waveSense: Ultra Low Power Gesture Sensing Based on Selective Volumetric Illumination

Anusha Withana, Shanaka Ransiri, Tharindu Kaluarachchi, Chanaka Singhabahu, Yilei Shi, Samitha Elvitigala, Suranga Nanayakkara
Augmented Human Lab, Singapore University of Technology and Design, Singapore
{anusha | shanaka | tharindu | chanaka | yilei | samitha | suranga}@ahlab.org

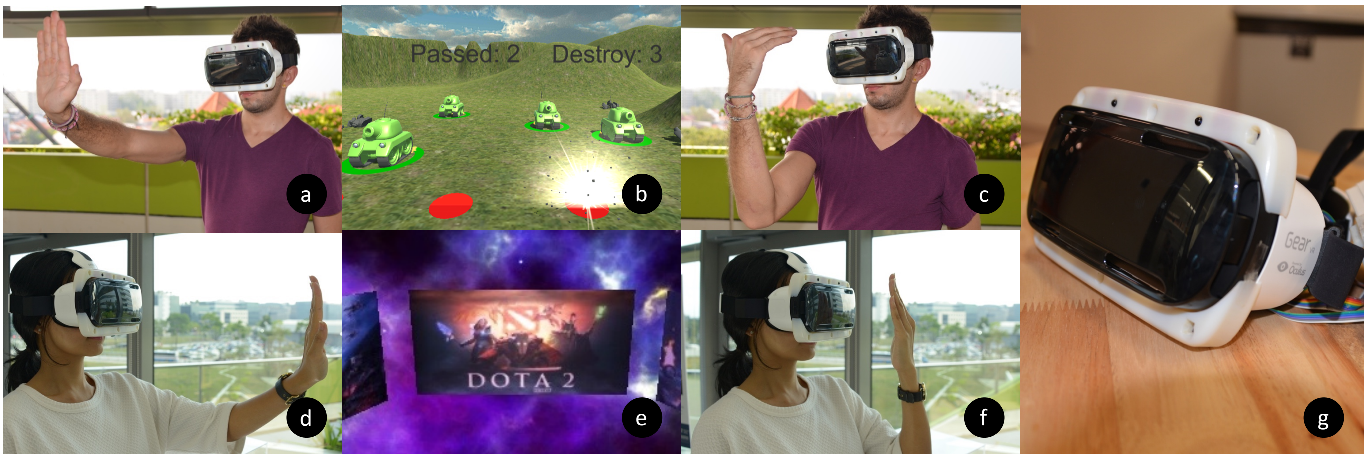

Figure 1: Hand gesture recognition with *waveSense*: a), b), c) A user playing a virtual reality (VR) game using waveSense where he uses an *extension of hand* gesture towards the target direction to destroy an enemy tank d), e), f) Usage of swipe gestures to browse through a virtual reality image gallery, left and right swipes help to navigate through images and push and pull gestures to select an image; g) *waveSense* sensing module for Samsung Gear VR

ABSTRACT
We present *waveSense*, a low power hand gestures recognition system suitable for mobile and wearable devices. A novel *Selective Volumetric Illumination* (SVI) approach using off-the-shelf infrared (IR) emitters and non-focused IR sensors were introduced to achieve the power efficiency. Our current implementation consumes $8.65mW$ while sensing hand gestures within $60cm$ radius from the sensors. In this demo, we introduce the concept and the theoretical background of *waveSense*, details of the prototype implementation, and application possibilities.

Author Keywords
Hand Gesture Recognition; Interacting with virtual reality devices; Smart wearables; Selective volumetric illumination; Compressive sensing

ACM Classification Keywords
H.5.2 User Interfaces: Input devices and strategies

UIST'16 Adjunct, October 16-19, 2016, Tokyo, Japan
ACM 978-1-4503-4531-6/16/10.
http://dx.doi.org/10.1145/2984751.2985734

INTRODUCTION
Introduction of virtual reality devices such as *Oculus* and *Samusng GearVR* calls for interaction methods that goes beyond the traditional touch / button based interfaces. Traditional computer vision based gesture recognition systems ([2, 3]) require larger physical space and energy resources that are extremely limited in such compact devices. Other sensing approaches such as non-focused IR ([1, 5]) or magnetic markers ([4]) has limited sensing range and expressivity making them unsuitable for many, including virtual reality applications. Alternatively, recent works such as *Mime* proposed a gesture based interaction technique that require low power; but its time of flight based sensing requires a significantly high processing power. In this work, we propose *waveSense*,a gesture sensing technology that consumes low battery power and processing power that is suitable for devices with limited resources.

To achieve this, we developed a novel gesture sensing technique based on Selective Volumetric Illumination (SVI), inspired from the non-linear spatial sampling (NSS) scheme introduced in *zSense* [6]. *zSense* switches multiple IR emitters to create a 2D spatial illumination patterns. Reflected light from a target was measured as an analog value (0 to 255) to estimate the depth. This works efficiently for low ranges ($\sim 15cm$). To extend the range with zSense, high power IR emitters are required (no power saving) and analog sensing is vulnerable to external noise (low accuracy). The key technical innovation in waveSense is that it excites IR emitters with discrete power levels (8) with a modulated signal, creating different 3D illumination patterns, which can be sensed digitally (reflected signal exist as 1 or no signal as 0). This allows us to extend the sensing range ($\sim 60cm$) while keeping emitters at low power due to improved noise immunity in digital sensing. The contributions of this paper can be summarized as, 1) Introduction of Selective Volumetric Illumination (SVI) for low power and low cost gesture sensing; 2) Prototype implementation of *waveSense*; 3) Implementation of various application scenarios of *waveSense* for vitual reality glasses.

