## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>3</td>
</tr>
<tr>
<td>Introduction</td>
<td>3</td>
</tr>
<tr>
<td>Current Gesture Sensor Trends</td>
<td>4</td>
</tr>
<tr>
<td>Gesture Sensor Technology</td>
<td>5</td>
</tr>
<tr>
<td>Non-Image Technology</td>
<td>5</td>
</tr>
<tr>
<td>Image Technology</td>
<td>5</td>
</tr>
<tr>
<td>Requirement of Gesture Sensing for Mobile Devices</td>
<td>8</td>
</tr>
<tr>
<td>Low-Power Consumption</td>
<td>8</td>
</tr>
<tr>
<td>Small Form Factor</td>
<td>8</td>
</tr>
<tr>
<td>Dynamic Vision Sensor</td>
<td>9</td>
</tr>
<tr>
<td>Summary</td>
<td>10</td>
</tr>
<tr>
<td>References</td>
<td>11</td>
</tr>
</tbody>
</table>
Abstract

Gesture sensor is used to detect and recognize meaningful gestures and expressions from the parts of the human body such as hands, arms, face, head and so on.

Most of the gesture sensors are mainly based on image sensor. And Image sensor technologies have been focused on depth information. The various kinds of image sensor for depth information are stereo type, structured light, time of flight, plenoptic and array sensor.

However, after the advent of smart phone, other technologies such as touch IC and proximity sensor are entering in the gesture sensing market with low power consumption. Mobile devices need high accuracy and low power consumption. Therefore, a second iteration of gesture sensor technologies has commenced.

This white paper provides a survey on gesture sensor and its trends for mobile devices. This paper also deals with existing challenges and future research possibilities. Finally, dynamic vision sensor is proposed as the most promising device for gesture sensor.

Introduction

After the popularization of smart phones, there has been a very keen interest in its user interface (UI). Among the smart phones, Apple® leads in touch screen UI. In October 2011, Apple introduced Siri® a new interface of smart phone, by using voice recognition[1]. Other mobile companies are struggling to adopt the new UI for smart phones and tablets.

Gesture recognition technology is one of the new UIs. This UI was launched in Xbox360® Kinect® of Microsoft® in November 2010 [3]. Thereafter, Samsung adopted gesture detection in TV in January 2012 and Pantech® brought this technology for their smart phone in December 2011[4][5].

Samsung released the new gesture UI by using the proximity IC in Galaxy S4 as illustrated in Figure 3[7].
Figure 3: Hidden Innovation in the Galaxy S4

However, the gesture-based UI is not as popular when compared to touch screen in mobile devices because of the undeveloped technology.

Current Gesture Sensor Trends

Typically, the meaning of a gesture depends on:

- Spatial information: Where it occurs
- Path information: The path it takes
- Symbolic information: The sign it makes
- Affective information: Its emotional quality

According to this definition, gesture sensor is important to detect spatial information that includes depth and distance.

Figure 4 illustrates the gesture sensor category to obtain the spatial information. The non-image-based sensor includes touch IC and proximity IC. The image-based sensor includes image sensor + software, time of flight (ToF), structured light.

Microchip Technology and Maxim® use non-image based gesture technology.

Microchip Technology released their new GestIC® Technology that enables the mobile-friendly 3D gesture information. This technology is the 3D gestures controller based on electrical field, which is the technology of touch IC. Maxim released an article, Touching Air: Wake your device with a simple gesture. This article demonstrates that the infrared (IR) proximity sensors can detect gestures by using MAX44000.

The three kinds of image-based gesture sensor are:

- 2D legacy image sensor and software: eyeSight™, PointGrap®, and XTR3D® are the companies that use this technology. The S/W algorithm detects the motion.
- ToF: PMD Technologies, MESA Imaging, and SoftKinetic® are the companies that use this technology. ToF provides high accuracy for motion detection when compared to other technologies.
- Structured light: PrimeSense™, which provides for Microsoft and Xbox360 Kinect, uses this technology. This is a well-commercialized technology for consumer electronics. It announced the advanced motion sensing by using IR light in May 2012. This technology is different category in Figure 4.

**Figure 5: Leap Motion™**

Gesture sensor is one of components for the gesture UI. Gesture sensor requires the software to control the gesture sensor itself and to recognize the motion. Therefore, most companies collaborate with others to secure the gesture-sensing solution.

