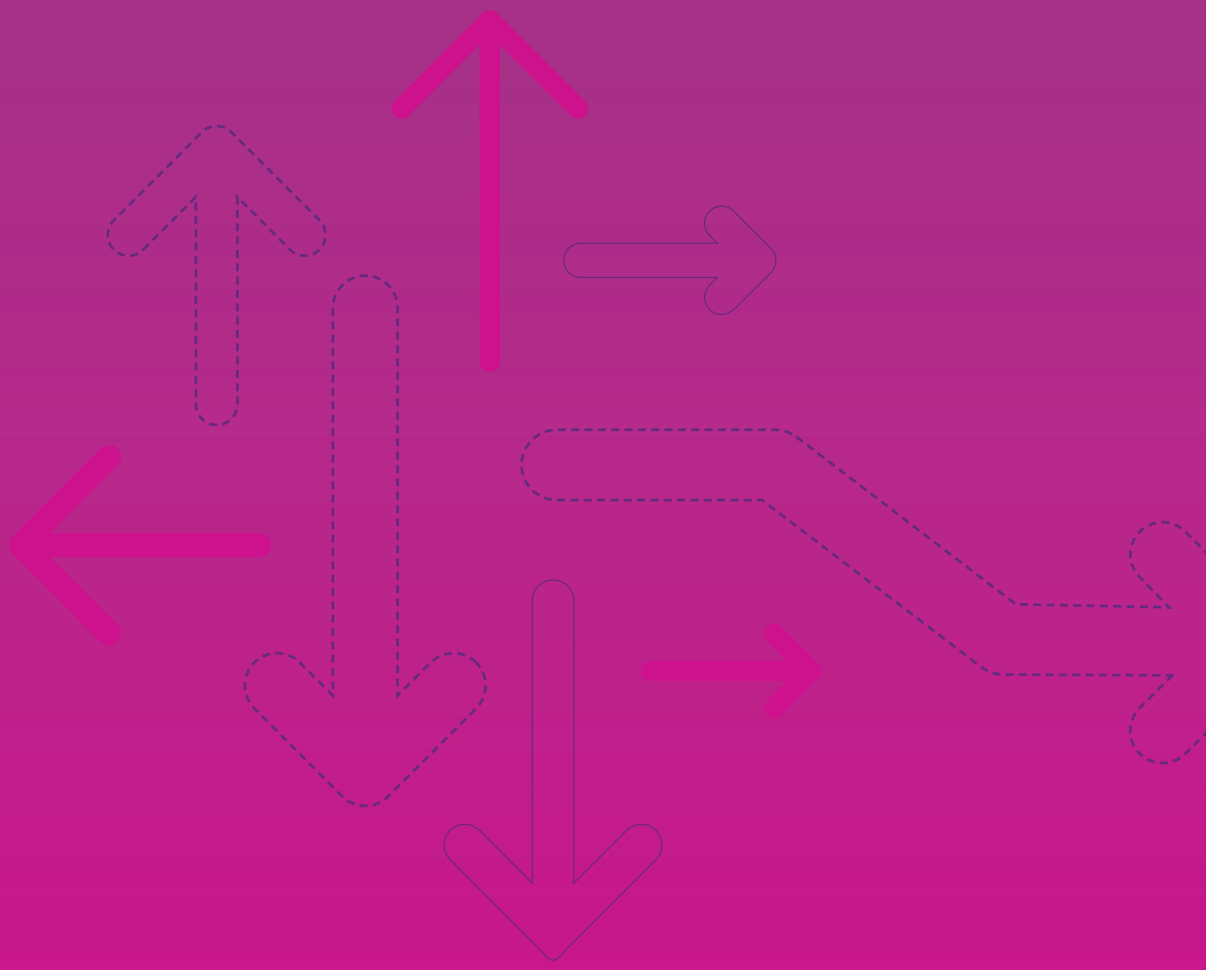
QUALCOMM®

Making Immersive Virtual Reality Possible in Mobile

March 2016



Disclaimer

Qualcomm Technologies Inc.

Qualcomm Snapdragon, Qualcomm Adreno, Qualcomm Hexagon, and Qualcomm Symphony System Manager are products of Qualcomm Technologies, Inc.

Qualcomm Technologies, Inc.
5775 Morehouse Drive
San Diego, CA 92121
U.S.A.

