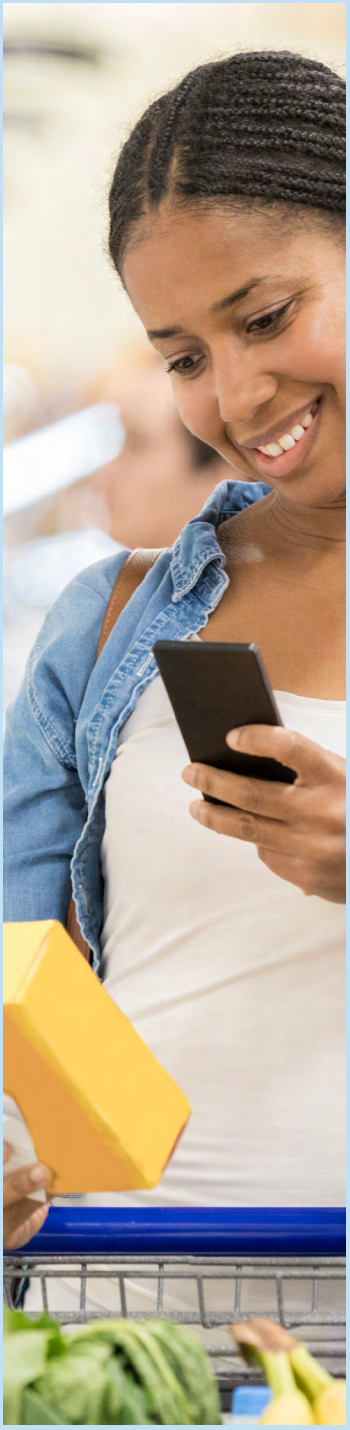
VIEW POINT



ARKIT ENHANCEMENTS, IMPACT AND WAY FORWARD

Abstract

Augmented Reality is one of the most trending technologies that is helping businesses to develop meaningful and creative solutions for their clients by superimposing digital content on their actual environment. Many organizations are working towards providing new capabilities in their AR framework that can aid developers in creating these innovative experiences. This document briefly outlines the inception and growth of Apple's ARKit framework. It mainly focuses on the recent enhancements that have been done in the framework and provides a view of how these advancements can be beneficial for the developer community.

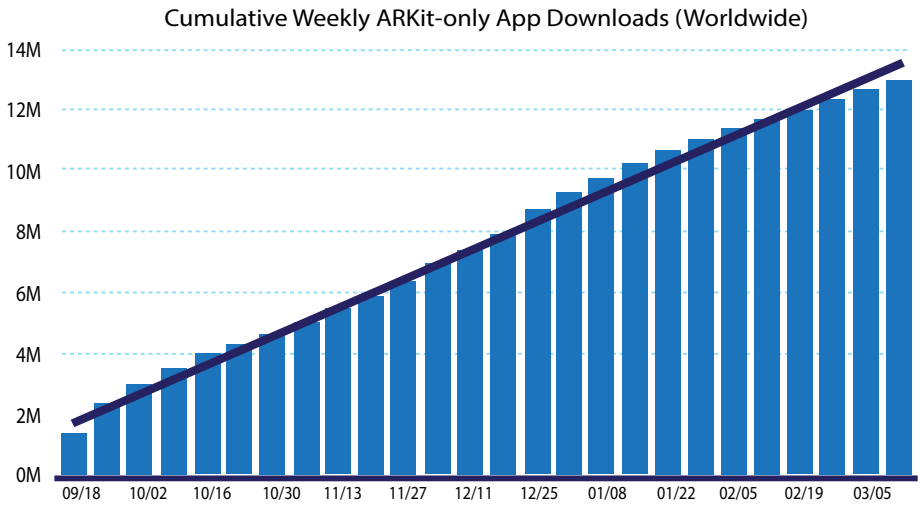
Infosys®
Navigate your next

ARKit Foundation and Growth

ARKit was introduced by Apple in 2017 as part of iOS 11 update. It is a leading-edge mobile AR platform for creating Augmented Reality applications on iOS. The initial features that were introduced as part of the first version included:

- *Tracking - To track the position and orientation of the device in the real world so that we can correctly place our digital content in the environment*
- *Scene understanding - To understand the attributes like lighting and texturing of the surroundings around the device to seamlessly integrate the virtual object in the physical environment*
- *Rendering - To place the digital object in the real world to create an illusion of realism*

Adoption rate of this framework has been huge since its inception with more than 13 million AR applications built using this framework within the first six months of its release. The graph shows weekly ARKit app downloads within these six months (September 2017 – March 2018, as per source).



