embedded VISION SUMMIT

Vision Based Gesture Recognition

Francis MacDougall, Sr.Director, Qualcomm, Inc.

QUALCOMM®

October 2, 2013

QUALCOMM®

Fortune 500 Company

> $100USD Billion Market Cap

Celebrating more than 25 years of driving the evolution of wireless communications.

Over 23,000 employees worldwide.

World's largest fabless semiconductor company, #1 in wireless.

Member of the S&P 100 & 500 Indexes

GREAT PLACE TO WORK® INSTITUTE INDIA : TOP 25 BEST companies to work for 2010

FORTUNE 500 2011

FAST COMPANY 2010 MOST INNOVATIVE COMPANIES

Institutional Investor

2011 All-America Executive Team Award Winner

Forbes 2010 WORLD'S LEADING COMPANIES GLOBAL 2000

FORTUNE 100 BEST COMPANIES TO WORK FOR 2011

COMPUTERWORLD 100 BEST PLACES TO WORK IN IT 2010

FORTUNE WORLD'S MOST ADMIRED COMPANIES 2011 #1 - NETWORK AND OTHER COMMUNICATIONS EQUIPMENT

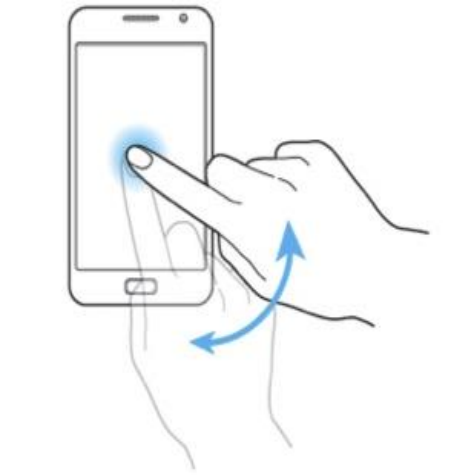
What is Gesture Recognition?

Perceptual User Interface (PUI) Technologies

PUI Tech	Overview	Technology
Face Detection	The ability to find any face in an image, often with orientation using head pose estimation.	2D/3D/IR Vision
Face Recognition	The ability to discern an individual's identity from an image.	2D/3D Vision
Eye Tracking	Determination of gaze direction and blink from a series of images.	2D/3D/IR Vision
Emotion Sensing	User emotional state determination using visual cues from one or more images. Includes smile detection, etc.	2D Vision
Lip Movement Recognition	Interpreting speech through lip reading from a series of images.	2D Vision
Gesture Recognition	Interpretation of hand and body pose and motion.	2D/3D/IR Vision Ultrasound / Dedicated Sensors
Speech Recognition	The ability to receive and interpret voice commands from an audio stream.	Audio