Contents

1 Executive summary

Immersion enhances everyday experiences, making them more realistic, engaging, and satisfying. Virtual reality (VR) will provide the ultimate level of immersion, creating a sense of physical presence in real or imagined worlds. VR will bring a new paradigm for how we can interact with the world, offering unprecedented experiences and unlimited possibilities that will enhance our lives in many ways.

This promise of VR has excited us for decades, but it is happening now due to the alignment of ecosystem drivers and technology advancements. In fact, it is mobile technologies that are accelerating VR adoption. To stimulate our human senses with realistic feedback, truly immersive VR places extreme requirements on several dimensions of the three pillars of immersion[1]—visual quality, sound quality, and intuitive interactions. Adding to the complexity, mobile VR requires full immersion at low power and thermals so that the headset is sleek, lightweight, and stylish while remaining cool.

Qualcomm Technologies, Inc. (QTI) is uniquely positioned to support superior mobile VR experiences by designing for these extreme requirements. By taking a system approach and custom designing specialized engines across the SoC, Qualcomm® Snapdragon™ processors are engineered to provide an efficient heterogeneous computing solution that is optimized from end-to-end for VR.

2 VR offers unprecedented experiences

Immersive experiences stimulate our senses—they draw us in, take us to another place, and keep us present in the moment. When done right, VR provides the ultimate level of immersion. By definition, VR creates a sense of physical presence in real or imagined worlds. For creating truly immersive experiences, VR places high emphasis on the three pillars of immersion to stimulate our senses with realistic feedback, providing:

- Visuals so vibrant that they are eventually indistinguishable from the real world
- Sounds so accurate that they are true to life
- Interactions so intuitive that they become second nature and you forget there is even an interface

Figure 1: Immersive VR stimulates our senses with realistic feedback

[1]https://www.qualcomm.com/documents/making-immersive-virtual-reality-possible-mobile

2.1 A new paradigm for how we play, learn, and communicate

By creating the sense of presence, VR will bring a new paradigm for how we interact with the world. VR will offer unprecedented experiences and unlimited possibilities that allow us to live out our dreams on-demand. We'll be able to be anywhere and do anything, with anyone. Our imagination and creativity are the only limits, but they will be stretched. It's truly an exciting time. VR will impact every aspect of our lives, redefining how we play, learn, and communicate.

Figure 2: VR offers unprecedented experiences and brings a new paradigm for how we interact with the world

How we play

Immersive movies, shows, and videos: Movies and shows will be brought to a whole new level with things like 3D 360° spherical video capture and playback. The viewer will become the director and be able to look wherever they want, whether snowboarding, sky diving, or walking on the moon.

Live concerts, sports, and other events: Imagine watching sports from the best seat in the stadium, whether it is basketball, football, the World Cup, or the Olympics (Figure 3). Or even imagine having the point of view of a player, such as a NBA player on a breakaway dunk.

Interactive gaming and entertainment: Video games already create computer-generated, 360° virtual environments that are amazing, but VR will take gaming to the next level by making you feel like you are actually a character inside the game.

Figure 3: Sports fans enjoy a soccer game from the best seat

How we learn and work

Immersive education: School education and classroom learning will be revolutionized when students are taught with the best materials and lectures, which will keep them engaged. Students can take a field trip to Paris, the Great Wall of China, or the Grand Canyon without ever leaving the classroom.

Training and demos: Training in many areas, such as healthcare, military, manufacturing, and even astronauts, will be more effective, more cost efficient, and safer with VR. For example, doctors can train in everything from surgical procedures to diagnosing a patient (Figure 4).

Figure 4: Doctors learn through realistic training

3D design and art: Design and visualization will be much more intuitive in VR. For example, car companies can build virtual prototypes of new vehicles, testing them thoroughly before producing a single physical part.

How we communicate and interact

Social interactions: Individuals can virtually meet, interact, and talk with one another while feeling as if they are physically located in the same place. Family get-togethers and work meetings from distant locations will never be the same.

Shared personal moments: The people you care about can virtually feel part of the moment, whether it is your child's first step, an amazing vacation, or wedding ceremony (Figure 5).

Empathetic storytelling and immersive visualization: When you become part of the story and part of the scene, the storytelling can be much more impactful, immersive, and empathetic. For example, you can know what it is like to live in a third-world country or to have a disability.

Figure 5: Grandparents feel like they are at their grandchild's recital

VR is about creating these immersive experiences. Furthermore, it is important to point out that virtual reality is not augmented reality. Although they share similar underlying technologies, they offer distinct experiences. VR simulates physical presence in real or imagined worlds, and enables the user to interact in that world. While AR superimposes content over the real world such that the content appears to a viewer to be part of the real-world scene. Advancements in VR technologies will help make AR possible, and vice versa.