Source : https://sensortower.com/blog/arkit-six-months

Second version of the framework was introduced in 2018 and included the following features:

- Saving and then loading maps for persistence and multi user experiences
- Environment texturing for realistically rendering the AR scene
- Image tracking for capturing and tracking 2D in real time
- Object detection for rendering digital content based on recognition of 3D objects
- Face tracking enhancements

Enhancements and Scope

Many new features have been released as part of ARKit 3 this year, which can help developers in creating many advanced and sophisticated applications in this area. Let's go in more detail to know what this latest version of the framework has in store for us on a broad level and how these amazing features can be leveraged in AR applications.

People Occlusion

- Digital content does not occlude people in the scene
- Achieved using advanced ML algorithms
- Works for multiple people in the scene that might be fully or partially visible

To make persuasive AR experiences, it's important that we create an environment that looks realistic to the user. This can be achieved by carefully placing the digital content in the real world so that it looks pragmatic even if there are some people standing in the scene. Content should not be occluding people if it is not supposed to do that else it will break the illusion. 'People Occlusion' helps in achieving that level of realism by calculating the distance of people from the camera and then placing the content appropriately. This is achieved by using advanced machine learning to do depth estimation. This is a significant change with respect to the earlier version of the framework where digital content placed in the environment appeared on top of the people in the scene. It not only works for multiple people in the scene but for fully and partially visible people as well.

With the capability to render content behind people, this feature aids the developers to display beautiful digital backgrounds in the scene. Applications like Snapchat can provide many cool features by making use of this advancement in the framework. People can even be realistically shown entering into the digital world and interacting with the surroundings.

Motion Capture

- Digital recording of movements in 2D and 3D
- Skeleton representation
- Helpful in entertainment applications

Motion capture refers to the technique of digitally recording the actions and movements of an actor.

ARKit3 provides the capability to track human body in 2D and 3D which can then be mapped to a virtual character in real time. It provides skeleton representation of the person. This is made possible by using advanced ML algorithms on Apple Neural Engine.

In addition to 3D body tracking and 2D body detection, this feature also provides the precision and orientation of the device along with selected world tracking features like plane estimation or image detection.

Motion capture can be very helpful in entertainment applications that use 3D animation. Further combining it with 'People Occlusion' might enable developers in creating experiences wherein users can be realistically interacting with the digital content in the real world e.g. retail customers can interact with the products in store to know more about them and even try them out.

Simultaneous Front and Back Camera

- Face tracking using front camera
- Scene tracking using back camera
- Useful in applications based on sentiment analysis

We have experienced and developed many AR experiences till now that involved the usage of back camera of the device for world tracking and then superimposing digital content on the top of it. For the first time this year, Apple announced the capability to use both the front and back cameras of the device simultaneously. It enables 'Face Tracking' with device orientation and position with six degrees of freedom using true depth camera system at the front.

It opens plethora of possibilities for AR developers where they can take the benefit of sentiment analysis in their applications. Front camera can help analyzing the face expressions of the user and based on that, the content can be rendered or functionality of the application can be altered. For instance, retail applications can provide various product options depending on user's likes/dislikes by capturing their emotions through expressions. The feature can further be enriched by integrating speech along with the expressions.

Let's hope to see more advancements in this area in the coming years to further enhance this feature by providing more scene understanding using front camera as well.

Collaborative Sessions

- Continuous synchronization of world maps
- Ability to leave and rejoin the same session
- Ability to share positional information with peers

Augmented reality as an experience gains more value when it involves collaboration with multiple users. In ARKit 2, multi user experience required users to record and load maps before the start of the collaboration. This enabled them to see virtual content placed by others only within the shared maps. However, it did not allow new 3D landmarks and ARAnchors to be shared at run time.

ARKit 3 overcomes the above restrictive experience by allowing users to freely explore the environment and update localized maps. By maintaining continuous sharing of the maps, ARKit is able to identify overlap regions of the individual users' maps and provides us a unified map of 3D landmarks and ARAnchors.

This now gives us a seamless augmented reality multi user experience allowing us to move freely without any restriction inside our environment. This feature is critical if augmented reality is to become mainstream as it overcomes a crucial barrier of sharing virtual content with anyone without having to perform a series of steps and instructions. This paves way for large scale augmented reality applications in superstructures like malls, factories, yards, etc. ARKit 3 takes the collaboration a step further by connecting users on a peer to peer basis within the same network layer. In other words, users can enter and exit the session at any given point of time. Users can even share their positional information with others. This allows certain actions to be triggered when a specific user enters a zone, say, display warning in hazardous area or exhibit an ongoing offer in a shopping complex.

To top all this, the collaboration mechanism works under the hood, allowing other features provided by ARKit form main part of the experience. By offering other features in a multi user experience, it creates a window of multitude immersive experiences.