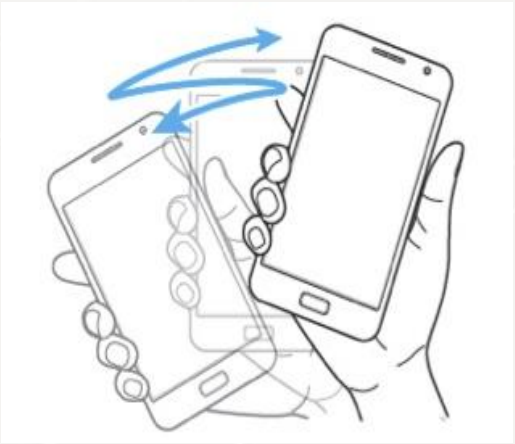
Gesture Types

Touch Gestures



- Pinch to Zoom
- Scroll
- Click & Drag

Motion Gestures

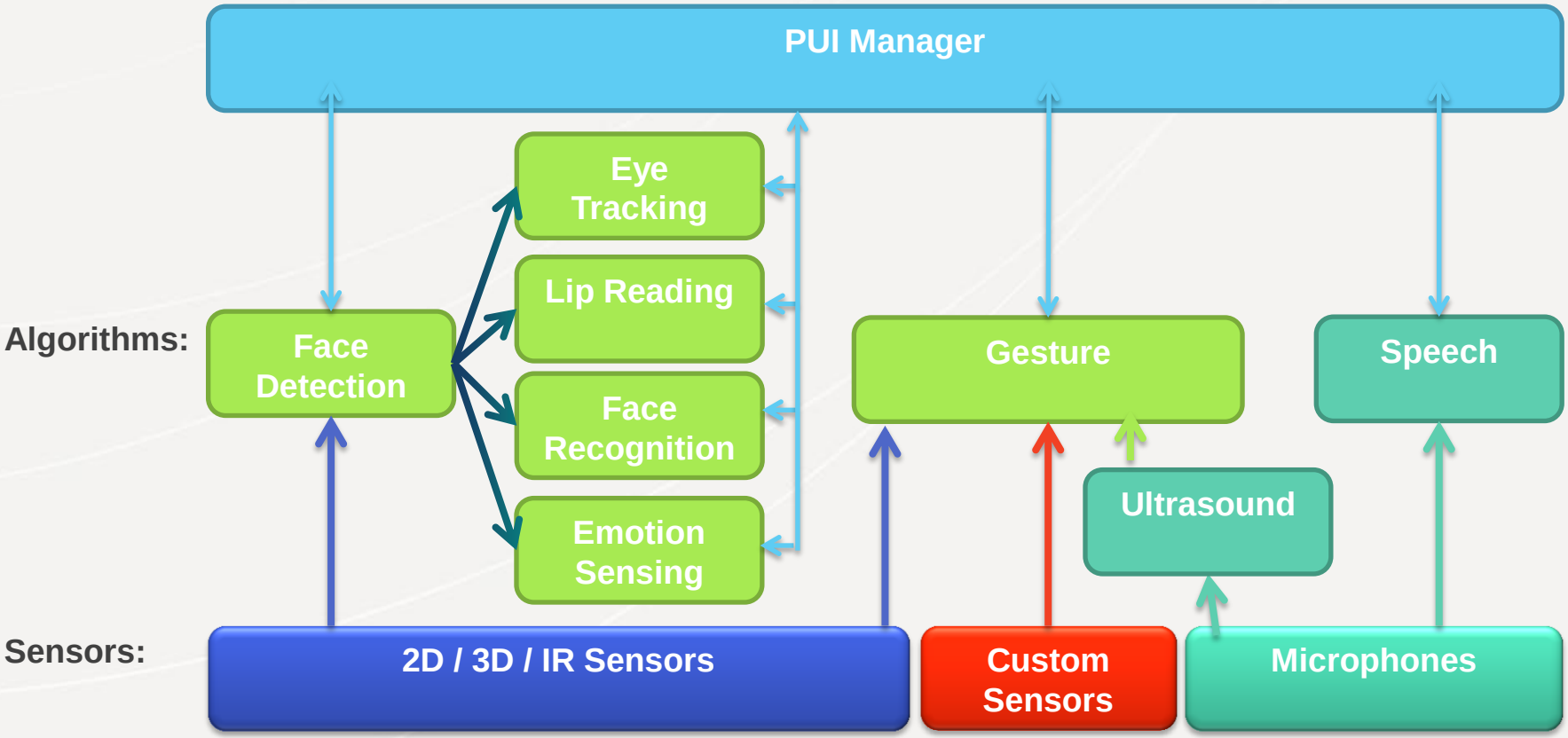


- Turn over to Mute/Pause
- Shake to Update
- Bump
- Rotate for landscape

Touchless Gestures

- Control devices from ½” to 15 feet away

Perceptual User Interface Architecture



Gestures use a variety of sensors, and can also leverage other PUI tech including Eye Gaze

Why Gestures?