The two categories of the collaboration are:

- Set development
- Solution development

The collaboration of Xbox360 Kinect between Microsoft and PrimeSense and the collaboration of CT510 between EEDOO and SoftKinetic is an example for set development.

**Figure 6: EEDOO CT510™**

There is the progress to migrate and collaborate for solution development. Texas Instrument® (TI) and Omek developed the engine of gesture recognition for OMAP™ 3530 in March 2011 [13]. Also, Qualcomm® acquired gesture-recognition assets from GestureTek in July 2011 and released Snapdragon™ S4 based on this acquired technology. CEVA® and eyeSight announced their partnership in November 2010 and released the gesture solution in June 2012.

### Gesture Sensor Technology

This section includes:

- Non-image technology
- Image technology

#### Non-Image Technology

GestIC of Microchip technology is a patented 3D sensor technology that uses utilizes an electric field for advanced proximity sensing [10]. It enables motion detection with low-power consumption. However, it is impossible to distinguish and detect gestures between the palm, fist, and fingers simultaneously.

The gesture sensing technology of Maxim is based on the proximity sensing as illustrated in Figure 7. However, more power is required to operate light source such as IR LED.

**Figure 7: Proximity Sensor**

#### Image Technology

According to information to detect motions, gesture sensing that based on image sensor is classified into:

- 2D imaging
- 3D imaging

In 3D imaging, based on the kinds of image sensors, it can be classified into:

- Active ranging camera: Uses additional light source such as IR LED and IR laser to gather complete information. Therefore, the specialized image sensor such as ToF and IR detection sensor are important.
- Passive ranging camera: Uses high performance image signal processing (ISP).
- Light field photography: Uses optical technology and requires specific ISP to analyze the optical information for depth map
Passive Ranging Camera

Passive ranging camera is categorized into:

- Single camera solution
- Stereo camera solution

• Single Camera Solution

In single camera solution, the key technology is ISP. It includes object detection, extraction, and tracking on 2D space. eyeSight, XTR3D and pointGrap are the companies that use single camera solution[14].

Figure 8: Samsung TV Gesture Control[15]

There is a technology based on 2D image to gather 3D information. Figure 9 illustrates four kinds of technologies to gather 3D information[16].

- Focus Control
- Visual Field Size
- Aerial Perspective
- Motion Parallax

Figure 9: Methodology of Monocular Disparity

Focus control method detects the distance by using the size and contrast of focused objects.

Visual field size method uses a relative size familiar with objects such as a tree and rock.

The principle of aerial perspective is that near objects are clearer than far objects.

Motion parallax detects the distance from the difference of speed between the near and far object.

• Stereo Camera Solution

A stereo camera is a type of camera with multiple lenses with a separate image sensor that enables the camera to simulate human binocular vision. It captures 3D images[17].

Figure 10: Stereo Camera Phone: SCH-8710[18]

Active Ranging Camera

Active ranging camera uses IR light source to obtain high accuracy depth information. It includes:

- Structured light
- ToF

• Structured Light

Structured light uses one or multiple sources of light to project a known pattern of pixels[19][20].

A structured light sensor is similar to a classical stereoscopic sensor. A light source such as a laser or a slide project replaces the camera. Single line, cross hair, parallel lines, single, square, dot line, 7x7 dot matrix, single circle, and seven concentric circles are the examples of patterns as illustrated in Figure 11[21].
**Time of Flight**

ToF extracts complete depth information from phase delay of modulated light, typically light emitting diode (LED) array of near-infrared wavelength after reflection from the light \(^{21}\).

**Light Field Photography**

Light field camera uses a micro lens array to capture 4D light field information about a scene \(^{22}\). Light field camera can be categorized into:

- **Plenoptics**
- **Array sensor**

**Plenoptics**

Plenoptics is achieved by inserting a micro lens array between the sensor and main lens \(^{22}\). Each micro lens measures the light that arrives along each ray instead of the total amount of light. By sorting the measured rays of light, which is measuring where they would have terminated in different synthetic cameras, Plenoptics can compute sharp photographs focused at different depth.

**Array Sensor**

Array sensor uses micro lens array with wafer-level optics. Similar to Plenoptics, array sensor acquires additional information to process 3D imaging. Pelican imaging uses the array sensor technology. At less than 3 mm thickness, the array camera of Pelican is about 50 percent thinner than current smart phone cameras. The architecture of Pelican does not require moving parts for immediate first shot with complete focus \(^{23}\).

