Low Latency Rendering in Augmented Reality Based on Head Movement

Jung-Bum Kim Joon-Hyun Choi Sang-Jun Ahn Chan-Min Park

Samsung Electronics
34 seong-chon, 06765, Seoul, Republic of Korea
{jb83.kim, jh53.choi, sjun.ahn, chanmin.park}@ samsung.com

ABSTRACT
In Augmented Reality, AR, the latency is a huge problem that disrupts immersive experience especially when head-worn devices are involved. Rendering of virtual objects generally accounts for major proportion of the latency in AR. In this paper, we identify the problem caused by the latency and observe human perception of virtual objects during head movement to find opportunities to reduce the latency. We propose solutions that reduce the latency of rendering by introducing Level of Detail (LoD) concept based on head movement. Experimental results show our approach is effective to decrease the latency of rendering.

Keywords
Augmented Reality, low latency rendering, Level of Detail, head movement.

1. INTRODUCTION
Augmented Reality, AR, has significantly emerged in recent years. The advent of head-worn devices including [Ocu12a, Goo12a, Sam15a] has demonstrated usefulness of AR technology in various areas such as education, game, and location-based service [Van10a]. Thanks to the immersive experience that head-worn devices provide, AR has been recently gaining enormous attention [Zho08a]. In this paper, we examine properties that prevent users from getting absorbed in AR when head-worn devices are engaged. Head-worn devices, in AR perspective, are very different from stationery devices and hand-held devices, because users change their view direction very quickly. Rapid change in head direction is one of the major sources of common problems in AR. Since AR requires a variety of operations including object recognition, object tracking, and rendering, it inherently yields some latency. In contrast, the latency of perceiving real world is nearly zero, which indicates there could be a gap between real world and virtual objects. We define this gap as *realism gap* which describes inconsistency of virtual objects overlaying onto real world. We concentrate on rendering to reduce the realism gap by analyzing human perception of virtual objects upon rapid head movement. Our observation implies that AR users with head-worn devices are vulnerable to notice degradation of quality of virtual objects when they rapidly rotate their head. In order to reduce the computation amount required for rendering, we apply Level of Detail concept [Heo00a] in accordance with head movement. That is, we intentionally decrease the quality of virtual objects to the level which AR users are not able to detect. In addition, virtual objects closer to eyes are more influenced by rapid head rotation. Therefore, in our approach, virtual objects closer to eyes are subject to more degradation of quality. To validate our approach, we establish a test platform that consists of a smartphone with a camera, various sensors, and a Gear VR which transforms a smartphone to a head-worn device. Experimental result shows that our approach is effective to reduce the latency of rendering during head movement.

2. RELATED WORK
The latency problem in AR systems has been identified by Donald T. Azuma *et al.* in [Azu97a]. The paper analyzed common problems caused by the latency in AR systems. Zhou *et al*. [Zho08a] have explored researches in a recent decade regarding major techniques in AR including tracking, interaction, and display.

In Computer Graphics, many researchers have proposed approaches to reduce the latency of rendering. Level of Detail [Heo00a] is a common solution for the purpose of rendering optimization. Various works including [Xia97a, Lin96a] have presented LoD based approaches for reducing the rendering latency. In addition to LoD, there have been various optimization techniques for rendering including real time occlusion culling analyzed by S. Coorg et al. [Coo97a] and a shader simplification technique proposed by P. Sitthi-amorn *et al.* [Sit11a].

More revolutionized head-worn devices with an optical see-through display such as Microsoft HoloLens [Hol16a] have been recently released and are expected to be the mainstream.

However, general rendering optimization approaches do not take into account AR and head worn devices.

3. PROBLEM OF LATENCY IN AR

Most AR systems recognize real objects and augment relevant information on the real objects by rendering virtual objects in various forms including images, texts, and 3D models. A virtual object has a relative location to a particular real object to provide specific information about it. A label on top of a real object can be an example of a virtual object that has a relative location [Azu03a]. For immersive experience, it is important to consistently preserve a relative location between a virtual object and a real object. However, operations in AR such as object recognition, and rendering of virtual objects inherently require certain amount of the latency to accomplish its purpose. On the other hand, the latency of perceiving a real object is nearly zero. Figure 1 illustrates the difference of the latency between real world and virtual objects.

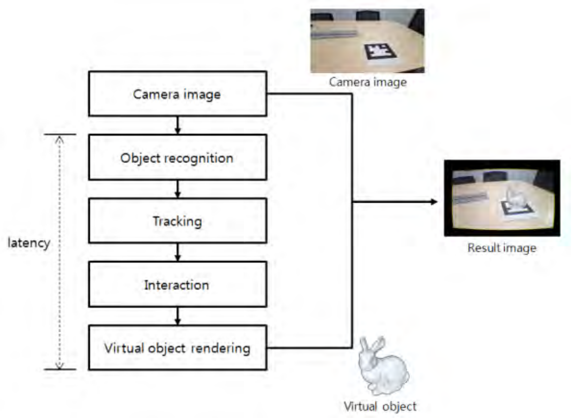

Figure 1. Latency difference between a camera image and a virtual object.