PROOF OF CONCEPT

Our main goal with *waveSense* was to develop a viable (low power, low processing overhead) gesture recognition system for resource limited interactive systems. *waveSense* trades off the hand tracking resolution to save power while keeping the capability of accurately recognizing a reasonable number of expressive gestures to interact with intended applications.

Prototype

Current prototype of *waveSense* shown in figure 1g for Samsung gearVR consist of 2 modulated ($57.6kHz$) IR emitters excited at 8 distinct power levels ($1.05mW$ to $18.3mW$) sequentially, creating a total of 16 different *SVI* patterns. Six demodulating digital IR receivers are embedded in the prototype with relative displacement in the directionality (Similar to spatial displacement in [6]). Each sensor captures a single bit of data for a given pattern, resulting a string of 96 bits of data per iteration. Forty such iterations are captured per second resulting 40fps capture rate.

The functional block diagram of *waveSense* is shown in figure 2a. The system is controlled using a Nordic's nRF51822 (16MHz, 3.3V) System on Chip (SoC) (Figure 2d) which has a built-in Bluetooth Low Energy (BLE) to communicate with the devices such as smartphones. The microprocessor modulates the IR emitters at $57.6kHz$ and a current control circuit is used to achieve *SVI*. Each *SVI* pattern is kept on for 210uS and the reflected light is captured by the IR receivers which are self contained with an Automatic Gain Controller (AGC) and a demodulator. Output of the sensor is read by the microprocessor and send to the host device for classification.

Similar to *zSense*[6], classification of the gestures are done in two stages. 1) *Level classification*: classify the x, y, z locations of a target using raw data. 2) *Gesture classification*: x,y,z locations are sent to another classifier using a time domain moving window to recognize the gestures (Figure 2a). Both steps use BayesNet SMO classifier from *Weka* toolkit. *Level classification* training is user independent (i.e. does not have to be done per user) and can be kept common for all the users. Per user training is necessary for the *gesture classification* to cater for subjective differences. Our pilot study shows the current prototype has a *level classification* accuracy of 95% for a $5 \times 4 \times 2$ grid in x, y, z axes resulting 40 distinct locations.

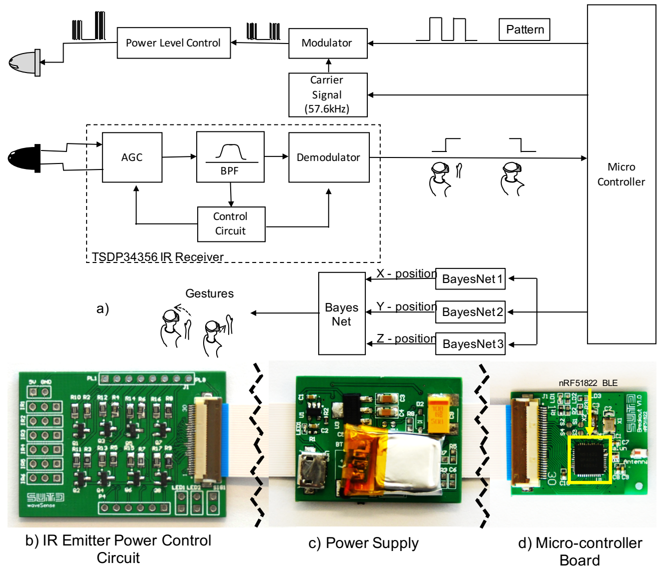

Figure 2: a) Hardware block diagram, and b), c), d) the implementation of embedded control system

INTERACTIVITY AND APPLICATIONS

We developed three applications using *Unity5* for Samsung GearVR; 1) *Tank destroyer* first-person game (figure 1a, b, c), 2) image gallery (Figure 1d, e, f) and 3) 3D fruit ninja game (see attached video) in order to demonstrate the potential interactions enabled by *waveSense*. Ten distinct gestures; four *horizontal swipes* (close range right and left, far right and left), two *vertical swipes* (towards left and right side) and four *push-pull* (in four radial directions) have been considered when developing these three applications. We believe these applications with *waveSense* based gesture interactions will allow users to have a truly immersive and entertaining experience with 3D virtual reality content.

REFERENCES

1. A. Butler, S. Izadi, and S. Hodges. Sidesight: multi- touch interaction around small devices. In *Proc. of UIST '08*, pages 201–204, 2008.
2. S. Gustafson, D. Bierwirth, and P. Baudisch. Imaginary interfaces: Spatial interaction with empty hands and without visual feedback. In *Proc. of UIST '10*, pages 3–12, 2010.
3. C. Harrison, H. Benko, and A. D. Wilson. OmniTouch: Wearable multitouch interaction everywhere. In *Proc. of UIST '11*, pages 441–450, 2011.
4. C. Harrison and S. E. Hudson. Abracadabra: wireless, high-precision, and unpowered finger input for very small mobile devices. In *Proc. of UIST '09*, page 121, 2009.
5. K. Nakatsuma, H. Shinoda, Y. Makino, K. Sato, and T. Maeno. Touch interface on back of the hand. In *proc. of SIGGRAPH '11*, pages 39:1–39:1, 2011.
6. A. Withana, R. Peiris, N. Samarasekara, and S. Nanayakkara. zsense: Enabling shallow depth gesture recognition for greater input expressivity on smart wearables. In *Proc. of CHI '15*, pages 3661–3670, New York, NY, USA, 2015. ACM.