Multiple Face Tracking

- Ability to track up to three faces simultaneously
- Persistent tracking ability

ARKit 2 allows applications to use the true depth front facing cameras to provide robust face tracking feature in real time. This ability provides face detection and positional tracking of the detected face with six degrees of freedom. It opens up the opportunities for various selfie effects like virtual tattoos, makeup tryouts, masks, jewelry, glasses, etc.

Face tracking also includes understanding facial expressions by categorizing pose specific features like eyelids, jaws, nose and so on. These coefficients representing facial features also known as blend shapes, can detect distinct 50 muscle movements. This allows us to use these detailed blend shapes to animate or rig a character which otherwise would be a complex task. Best example use case is Animoji app.

In addition to the above features, ARKit 3 allows face tracking up to three faces simultaneously. This tracking is also persistent, which means that in an active session, the tracking will be able to uniquely identify a face even when a person moves on and off the screen.

To add to the realism of the augmented content, rendering of face geometry is provided with color data, depth image, light intensity and direction information. On a side note, face tracking allows applications to record audio samples, which is disabled by default.

Scene Understanding Improvements

- Advanced image detection with up to 100 images
- Quality feedback of the image
- Advanced object detection using ML algorithms
- Advanced plane estimation

Scene Understanding has been a significant feature in the framework from the beginning and we have been observing various refinements in this area every year.

This year Apple has announced image detection capability advances by giving the ability to detect up to 100 images at the same time in conjunction with automatic scale estimation and providing quality feedback of the image at run time.

Object Detection has also been drastically improved by incorporating machine learning algorithms for detection. Objects can now be recognized faster in more vigorous environments. Body parts recognition apart from face and hand recognition has been a complex area due to the unavailability of sufficient feature points for identification. Hopefully, we should be able to achieve this better by improvements in object detection area.

ML-supported plane estimation also provides more accurate and faster plane detection along with the ability to detect walls surrounding the plane without any feature points.

These improvements look promising and hopefully we will witness lots of complex and entertaining AR applications in the near future.

RealityKit and Reality Composer

RealityKit and Reality Composer have been designed to aid the developers in creating AR content and applications without the need of having any complex gaming engines or data modelling experience.

RealityKit is a high-level framework that provides the capabilities to the developer to render the AR content efficiently and seamlessly. It leverages information provided by ARKit to smoothly integrate virtual objects into the actual physical environment.

Reality Composer is a new application developed for iOS and Mac featuring drag and drop interface that makes it easy to model AR content. It has a built-in library of high quality objects and animations.

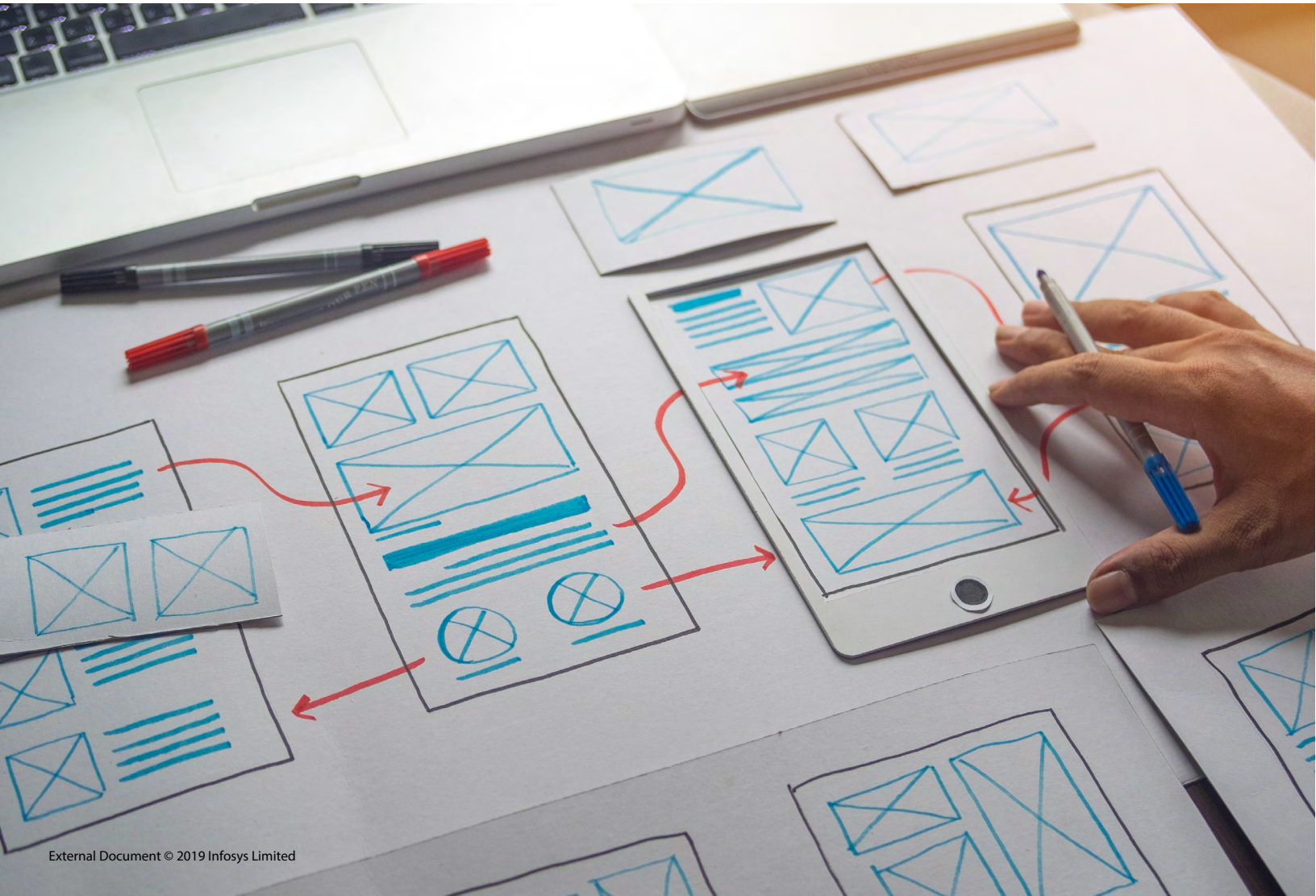
It provides a very useful capability of recording and replaying an experience that can help the developers in creating and debugging AR applications. They can simply record their AR experience while testing and can replay the movie file at their desk to be able to continue with the development.