3 The time is right for VR

The promise of VR has excited us for decades, but introducing a new product category to the world is always challenging. However, the time is right for VR and it is happening now. Ecosystem drivers and technology advancements are aligning to make VR possible.

Figure 6: Ecosystem drivers and technology advancements are aligning for VR

3.1 Ecosystem drivers and technology advancements

The key ecosystem pieces—such as device availability, software infrastructure, and content creation and availability—are in place due to industry collaboration.

Device availability: Besides the upcoming high-profile VR product launches from multiple OEMs in 2016, smartphone-powered VR headsets are already available and taking off. Google Cardboard, which launched in 2014, has been adopted by consumers and enterprises. In fact, over 5 million Google Cardboard viewers have shipped[2]. Mobile VR headsets, as opposed to tethered, will drive mass adoption and provide the freedom to enjoy VR anywhere.

Software infrastructure: The entire VR software stack is being optimized to remove bottlenecks so that VR runs well. The software infrastructure includes many components, such as the drivers, operating system, middleware engines, and tools & SDKs. QTI, for example, is helping to optimize application programming interfaces (APIs) across the entire software stack.

[2]https://googleblog.blogspot.com/2016/01/unfolding-virtual-journey-cardboard.html

Content creation and availability: With hardware and software available, VR content is also being created and deployed. Content developers are experimenting with VR to create new and unprecedented experiences. App Stores, such as the Google Play Store and Oculus Store, and video stores, such as YouTube 360 and Facebook 360, are starting to fill up. For example, there are over 1000 Google Cardboard apps on Google Play store[2]. A broad spectrum of VR video, including TV content such as Saturday Night Live's 40th anniversary show, sports content such as the NBA season opener of the Warriors versus Pelicans, or concerts such as Paul McCartney is already being distributed.

In terms of technology advancements, exponential improvements in many technologies are making VR possible, such as:

Multimedia technologies: Graphics, video, and audio processing have improved significantly. For example, graphics processing for parallel tasks has increased exponentially for many years, enabling support for photorealistic graphics and visual quality.

Display & sensor technologies: Displays have significantly increased pixel density, power efficiency, and visual quality. Sensors, such as gyroscopes, accelerometers, and tracking cameras, are smaller, higher precision, lower power, and lower cost.

Power & thermal efficiency: Architecture innovations, such as heterogeneous computing, have improved power and thermal efficiency. Integration efficiency has improved due to better transistors and Moore's Law. Optimized algorithms, such as motion tracking, run more efficiently on the hardware.

3.2 The mobile industry is accelerating VR adoption

Many of the technology advancements mentioned above have been driven by smartphones, and the VR ecosystem development will mirror what happened in the mobile industry. The mobile ecosystem has characteristics that make the proliferation of new technologies very feasible, such as:

Innovation at scale, which brings both cutting edge technology and cost advantages. Over a billion smartphones are shipping globally per year, which brings tremendous scale and innovation to mobile.

Rapid design cycles, which bring fast adoption of those cutting edge technologies. Smartphone OEMs have a cadence of shipping an upgraded model every year.

Mass adoption, which means that smartphone usage has created a broad appeal for mainstream consumers. Smartphone users are adventurous and willing to try new things, such as downloading new apps from an app store.

4 Truly immersive VR has extreme requirements

VR places extreme requirements on the three pillars of immersion: visual quality, sound quality, and intuitive interactions. Truly immersive VR experiences can only be achieved by simultaneously focusing on the broader dimensions of these pillars. Adding

to the complexity, mobile VR requires full immersion at low power and thermals so that the headset can be sleek, lightweight, and stylish while remaining cool.

