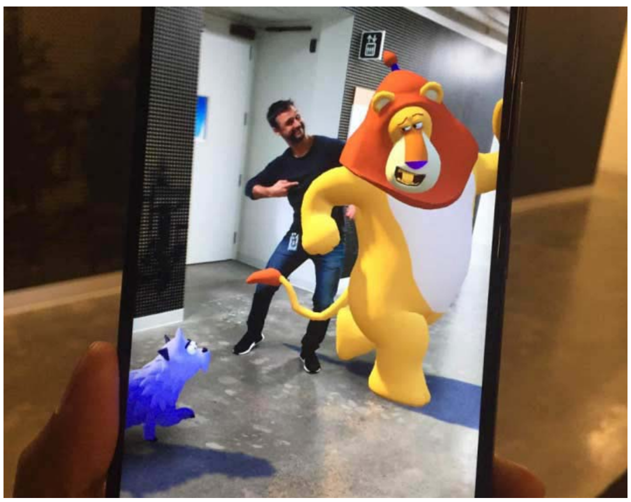

**Figure 2.** Camera app featuring virtual characters.

***Area Learning*** - If some showed you a picture of your kitchen, you could probably estimate the approximate location and orientation of where the camera was when the photo was taken. This utilizes your spatial memory to both recognize that it is your kitchen, and your familiarity with the shape of your kitchen to estimate the position of the camera. Area Learning is analogous to spatial memory. The system attempts to recognize places that it has been before so that it can associate current position to an absolute location the real world. Without it, every time the system turns on, it believes it is in a brand new space. Area Learning is an important function in allowing virtual objects to persist in a physical location, even when the device is turned off. This spatial memory can also be shared across devices, allow allowing multiple people to see a virtual object in the same location using separate devices. Venues that wish to anchor virtual content in their physical space can use this capability to share that location information with their customers. It effectively becomes a centimeter precision indoor positioning system, which enables a variety of application in retail and enterprise applications.

**Figure 3.** Augmented reality indoor navigation.

## 4. Conclusion

By combining these technologies together in a seamless manner, the user is able to have an experience that relates to their physical environment in a much more detailed and richer way than before. On the application side, the 3D assets and spatial information necessary to create rich augmented reality experiences are the same regardless if the viewing experience is in the hand or through a headset. While we wait for optical, display, and power technologies to advance sufficiently to enable wearable devices that are widely adopted, many of these experiences and the value that augmented reality provides can be done in a form factor you carry in your pocket today.