Gaming / Fitness

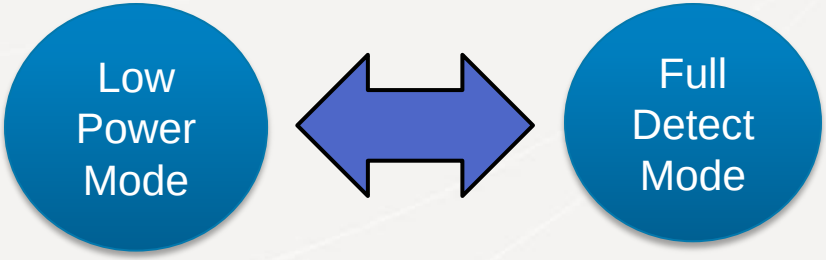
KINECT™

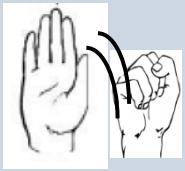
for XBOX 360

Samsung SmartHub Televisions

Gesture Application Types: Always On

Engagement gesture followed by mouse emulation with click.



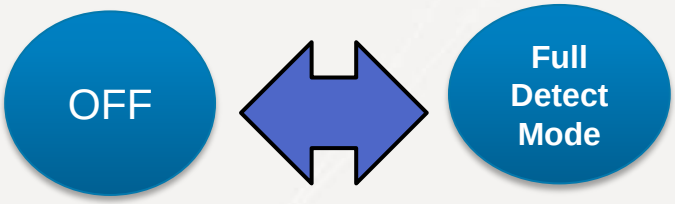
Gesture	Description	Uses
Hand Wave Engagement 	Detects a forward facing palm pose in the camera view.	Initial engagement of gesture system.
Hand position Tracking 	Hand position is tracked and provides (x,y) positions.	On-screen mouse cursor control.
Hand close For “click” 	Hand moves from open palm position to closed fist position to effect a click event.	Mouse click emulation.

Samsung S4 – the next big thing is here…

Gesture Application Types: Event Driven

"***Event Driven***" applications:

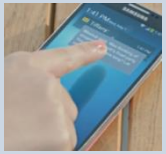
- Can run sensor at full speed for short duration and detect gestures



Call Answer

- Turns on sensor when call is received
- Swipe to answer call and turn on speakerphone
- Alternate swipe can send call to v-mail
- Secondary swipes to mute and hang up

Notification View

Gesture	Description	Uses
Hover 	Tracks a fingertip hovering above the display within 1 inch of the display	View notifications, view extra button information (tool tips), emulate hover on web pages. Zoom in the area below the finger to make a click easier.
Near Swipe 	Tracks a hand passing within 6 inches in front of the camera sensor. Detects swipes across the sensor.	Touch-free call answer, in-car gesture mail "no-look" control, page turns, etc.

Samsung S4 – touch free gestures…

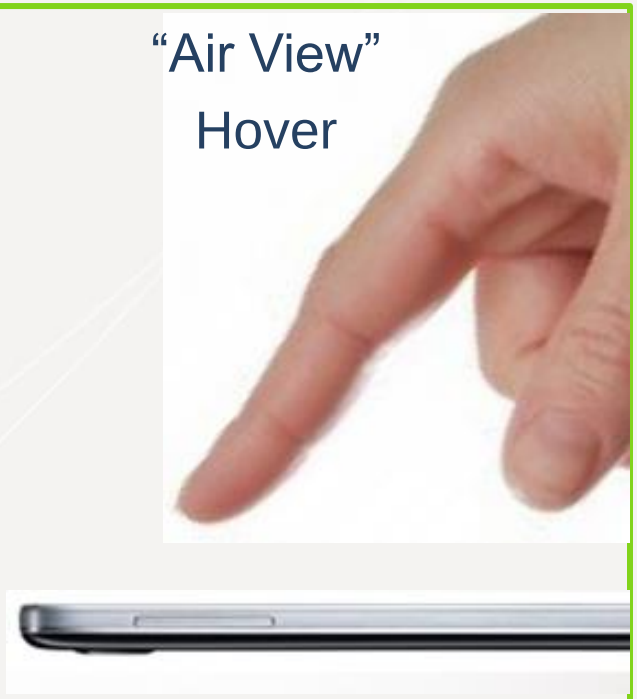
"Air Gestures"