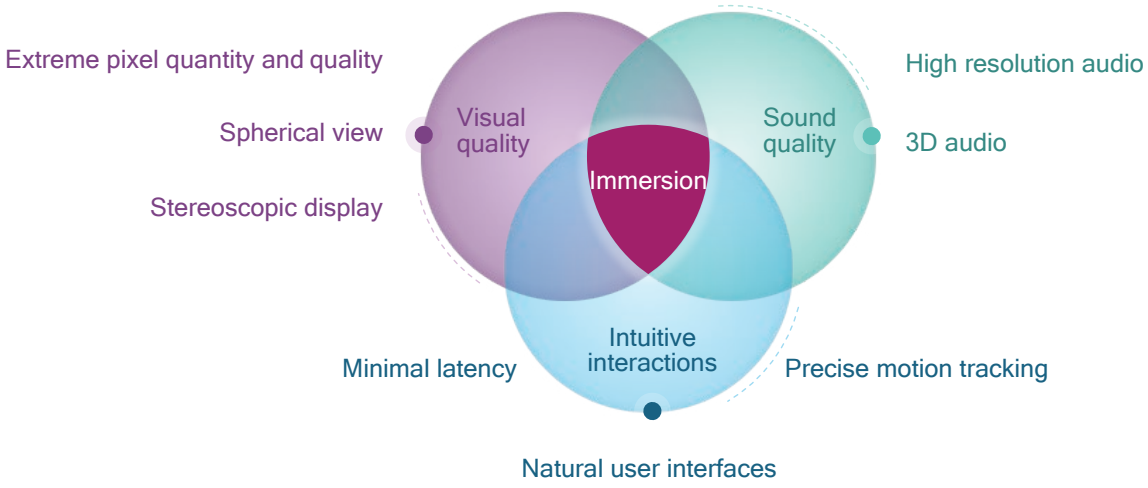


Figure 7: Truly immersive VR has extreme requirements

4.1 Visual quality

To create visuals so realistic that they are eventually indistinguishable from the real world, VR requires extreme pixel quantity and quality, spherical view, and stereoscopic display.

Extreme pixel quantity and quality

Extreme pixel quantity and quality are required on VR headsets for several reasons, most of which can be explained by human visual system. For immersive VR, the entire field-of-view (FOV) needs to be the virtual world, otherwise you will not believe that you are actually present there. The combination of the human eye having a wide FOV and the fovea[3] having high visual acuity means that an extreme number of pixels are required. In a VR headset, the display is brought close to the eyes and biconvex lenses help magnify the screen further so that the virtual world is the entire FOV. As the screen takes up more of the FOV, pixel density and quality must increase to maintain presence. Otherwise, you will see individual pixels—known as the screen door effect—and no longer feel present in the virtual world.

One potential approach to help reduce the quantity of pixels processed is foveated rendering. Foveated rendering exploits the falloff of acuity in the visual periphery by rendering high resolution where the eye is fixated and lower resolution everywhere else, thus reducing the total pixels rendered.

[3]The fovea is a small portion of the retina in the eye and is responsible for sharp vision.

The lenses in the VR headset, which help provide the full FOV, also create visual quality challenges (Figure 8). A wide-angle biconvex lens creates pincushion distortion, so a barrel distortion must be applied to the rendered image to compensate. In addition, further visual processing is required to correct for chromatic aberration, which causes colors to be focused at different positions in the focal plane.