AR Quick Look

AR Quick Look was introduced in 2018 that helps in previewing the virtual objects in 3D or AR space by simply sharing 3D model files of USDZ type. The idea of designing this feature was to enrich any application on iOS with the AR content using Quick Look technology without the need of understanding the details of AR.

AR Quick Look View can be integrated in iOS applications using Quick Look API. These views can also be integrated in websites in Safari.

This year in 2019, Apple has taken a step forward and provided the integration of AR Quick Look with the new Reality Composer App so that the developers and designers can quickly view the produced content. Improvements have also been done to the visual rendering and experience with AR Quick Look for providing more realism and greater immersion.



Conclusion

With the advancements that have been done in this area, Apple has given the developers an ability to create amazing augmented reality experiences that have not been possible earlier.

Augmented Reality as a technology has created an impact on every field whether its retail, healthcare, education, real estate or any other that we can think of. Many businesses are using it for training their staff and providing maintenance assistance to the technicians while they are on the job. This technology is still in its initial stages and there is a lot of innovation that we might observe in the coming years.

References:

1. https://www.techopedia.com/definition/4776/augmented-reality-ar
2. https://developer.apple.com/videos/all-videos/
3. https://developer.apple.com/augmented-reality/
4. https://developer.apple.com/augmented-reality/reality-composer/
5. https://developer.apple.com/augmented-reality/quick-look/
6. https://venturebeat.com/2019/06/03/apple-reveals-arkit-3-with-realitykit/
7. https://medium.com/predict/the-future-of-augmented-reality-90143b98f7a3
8. https://www.forbes.com/sites/timbajarin/2019/06/14/can-augmented-reality-make-everyone-experts/#7517e7f147d7
9. https://en.wikipedia.org/wiki/Motion_capture
10. https://uploadvr.com/2019/06/03/arkit-3-realitykit-wwdc-2019/

About the Authors

Kavita Sabharwal Kanwar , *Technology Architect, Infosys*

Kavita is an Immersive Technologies and Mobile Solution Architect with 16+ years of experience. She has extensive experience in architecture, design and implementation of B2B, B2C mobile solutions for leading clients in diverse domains. She is also responsible for architecting and implementing AR/VR solutions as part of Immersive Technologies COE.

kavita.kanwar@infosys.com

https://www.linkedin.com/in/kavita-sabharwal-kanwar-39074833/

Bharadwaj Balraj , *Technology Lead, Infosys*

Bharadwaj is an Immersive Technologies Lead with 6+ years of experience in IT Industry. His role in Immersive technologies COE involves design and development of Enterprise Augmented and Virtual reality solutions. He specializes in Unity3D content development and is an active AR/VR evangelist.

Bharadwaj_Balraj@infosys.com

https://www.linkedin.com/in/bharadwaj-balraj-4ab619145/

For more information, contact askus@infosys.com

Infosys®
Navigate your next