Near Swipe

Operates within 20cm of sensor

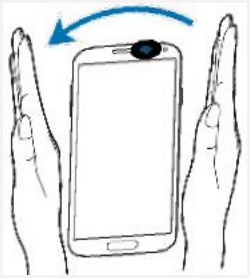
"Air View"

Hover



Operates within 2cm of screen

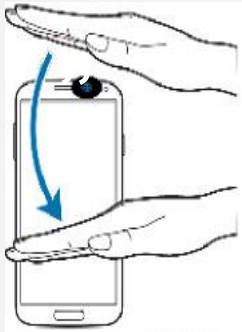
Samsung S4 – “air gestures



Air Browse

Move your hand to the left or right across the sensor to browse images, webpages, songs, or memos. time and date, and more.

Left/Right Single Swipe



Air Jump

While viewing emails or webpages, move your hand up or down across the sensor to scroll the page up or down.

Up/Down Single Swipe

Air Call-Accept

When a call comes in, move your hand to the left, and then to the right across the sensor to answer the call.

Left/Right Double Swipe

Quick Glance

When the screen is turned off, move your hand above the sensor to view notifications, missed calls, new messages, time and date, and more.

Cover

Samsung S4 – “air view” or hover

The Layered UI matches the user’s expectation of depth based on their hand position.

Enhances basic touch with “mouse over”.

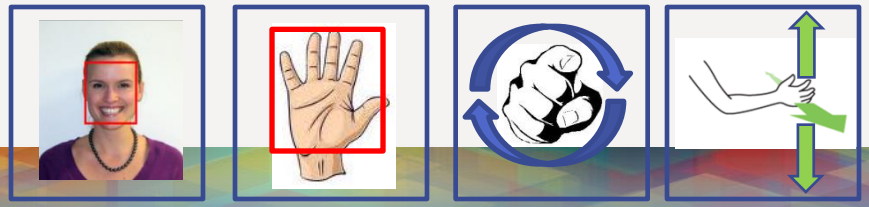
Additional hw could allow multi-level layers.

Application	Use case
Picture Browser	Show thumbnail images of images within folder by hovering over folder
Email	Show text of email by hovering over message heading
Calendar	Show individual schedule items
Notifications	Show full text of latest notification
Browser	Zoom in to show text below finger in a larger font Standard “mouse over” behavior can also be utilized

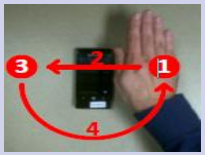
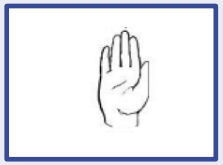
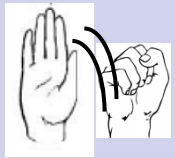
Gesture Use Cases: Tablet

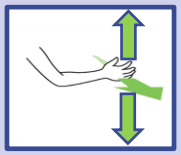
Tablets lend themselves to be in stands for a "lean back" use case where touch is awkward, and to being laid on a table where certain features do not function.

Use Case	Description
Media / Music Control	Pause / play movie Pause / play / skip song
Email / Facebook reading	Step through messages in email quickly without touching the screen.
Meeting room control	Step through PowerPoints with a hand wave…
Screen Orientation	When a tablet is on its back on a tablet, its orientation can be shifted by placing your hand above it in your desired orientation.
Multiple user identification for games and shared interactions	Use the hand orientation to determine the direction they are approaching the tablet from to identify who they are.



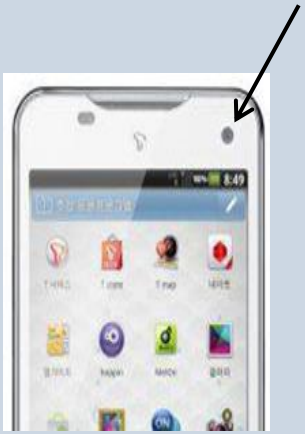
A “Standard” Gesture Alphabet

Gesture	Description
	“Near” Swipes, left / right / up / down motions across the device.
	“Pause or Mute”, holding the hand still over the device without touching it.
	“Zoom in / out”, moving your hand slowly towards or away from the device
	“Pointer Control”, gives mouse like control using the fingertip.
	“Click & Drag”, either by thrusting a finger forward or closing the hand.