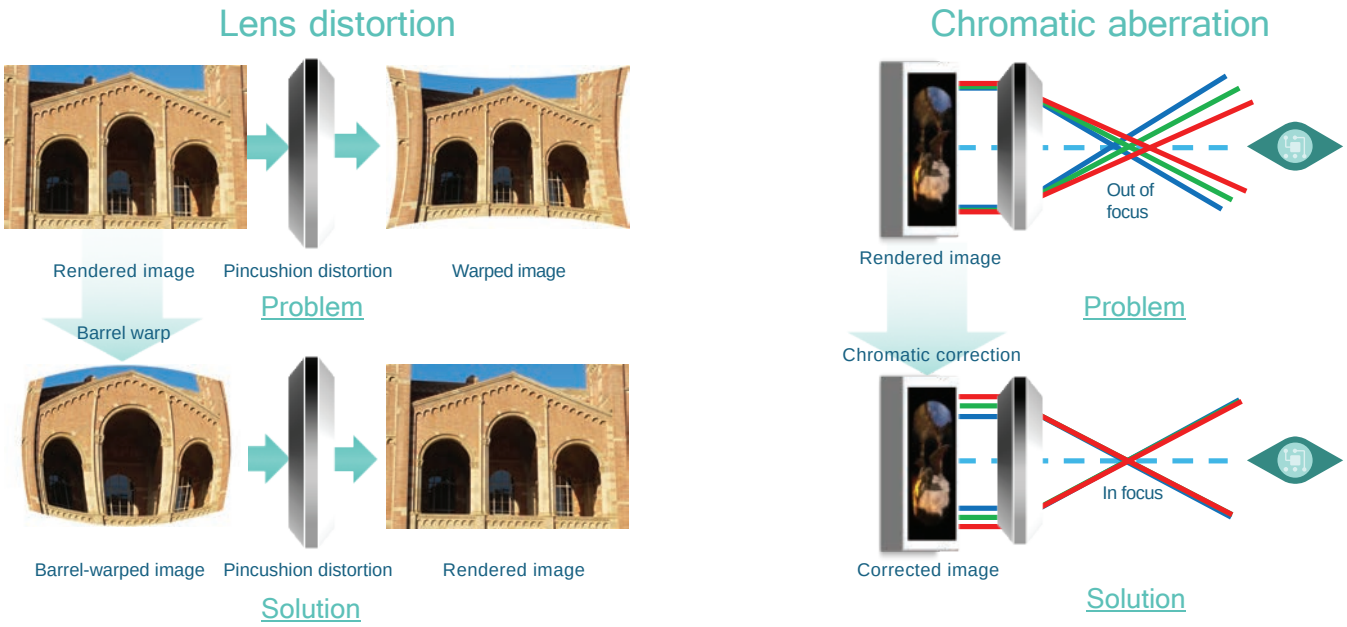


Figure 8: Problem and solution for both lens distortion and chromatic aberration.

Spherical view

While wearing a VR headset, the user is free to look anywhere in the virtual world. As a result, VR needs to provide a full 360° spherical view, which means generating even more pixels as compared to a fixed view of the world. Although a 360° spherical view is not new for gaming, it is new for video and is being driven by VR. Multiple 360° spherical video formats exist in the industry, such as equirectangular, cube-map, and pyramid-map, so it important to support them all.

Besides the video decoder being able to handle a high bit-rate, the GPU must also warp each image of the video so that it maps correctly to the display. Most premium 360° spherical video will be high resolution and content protected, so it is also important that the GPU supports digital rights management (DRM) extensions.

Stereoscopic display

Seeing the world in 3D is key to immersion. Humans see in 3D due to our binocular vision, which allows us to see objects in the scene at the right depth. To replicate this on a VR headset, a stereoscopic display shows slightly different images to each eye. These images are generated from different viewpoints, ideally the actual separation between the human eyes, so that objects in the scene appear at the right depth. Since stereographic display requires generating and processing an image per eye, this is another reason for extreme pixel quantity. For VR, we need to generate the appropriate view for each eye with stereographic rendering and video.

4.2 Sound quality

To create sound so accurate that it is true-to-life and completely in synch with the visuals, VR requires high resolution audio and 3D audio.

High resolution audio

The sampling frequency and bits-per-sample need to be up to our human hearing capabilities to create high-fidelity sound that is truly immersive. Increased sampling rates capture both the low frequency sounds, such as water dripping, and high frequency sounds, such as birds chirping, so that the entire audio environment can be reproduced. Increased precision, or bits-per-sample, improve audio fidelity. More bits allow the analog sound signal to be reproduced more precisely when converted to digital.

3D audio

Realistic 3D positional audio makes sound accurate to the real world and much more immersive. This goes beyond surround sound since the sound adjusts dynamically based on your head position and the location of the sound source. For example, if a plane is flying by, then the sound from the plane will keep adjusting as both you and the plane move. Positional audio is achieved by understanding how humans hear. A head related transfer function (HRTF) attempts to model how humans hear sound. The HRTF takes into account typical human facial and body characteristics, such as the location, shape, and size of ear, and is a function of frequency and three spatial variables. Positional audio requires the HRTF and a 3D audio format, such as scene-based audio or object-based audio, so that that the sound arrives properly to the ears.

Reverberation also makes sound more realistic. In real life, sound reflections spread and interact with the environment appropriately. Reverberation is function of sound frequency, material absorption, room volume, and room surface area. Sophisticated models of the environment can be developed to create real-time convolutional reverberation.

4.3 Intuitive interactions

To create interactions so intuitive and responsive that they become second nature and you forget that you are even dealing with an interface, VR requires natural user interfaces, precise motion tracking, and minimal latency.

Natural user interface

For a VR headset, a primary input method is head movement. Turning your head to look around is very natural and intuitive. Other input methods, such as gestures, a control pad, and voice, are also being investigated by the VR industry. Figuring out the right input method is an open debate and a large research area.