Image-based 3D technologies are compatible to conventional system. Therefore, it makes vibrant researches in academia and the industrial world.
Requirement of Gesture Sensing for Mobile Devices

The two kinds of requirements in mobile device are:

- Power consumption
- Size

Low-Power Consumption

Table 1 lists the legacy smart phone features. According to this table, hardware features maintain the high performance of the smart phone.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Galaxy S4@ 2013</th>
<th>Galaxy S3@ 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Quad core 1.9 Ghz</td>
<td>Quad core 1.4 Ghz</td>
</tr>
<tr>
<td>Display</td>
<td>5.0 inch (1080p)</td>
<td>4.8 inch (720p)</td>
</tr>
<tr>
<td>Camera</td>
<td>13 MP</td>
<td>8 MP</td>
</tr>
<tr>
<td>Battery</td>
<td>2600 mAh</td>
<td>2100 mAh</td>
</tr>
</tbody>
</table>

Table 1. Smart Phone Features

However, smart phone users complain low battery capacity. Battery technology cannot keep pace with other technology development such as display, application processor, speed and so on. For mobile device, every component requires the low-power consumption to adopt in mobile system.

Therefore, low-power consumption is mandatory for gesture sensor because the UI is always in use in mobile devices.

Although there are some gesture sensors such as touch IC and proximity sensor to fulfill the criteria of the power consumption, the performance is not up to the mark.

In image-based technology, additional light source or sensor is physically required or software is required systemically to obtain 3D information in Table 2.

There is a compromise between the low-power consumption and performance. This creates the first barrier to expand gesture UI in mobile devices.

Small Form Factor

The design of the smart phones should be attractive to the end users. Therefore, there are two kinds of trends in the designs:

- Slim design
- Zero bezel design

The display of Galaxy S4 (5 inches) is bigger than the display of Galaxy S2 (4.3 inches) by 16 percent. However, dimension of Galaxy S4 has increased by 8 percent when compared to Galaxy S2. Galaxy S4 (7.9 mm) is thinner than Galaxy S2 (8.49 mm) and the width is almost same. This comparison indicates the trends of slim design and the zero bezel design.

These design trends create second barrier to adopt gesture sensor in mobile devices. Especially, the size limitation is the big barrier for image-based technologies. As described in Table 2, most of the image-based gesture sensors require additional light source or sensor or specialized optics to create a bigger module. However, the size limitation is an opportunity for touch-based gesture sensor and proximity-based gesture sensor because these kinds of gesture sensor can be merged into touch IC and commodity proximity sensor.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Galaxy S2</th>
<th>Galaxy S4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display</td>
<td>4.3 inch</td>
<td>5.0 inch</td>
</tr>
<tr>
<td>Dimension</td>
<td>125.3x66.1x8.49mm</td>
<td>136.6x69.8x7.9mm</td>
</tr>
</tbody>
</table>

Figure 15. Galaxy S2 and Galaxy S4

Table 2. Image-Based Gesture Sensor
Dynamic Vision Sensor

Gesture sensor in mobile devices is at a nascent stage. Although many kinds of sensors are considered for gesture UI applications in mobile devices, none of them could take a leading position because of the performance issues. Conventional image-based methods have been focusing on the high performance such as high accuracy of depth. However, it consumes high power, delays in processing the images and requires additional component. Alternatively, touch-based technology and proximity-based technology are used for the gesture sensor with low power consumption. However, they also include disadvantages, such as low accuracy and small detection area. Therefore, gesture sensor is required for high performance and low power consumption with small form factor.

For the high performance, the gesture sensor output is compatible with the conventional gesture algorithms. The gesture sensor output includes feature extraction, tracking, matching and recognition. Gesture sensor consumes more power to run the algorithm because it generally requires computational steps such as feature extraction, tracking, matching, recognition, and so on as illustrated in Figure 16.

If the gesture sensor performs any of these processing step by itself, you can reduce the load of main processor which runs the gesture-sensing algorithm considerably. In this view, Dynamic Vision Sensor (DVS) is proposed to be the most promising device for gesture sensor. DVS can effectively detect a moving object in a cost-effective way with fast response by providing events asynchronously on its edges as illustrated in Figure 17. The DVS only sends the information of local pixels. The relative intensity changes instead of entire images at fixed frame rates. It can also detect illumination changes that occur in a few microseconds, which makes fast detection of movement possible.