There is insignificant latency to acquire an image of real world. In order for a virtual object to be rendered, a large amount of the latency for various operations including object recognition and virtual object rendering is inevitable.

It is very difficult to retain assigned relative locations of virtual objects to real objects because of the difference of the latency. In other words, there is a spatial gap between virtual objects and real objects. We define this gap as *realism gap* which describes inconsistency of virtual objects overlaying onto real world. It gets worse when a head-worn device such as a smart glass and a head mount display is involved, because users change their view direction by head movement. Suppose that a user makes a rapid head turn. What a user sees in real world obviously changes in accordance with the transition. However, due to the latency, a virtual object still remains at a location where it was until rendering of a virtual object finishes. This realism gap not only disturbs immersive experience but also causes motion sickness. Reducing the latency to alleviate the realism gap is one of the keys for successful implementation of AR.

4. REDUCING LATENCY OF RENDERING

In order to reduce the latency, we concentrate on rendering of virtual objects. We observe how a user perceives a virtual object in rapid head movement. The first observation is that human visual system is vulnerable to notice insignificant change in a virtual object in the middle of rapid head movement. In fact, researches including [Bri75a] in various areas such as Vision, and Neurosciences have already explored a phenomenon called saccadic suppression which refers inability to detect changes in objects during rapid change in view direction. Our first observation complies with the researches. The second observation is that it is more difficult to notice change in a closer virtual object while a user rapidly rotates a head. From the observations, we conclude that it's acceptable to control quality of virtual objects in the middle of head movement to reduce the latency for rendering so that human visual system is unable to detect change in quality.

In Computer Graphics, Level of Detail, LoD, is one of the well-known approaches to reduce the amount of computation by controlling quality of virtual objects. Generally, the way LoD applies is based on distance of virtual objects from a camera [Lin96a]. LoD is able to reduce the amount of computation for rendering by decreasing quality of virtual objects distant from a camera. In this paper, to reduce the latency for rendering of virtual objects in AR, we apply Level of Detail particularly based on head movement in two aspects: 1) the amount of angular velocity of head rotation, 2) distance of virtual objects from a head.

The first criterion, angular speed of head rotation, makes sense according to numerous researches about saccadic suppression. Angular speed of head rotation determines quality of virtual objects by the following exponential function:

$$y = -b^{x-d} + 1 \quad (1)$$

x is angular speed of head rotation. We define angular speed of head rotation as the amount of change in angle in degree in a second. *y* represents quality of virtual objects. Values of quality of virtual objects are normalized. And value 1 indicates the original quality. *b* plays a role to determine a curvature of the function. *d* is a constant that denotes the maximum angular speed of head rotation. Maximum angular speed of head rotation is obviously limited because a human's ability to move a head is constrained. In case of head-worn devices, angular speed of head rotation practically ranges from 0 to 90 degree per second. We experimentally conclude that exponential functions are appropriate to account for the human visual perception in the middle of head rotation. Experiments we perform imply that decline of the ability to perceive change in virtual objects accelerates, as angular speed of head rotation increases.

The second criterion, distance of a virtual object from a head, also helps reduce quality of virtual objects. A projection that transforms virtual objects is perspective, which means 3D space distorts in such a way that closer areas are larger than farther areas after transformation. Therefore, it is more difficult to recognize movement of closer objects in the middle of head rotation. This approach that further decreases quality of closer objects is the opposite of general application of LoD. In general, farther objects have lower quality because they are less noticeable. However, in case of head-worn devices, we decrease quality of closer objects more, because rapid head rotation has more influence on closer objects. The function that determines quality of virtual objects from distance of virtual objects is as follow:

$$y = -\frac{1}{ax} + \frac{1}{f-n} + 1 \quad (2)$$

x is distance of a virtual object as an input. *f* denotes the farthest distance from a head. *n* is the nearest distance from a head. A perspective projection in 3D rendering typically accompanies a farthest plane and a nearest plane. *a* is a coefficient that determines a curvature of the function.

The final quality of a virtual object is defined as multiplication of outputs from (1) and (2). Therefore, quality of a virtual object is finally defined as

$$y = (-b^{x_1-d} + 1)(-\frac{1}{ax_2} + \frac{1}{f-n} + 1) \quad (3)$$

x_1 is angular speed of head rotation and x_2 is distance of a virtual object. The rest of variables and constants are already explained above.