The VR headset should be free from wires so that the user can move freely and not be tethered to a fixed-location power outlet or computer. The headset also needs to be lightweight and comfortable for the user to wear it for long periods of time. In addition,

since the headset is a wearable device directly contacting the skin of your head, it must remain cool to the touch.

To meet the extreme VR processing requirements on the device within the thermal and power constraints of a sleek wearable device, very efficient processing is required. A heterogeneous computing approach[4], which utilizes specialized engines across the SoC for efficient processing, provides high performance at low power and thermals. To enable a connected mobile experience, a high-bandwidth wireless connection is required. High bandwidth technologies like LTE Advanced, 802.11ac Wi-Fi, and 802.11ad Wi-Fi allow for fast downloads and smooth streaming.

Precise motion tracking

The user interface needs accurate on-device motion tracking so that the user can interact with and move freely in the virtual world. For example, when turning your head to explore the virtual world, accurate head tracking will provide the pose to generate the proper visuals (and sounds). Similarly, if your head is stable and not moving, the visuals need to stay completely still, otherwise it may feel like you are on a boat. Motion is often characterized by how many degrees of freedom are possible in movement: either 3 degrees of freedom (3-DOF) or 6 degrees of freedom (6-DOF).

3-DOF detects rotational movement around the X, Y, and Z axis—the orientation. For head movements, that means being able to yaw, pitch, and roll your head (Figure 9), while keeping the rest of your body in the same location. 3-DOF in VR allows you to look around the virtual world from fixed points—think of a camera on a tripod. For many 360° spherical videos, 3-DOF will provide very immersive content, such as viewing sporting events or nature.

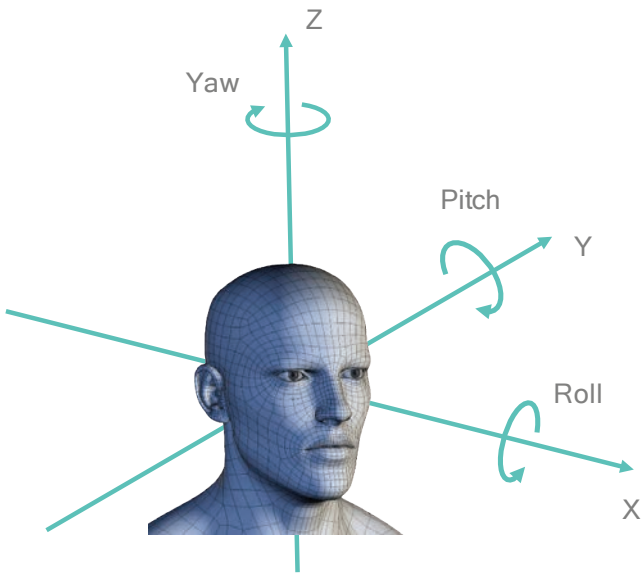


Figure 9: 3-DOF and 6-DOF

6-DOF detects rotational movement and translational movement—the orientation and position. This means that your body can now move from fixed viewpoints in the virtual world in the X, Y, and Z direction. 6-DOF in VR is very beneficial for experiences like gaming, where you can move freely in the virtual world and look around corners. However, even simple things, like looking at objects on a desk or shifting your head side-to-side can be compelling with 6-DOF.

One solution to provide precise on-device motion tracking is visual-inertial odometry (VIO). VIO fuses on-device camera and inertial sensor data to generate an accurate 6-DOF pose. The on-device solution allows the VR headset to be completely mobile, so you can enjoy room-scale VR and not worry about being tethered to a PC or getting tangled in wires.

Minimal latency

Any lag in the user interface, whether it is the visual or audio, will be very apparent to the user and impact the ability to create immersion—plus it may make the user feel sick. Reducing system latency is key to stabilizing the virtual world as the user moves. One of the biggest challenges for VR is the amount of time between an input movement and the screen being updated, which is known as "motion to photon" latency. The total motion to photon latency must be less than 20 milliseconds (ms) for a good user

[4]https://www.qualcomm.com/news/onq/2013/07/03/diversity-only-makes-you-stronger

experience in VR. To put this challenge in perspective, a display running at 60 Hz is updated every 17 ms, and a display running at 90 Hz is updated every 11 ms.

There are many processing steps required before updating the display. The end-to-end path includes sampling the sensors, sensor fusion, view generation, render / decode, image correction, and updating the display.