Also, conventional gesture algorithm can use the output of DVS to detect and analyze gestures without large modification. The DVS can track the finger tip with adaptive interval time and detect direction of hand motions with high response time with 98.7 percent accuracy as illustrated in Figure 18. Also, adaptive interval time can consumes low power during idle mode.

The power consumption of DVS chip, 128x128 resolution with 40µm x 40µm pixel is 23 mW. This is at least five times lower than the legacy front facing camera.

Figure 19 illustrates two benefits of DVS for low power consumption. Firstly, it reduces the burden of processing by using edge detected images. Secondly, the low power consumption supports idle mode based on the adaptive interval time to detect the moving object.
DVS does not require additional light source and specialized optics to detect gestures when compared to stereo 3D, TOF, structured light, plenoptics and array sensor. Also, the conventional output of the image sensor is in successive frames that contains redundant information. This information consumes memory. However, the output of DVS is optimized for gesture detection and does not require the memory for redundant information. Therefore, it makes small form factor to adopt in mobile devices.

The dynamic range of DVS extends 120 dB from 60~70 dB. The DVS can detect the moving object in extreme low light and ultra high light. However, the IR-based gestures sensors such as ToF and structured light have a limitation to obtain the information in outdoors.

**Summary**

This paper describes the current gesture sensor trends such as non-image based gesture sensor and image-based gesture sensor and it provides details of gesture sensor technologies.

In mobile devices, low power consumption and small form factor are the two kinds of barrier. Also, it is unsure that high accuracy of gesture sensing is mandatory in mobile devices. If a new UI-based on gesture has launched, then the high accuracy of gesture sensing is an important factor to select gesture sensor.

This paper describes DVS as the most promising device for gesture sensor. DVS has benefits of low power consumption and small form factor while maintaining high accuracy.

The next generation of gesture sensor will be launched into the mobile devices based on an indepth understanding mobile device requirement such as low-power consumption, small form factor, and trends of new UI.
References


About Samsung Electronics Co., Ltd.

Samsung Electronics Co., Ltd. is a global leader in technology, opening new possibilities for people everywhere. Through relentless innovation and discovery, we are transforming the worlds of televisions, smartphones, personal computers, printers, cameras, home appliances, LTE systems, medical devices, semiconductors and LED solutions. We employ 236,000 people across 79 countries with annual sales exceeding US$187.8 billion. To discover more, please visit www.samsung.com.

For more information

For more information about Samsung Semiconductor, visit http://www.samsung.com/global/business/semiconductor/

For more information about Samsung Exynos Processor, visit
http://www.exynos.com
http://facebook.com/SamsungExynos
http://twitter.com/SamsungExynos
http://youtube.com/SamsungExynos

Copyright © 2013 Samsung Electronics Co. Ltd. All rights reserved. Samsung is a registered trademark of Samsung Electronics Co. Ltd. Specifications and designs are subject to change without notice. Non-metric weights and measurements are approximate. All data were deemed correct at time of creation. Samsung is not liable for errors or omissions. All brand, product, service names and logos are trademarks and/or registered trademarks of their respective owners and are hereby recognized and acknowledged.

Apple and Siri are registered trademarks of Apple Inc.
Microsoft, Xbox360 and Kinect are registered trademarks of Microsoft Corp.
Pantech is a registered trademark of Pantech Co., LTD.
Maxim is a registered trademark of Maxim Integrated Products, Inc.
GestIC is a registered trademark of Microchip Technology Incorporated
PointGrab is a registered trademark of of POINTGRAB LTD.
XTR3D is a registered trademark of EXTREME REALITY LTD.
SoftKinetic is a registered trademark of SoftKinetic.
Prime Sense is a trademark of PrimeSense Ltd.
Leap Motion is a trademark of Leap Motion, Inc.
Texas Instrument and OMAP is a trademark or registered trademark of Texas Instruments.
Omek is a registered trademark of Omek Interactive.
Qualcomm and Snapdragon are trademarks of Qualcomm Incorporated, registered in the United States and other countries.
CEVA is a registered trademark of CEVA, Inc.
eyeSight is a trademark of eyeSight Mobile Technologies.
Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc.,
Google and the Google Logo are registered trademarks of Google Inc.
Lytro is a registered trademark of Lytro, Inc.
Pelican Imaging is a trademark of Pelican Imaging Corporation.

Samsung Electronics Co., Ltd.
416, Maetan 3-dong,
Yeongtong-gu
Suwon-si, Gyeonggi-do 443-772,
Korea

www.samsung.com