Gesture	Description
	“Far swipes”, left / right / up / down motions from 30cm to 5m from the device.
	“Wave”, typically used to tell the system that you want to engage gesture mode.
	“Open Palm”, holding steady for engagement. Note: also gives hand size for zoom.
	“Circle Gesture”, moving a finger or hand in a circle to rapidly scroll or set volume
	“Hand Pose”, holding the hand still in a variety of known poses.

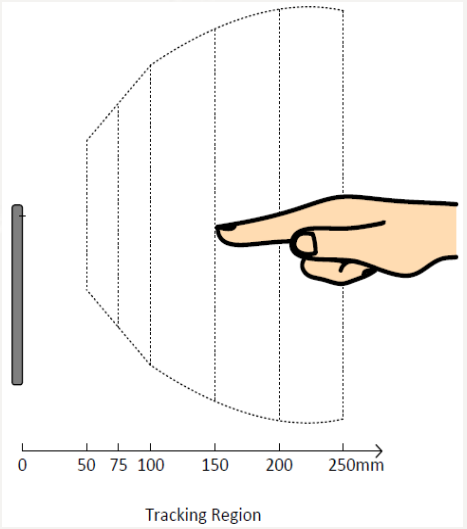
Which Gesture Hardware?

Gesture Hardware

2D Image Sensors	3D Image Sensors	Other
		
Built-in user facing camera	Structured Light Time Of Flight Hybrid Stereo/Triangulation	Ultrasound Near Field Imaging Dedicated sensors
Existing 2D sensors are low cost and widely deployed. Infrared to follow: 2014	Cost of 3D sensor solutions delays CE deployment 2013: Accessories 2014: Embedded	Low cost and low power solutions will be deployed as gesture specific hardware: Ultrasound / NFI: late 2013

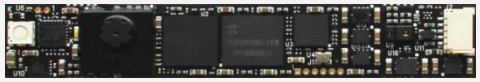
Standard 2D Sensor Vision Based Hardware

- Ubiquitous deployment of user facing 2D cameras makes this the easiest and lowest cost system to deploy.

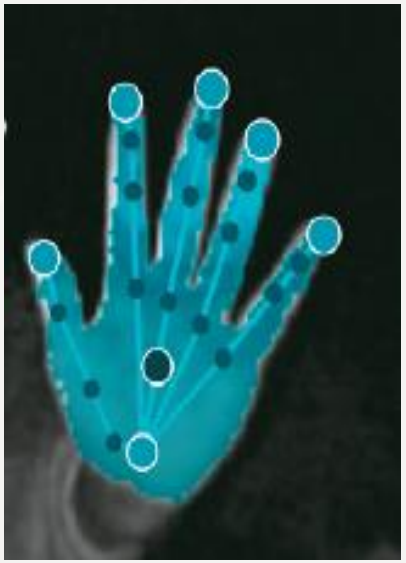


Range	Display Alignment	Low Power Mode	Swipes / Z	Finger Pointing	Hand Pose	Issues
0 to 5m	Top Alignment	Ultra low frame rate	Yes, with limited Z	Yes	Yes	Low ambient light. Complex environments. Lack of true Z.