For practical application, for head rotations with angular speed less than a particular value, it is possible for users to notice change of quality. Therefore, quality of virtual objects retains at angular speed of head rotation less than a threshold. We plan to experimentally identify an optimal threshold value.

To degrade quality of virtual objects, our approach introduces mesh simplification technique from [Tur92a]. The technique shrinks a mesh by merging a particular number of vertices into one vertex.

Although our approach does not involve the general way of applying LoD that decreases quality of farther objects, combined with the general way of applying LoD, more reduction of computation is possible.

5. EXPERIMENTS

To validate our approach, we establish a test platform that consists of a smartphone and a Gear VR which transforms a smartphone to a head-worn device. This headset is an affordable type of AR devices, as smartphones typically contain various sensors, a high resolution display, and a camera. Figure 2 shows the device included in the test platform. The test platform also has software implementation of AR including object recognition, object tracking, and rendering.

Figure 2. A head-worn device used for experiments

The detail of device specification is as follows. CPU is a combination of Cortex-A53 quad-core 1.5 GHz and Cortex-A57 quad-core 2.1 GHz. GPU is Mali-T760MP8 which includes 8 cores. The dataset used contains a set of virtual objects in a form of 3D meshes. The 3D meshes consist of 3,352,500 vertices. As head-worn devices require stereo images, our implementation renders two images for left and right eyes respectively. The resolution of each image is 1024x1024. Virtual objects in the dataset are lit by one directional light.

We measure the amount of time that rendering takes as the latency, in order to evaluate the improvement of our approach. Experiments also involve a conventional approach which renders virtual objects with constant quality for comparison. In experiments, constants b, d, a, f, and n in the equation (3) are set to 1.05, 90, 0.1, 2000, and 0 respectively.

The experimental result in figure 3 represents the relationship between angular speed of head rotation and the latency. We also alter distance values to observe the effect of distance to the latency. Distance values used are 15, 30, and 60. The latency of the conventional approach remains constant regardless of angular speed of head rotation or distance of a virtual object. The proposed approach significantly reduces the latency of rendering, as angular speed of head rotation increases. The latency decline accelerates to approximately 60 degree/sec, than decelerates. When the distance value is 60, the distance of a virtual object does not affect the latency. However, the latency at 10 degree/sec drops from 222 msec to 142 msec, when angular speed of head rotation increases to 55 degree/sec. Distance of a virtual object contributes to reduction of the latency as well. When angular speed of head rotation is 60 degree/sec, the latency at distance value 60 is roughly 22% of the latency at distance value 15.

Figure 5 illustrates output images of the proposed approach. A virtual object, Stanford Bunny, is rendered on top of a real object, a marker. It is possible to observe that a virtual object with the original quality in Figure 5 (a) changes to a version of lower quality in Figure 5 (b) during head rotation. We observe that it's hard to notice the quality degradation during head rotation.

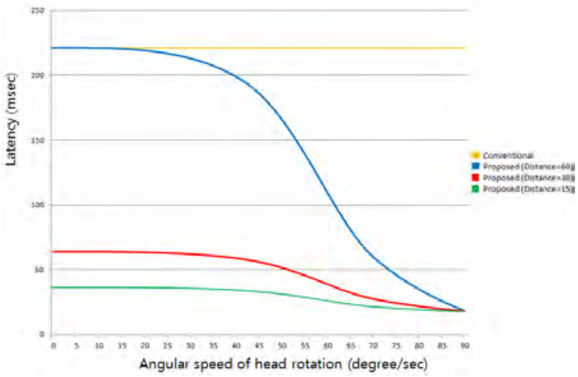

Figure 3. The latency of rendering from angular speed of head rotation and distance of a virtual object.