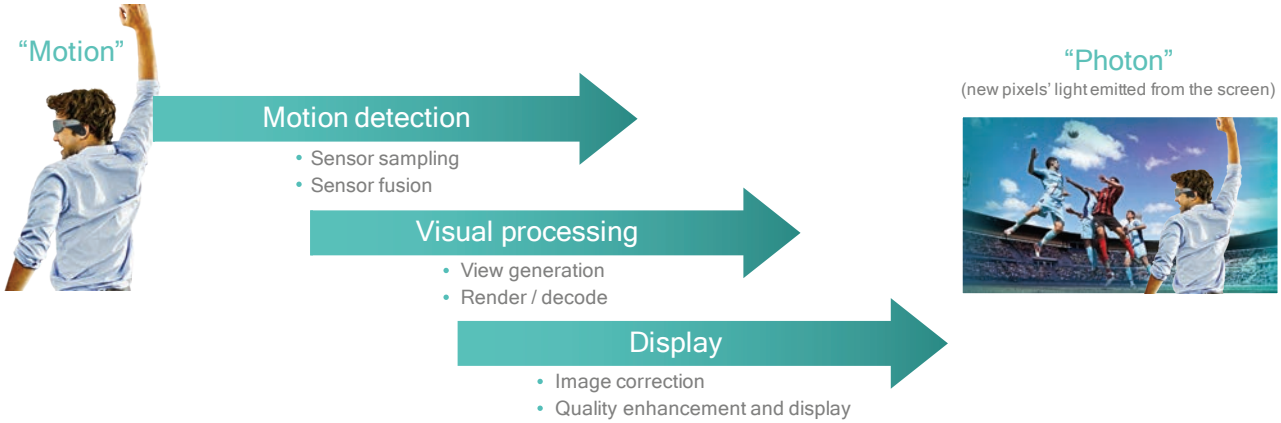


Figure 10: The motion to photon path includes many tasks.

To minimize the system latency, an end-to-end approach that reduces the latency of individual processing tasks, that runs as many tasks in parallel as possible, and that takes the appropriate shortcuts for processing tasks is needed. An optimized solution requires hardware, software, sensors, and display all working together in harmony. Knowledge in all these areas is required to make the appropriate optimizations and design choices. Possible optimizations to reduce latency include techniques, such as high sensor sampling speeds, high frame rates, hardware streaming[5], late latching[6], asynchronous time warp[7], and single buffer rendering[7].

5 QTI is uniquely positioned for VR

We are at just the beginning of the VR journey, and we expect the mobile VR headset to significantly evolve and improve over time. The industrial design will become sleeker, lighter, and more fashionable as technologies advance (Figure 11). To make this possible,

Figure 11: Sleeker, lighter, and more fashionable VR devices are coming

[5]Hardware streaming refers to processing engines that stream data directly to each other, potentially bypassing DRAM and CPU interactions. For example, the camera that streams directly to the DSP saves memory bandwidth and CPU cycles.
[6]Late latching refers to processing engines utilizing the latest head pose (rather than an older version) to generate more accurate visuals and sounds.
[7]Described below in section 5.1

we foresee continued improvements coming in many areas, such as power and thermal efficiency, pixel density and pixel quality, sound quality, intuitive interactions, and cost efficiency.

QTI is uniquely positioned to meet the extreme requirements of immersive VR as the headset evolves. We expect to make VR experiences more immersive by designing efficient solutions that meet the device constraints and by helping the ecosystem bring differentiated products to customers.

We see opportunities to improve VR experiences by focusing on the three pillars of immersion. For visual quality, we believe it is important to provide consistent accurate color, high resolution & frame rates, and 3D 360° spherical view. For sound quality, we believe it is important to provide realistic 3D positional audio, 3D surround sound, and noise removal. For intuitive interactions, we believe it is important to provide minimized system latency and precise motion tracking, while delivering intelligent contextual interactions.

Because of our heritage and technology leadership in smartphones, we are applying our mobile expertise to VR and are well positioned in the areas of:

- Faster development cycles than other types of devices, with customers who increasingly require more complete solutions.
- Sleek, passively cooled form factors that get thinner and more challenging each generation.
- Reducing cost of technologies each generation to lower price points, so that our customers can deploy new, more immersive experiences to consumers worldwide.