3D Sensor Vision Based Hardware

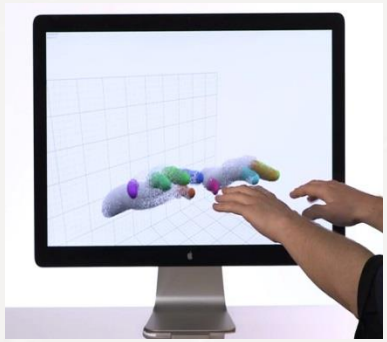


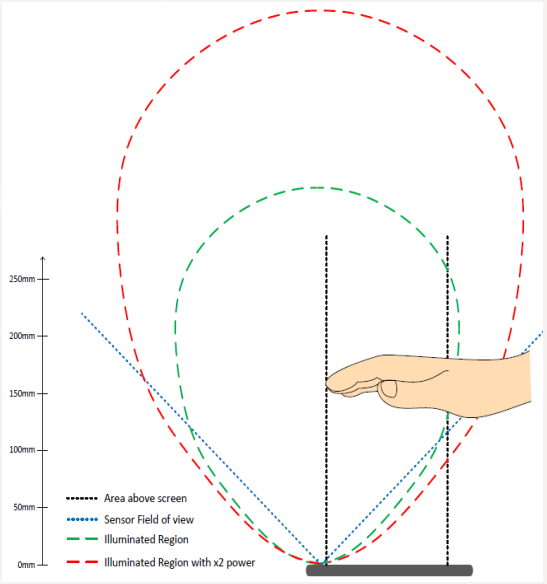
- Time of Flight sensors are being implemented in laptops for 2014 and will provide unprecedented robustness and accuracy of hand and finger tracking for gesture control
- SoftKinetic with TI and PMDtec with Infineon are the market leaders
- Power use and heat generation are quite high for mobile use in current era hardware



Range	Display Alignment	Low Power Mode	Swipes / Z	Finger Pointing	Hand Pose	Issues
0 to 500cm	Top Alignment	Not yet	Yes!	Yes!	Yes!	Power requirement is main restriction today.

3D Hybrids with IR





- Leap Motion has created a consumer market gesture accessory using two sensors, two IR emitters, and triangulation to provide 3D finger positions
- This device does not work well when pointed “at” people and hence is more suitable for laptops and desktop use than other devices

Range	Display Alignment	Low Power Mode	Swipes / Z	Finger Pointing	Hand Pose	Issues
0 to 50cm	Top Alignment	No	Yes	Yes	Yes	Orientation of device to user is main restriction.

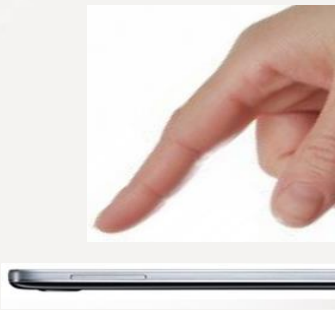
Samsung S4 – dedicated gesture sensor

Dedicated IR emitter/receiver pair
Similar to proximity sensor

Range	Display Alignment	Low Power Mode	Swipes / Z	Finger Pointing	Hand Pose	Issues
0 to 20cm	Top Alignment	Yes	Yes, no Z	No	No	Allows simple gesture only. Low fidelity. Easily fooled.

“Hover” Technology: extended projected capacitance

Samsung partnered with Synaptics to bring hover to the Synaptics ClearPad™ product line used in the S4 for AirView



Sony’s Xperia™ sola includes floating touch™ created in partnership with Cypress Technologies.

Combines self capacitance and mutual capacitance to deliver both touch and hover capability.

Range	Display Alignment	Low Power Mode	Swipes / Z	Finger Pointing	Hand Pose	Issues
0 to 2cm	Aligned	Yes	Difficult	Yes	No	Limited operating distance. Influenced by hand holding device.

Elliptic Labs - Ultrasound

ellipticlabs

Wide Field Of View

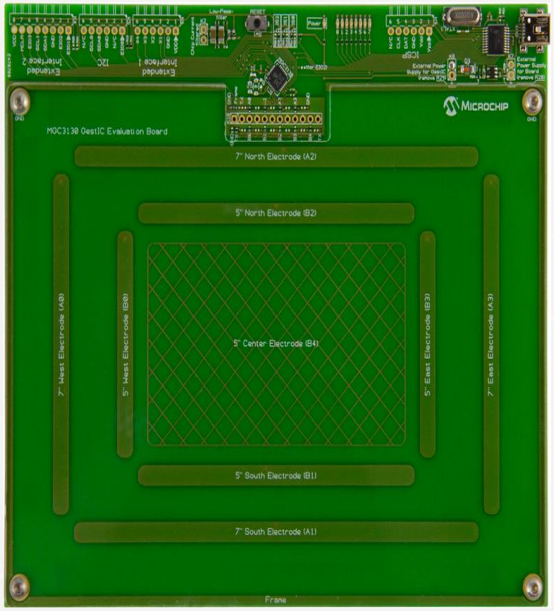
- Goes beyond the screen edges for best coverage