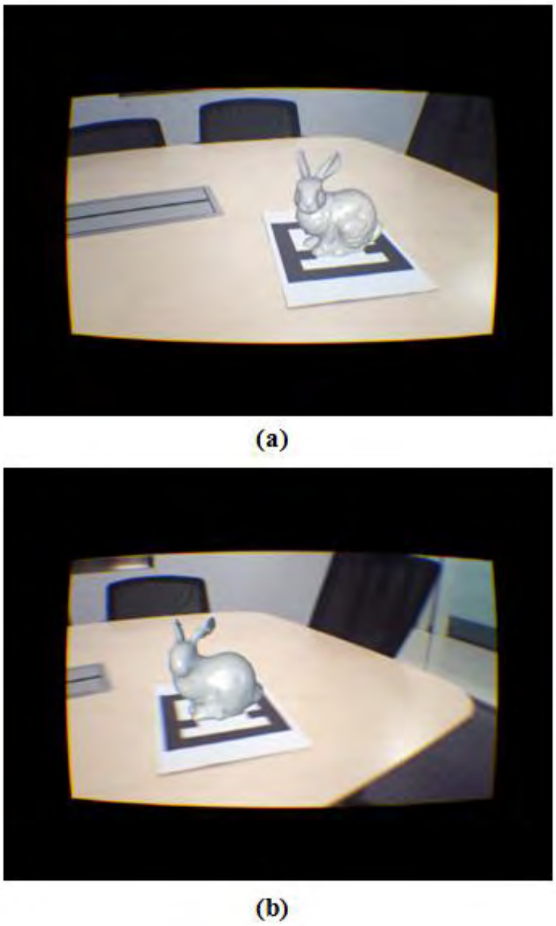

Figure 4. Quality change of a virtual object before (a) and during (b) head rotation.

6. Conclusion

The latency is one of critical problems in Augmented Reality. In this paper, we concentrate on reducing the latency of rendering, because rendering of virtual objects accounts for significant proportion of the entire latency in AR. By applying LoD based on head movement, the proposed approach successfully reduces the latency of rendering of virtual objects to 8% at best case.

7. Future work

We plan to continue developing and extending our approach for practical application. Since rendering of virtual objects is more complicated in practical cases, more decent control of quality is essential for efficient rendering of virtual objects. Our approach is expected to apply LoD to more factors such as display resolution, texture quality, ray tracing and so forth, in addition to the complexity of meshes. With more sophisticated control of quality, we plan to use real and practical datasets to present the effectiveness of the proposed approach.

ISBN 978-80-86943-58-9

In addition, to enhance other components of our implementation such as object recognition and object tracking, we consider employing a better solution such as Vuforia [Vuf12a].

8. REFERENCES

[Ocu12a] Oculus Rift, https://www.oculus.com, 2012.

[Goo14a] Google Cardboard Viewer,

https://www.google.com/get/cardboard/, 2014.

[Sam15a] Samsung Gear VR, http://www.samsung .com/global/galaxy/wearables/gear- vr/, 2015.

[Van10a] Van Krevelen, D. W. F., & Poelman, R. A survey of augmented reality technologies, applications and limitations. International Journal of Virtual Reality, 9(2), pp. 1-20, 2010.

[Zho08a] Zhou, Feng, Henry Been-Lirn Duh, and Mark Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality. IEEE Computer Society, pp. 193-202, 2008.

[Heo00a] Heok, Tan Kim, and Daut Daman. A review on level of detail. Computer Graphics, Imaging and Visualization, CGIV 2004. Proceedings. International Conference on. IEEE, pp. 70-75, 2004.

[Azu97a] Azuma, Ronald T. . *A Survey of Augmented Reality*. Teleoperators and Virtual Environments 6.4, pp. 355-385, 1997.

[Xia97a] Xia, Julie C., Jihad El-Sana, and Amitabh Varshney. Adaptive real-time level-of-detail based rendering for polygonal models. Visualization and Computer Graphics, IEEE Transactions on 3.2, pp. 171-183, 1997.

[Lin96a] Lindstrom, Peter, et al. Real-time, continuous level of detail rendering of height fields. Proceedings of the 23rd annual conference on Computer graphics and interactive techniques. ACM, pp. 109-118, 1996.

[Coo97a] Coorg, Satyan, and Seth Teller. Real-time occlusion culling for models with large occluders. Proceedings of the 1997 symposium on Interactive 3D graphics. ACM, pp. 83-ff, 1997.

[Sit11a] Sitthi-Amorn, Pitchaya, et al. Genetic programming for shader simplification. ACM Transactions on Graphics (TOG) 30.6, Article No.152, 2011.

[Hol16a] Microsoft HoloLens, https://www.microsoft .com/microsoft-hololens/en-us, 2016.

[Azu03a] Azuma, Ronald, and Chris Furmanski. Evaluating label placement for augmented reality view management. Proceedings of the 2nd IEEE/ACM international Symposium on Mixed and Augmented Reality. IEEE Computer Society, pp. 66, 2003.

[Bri75a] Bridgeman, Bruce, Derek Hendry, and Lawrence Stark. Failure to detect displacement of the visual world during saccadic eye movements. Vision research 15(6), pp. 719-722. 1975.

[Lin96a] Lindstrom, Peter, David Koller, William Ribarsky, Larry F. Hodges, Nick Faust, and Gregory A. Turner. Real-time, continuous level of detail rendering of height fields. In Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, ACM, pp. 109-118, 1996.

[Tur92a] Turk, Greg, and Marc Levoy. Zippered polygon meshes from range images. Proceedings of the 21st annual conference on Computer graphics and interactive techniques. ACM, 1994.

[Vuf12a] Vuforia AR, http://www.vuforia.com, 2012.