We also enable the ecosystem to commercialize devices and experiences via Snapdragon solutions and via ecosystem enablement. For our Snapdragon solutions, we've made the appropriate VR tradeoffs and focused on the right dimensions to design an efficient, differentiated SoC through:

- Custom designed processing engines: Rather than licensing off-the-shelf processing engines, we have custom designed several processing engines, such as the Qualcomm® Adreno™ GPU and Qualcomm® Hexagon™ DSP, to be optimized for specific tasks and efficiency. In addition to designing superior processing engines, we also gain tremendous knowledge that we use to make system level optimizations.
- Efficient heterogeneous computing: Our system approach allows us to design an optimal heterogeneous computing solution, taking into account the applications and tasks that need to be accomplished so that we can make appropriate hardware decisions. We then run the tasks on the most appropriate engines through optimized system-level software. Snapdragon 820, for example, utilizes specialized engines across the SoC for efficient processing of the diverse VR tasks, providing high performance at low power and thermals.
- Optimized end-to-end solutions: By thinking holistically at the system level, understanding the challenges, and working with other companies in the ecosystem, we develop comprehensive VR solutions. For example, we are able to optimize the entire system to reduce motion to photon latency. Removing the latency bottlenecks requires making optimizations across the entire SoC and system software. With the knowledge and design expertise in house, we can quickly come up with innovative solutions and introduce new capabilities.

For ecosystem enablement, QTI helps the ecosystem to quickly bring products to customers through app development tools, device optimization tools, development platforms, and ecosystem collaboration. For VR specifically, we are introducing the Snapdragon VR SDK[8], which has new APIs optimized for VR. Developers will be able to use commercially available Snapdragon 820 VR devices to see how VR applications run on real hardware.

[8]Snapdragon VR SDK is expected to be available to developers in the second quarter of 2016.

5.1 Snapdragon VR enhancements

Along with the Snapdragon 820 processor, which was designed with VR in mind, QTI announced the Snapdragon VR SDK. The Snapdragon VR SDK provides developers access to advanced VR features so that they can simplify development, optimize

Figure 12: Snapdragon VR SDK offers simplified development, optimized performance, and efficiency

performance, and improve power efficiency while reducing development time.

The Snapdragon VR SDK introduces new APIs optimized for VR, such as:

- DSP sensor fusion: Provides precise low-latency 6-DOF motion tracking, real-time and predicted, via fusion processing on the Hexagon DSP for improved responsiveness.
- Asynchronous time warp: Warps the image based on the latest head pose just prior to scan out.
- Chromatic aberration correction: Corrects color distortion based on lens characteristics.
- Lens distortion correction: Barrel warps the image based on lens characteristics.
- Stereoscopic rendering: Generates left and right eye view, up to 3200x1800 at 90 fps.
- Single buffer rendering: Renders directly to the display buffer for immediate display scan out.
- UI layering: Generates menus, text, and other overlays such that they appear properly in a VR world (undistorted).
- Power and thermal management: Qualcomm® Symphony System Manager SDK provides CPU, GPU, and DSP power and performance management to consistently hit 90 FPS.

Other new Snapdragon 820 features that benefit VR include:

- Several enhancements to achieve < 18 ms motion to photon latency
- High-quality 360° spherical video at 4K 60 fps, for both HEVC and VP9
- 3D positional audio for more immersive video and gaming experiences
- Ultra-fast connectivity through LTE Advanced, 802.11ac, and 802.11ad

By offering hardware, software, tools, and devices all optimized for VR, QTI is uniquely positioned to lead in VR. Snapdragon 820 is the ideal solution for today's mobile VR.

6 Conclusion

We are on the verge of consumer VR becoming a reality. After several false starts, ecosystem drivers and technology advancements are aligning to make VR possible—and the mobile industry will lead the way. VR offers such new and compelling experiences that it is going to transform the way we interact with the world. Making those VR experiences truly immersive requires simultaneously meeting several extreme requirements across visual quality, sound quality, and intuitive interactions. Mobile VR, which will drive mass consumer adoption, adds additional power and thermal requirements since the headset needs to be comfortable, lightweight, and cool to the touch.

QTI is uniquely positioned to meet these extreme requirements by designing efficient solutions through custom designed processing engines, efficient heterogeneous computing, and optimized end-to-end solutions. For example, the Snapdragon 820 was designed with VR in mind to make immersive VR possible on a mobile device. The Snapdragon VR SDK provides developers access to advanced VR features on the Snapdragon 820. QTI will continue to drive the industry forward as the mobile VR headset continues to evolve. Enabling truly immersive VR on mobile devices is yet another example of how Qualcomm Technologies is once again re-inventing the mobile world we live in.

To get most updated information about virtual reality, please visit:
www.qualcomm.com/VR