Distributed Sensing

- Redundant microphones and emitters enhance robustness

Range	Display Alignment	Low Power Mode	Swipes / Z	Finger Pointing	Hand Pose	Issues
0 to 30cm	Aligned + offscreen	Yes	Yes, with Z	Yes	No	Microphone placement can compromise design

GestIC from Microchip – Electric Near Field Imaging



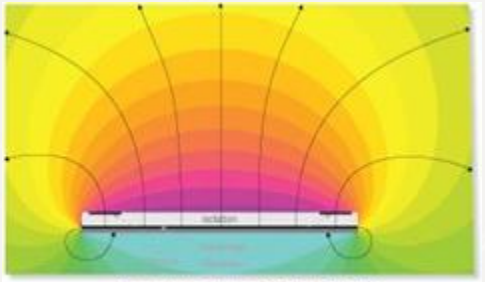


Figure 1: Undisturbed E-field distribution

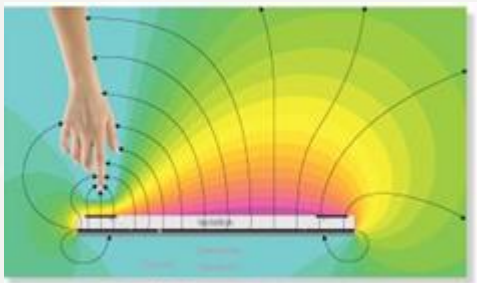
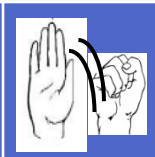
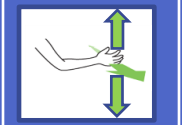
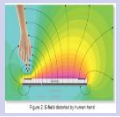


Figure 2: E-field distorted by human hand

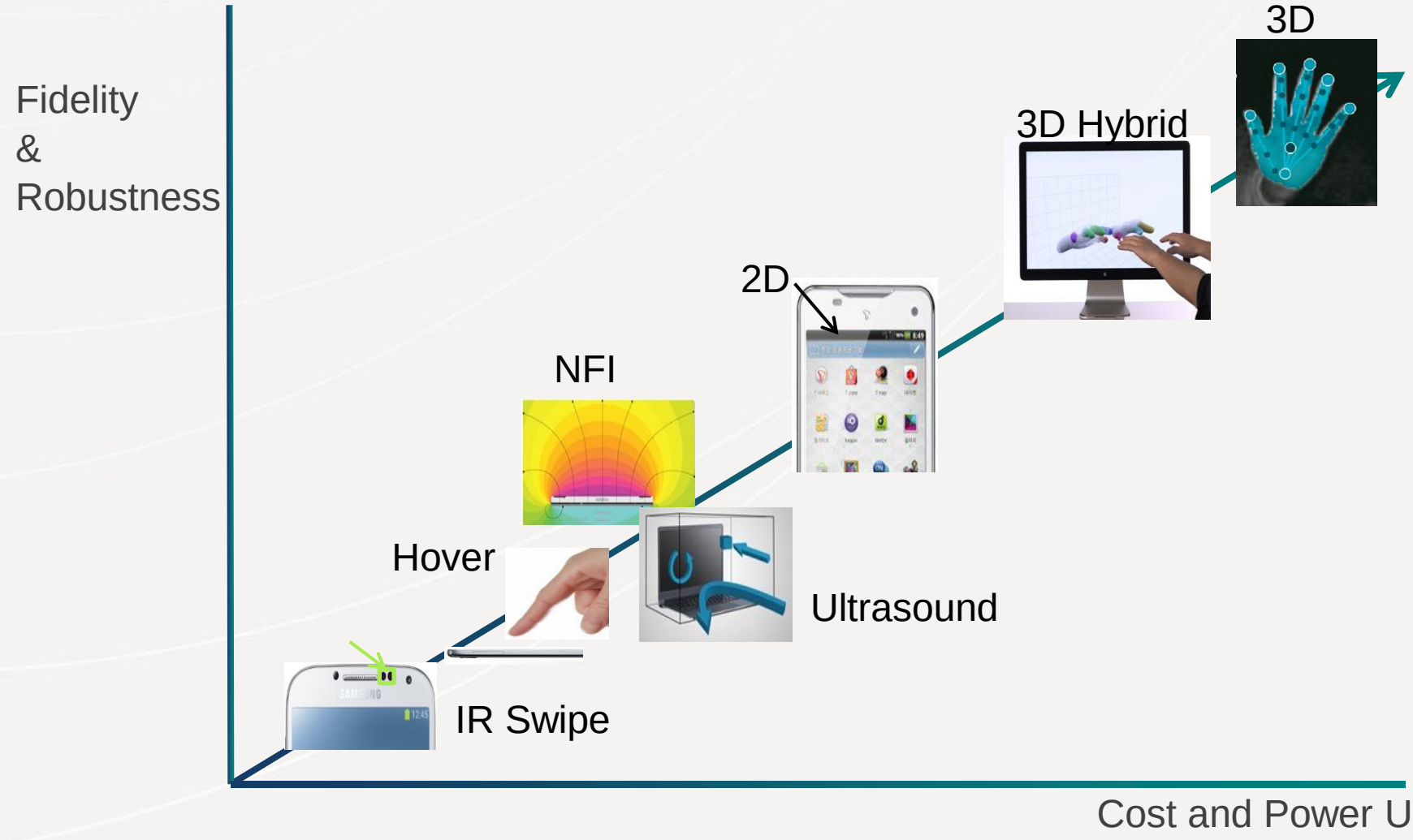
100kHz frequency emissions create a quasi static electrical near field used to sense conductive objects including the human body.

Range	Display Alignment	Low Power Mode	Swipes / Z	Finger Pointing	Hand Pose	Issues
0 to 15cm	Aligned + offscreen	Yes	Yes, with Z	Yes	No	Electrode placement can compromise design

Gesture Alphabet Hardware Capability Matrix

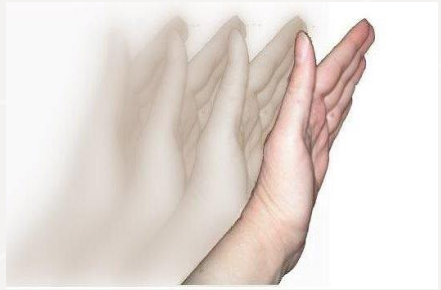
										
	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
	YES	YES	NO	NO	NO	NO	NO	NO	NO	NO
	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO
	YES	YES	YES	YES	NO	NO	NO	NO	YES	NO
	YES	YES	YES	YES	NO	NO	NO	NO	YES	NO
	YES	YES	YES	YES	YES	NO	NO	YES	YES	YES
	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Gesture Hardware Fidelity vs Cost and Power



How hard is vision based gesture recognition?

Depends on the hardware…

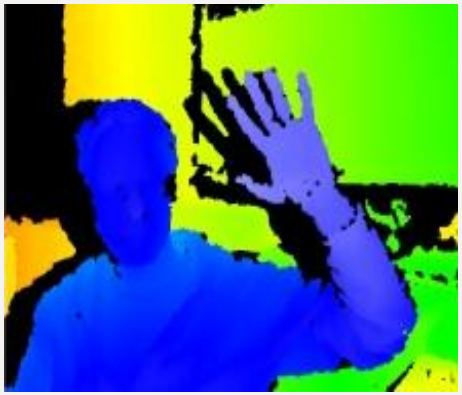


Lighting

Exposure

IR Augmentation

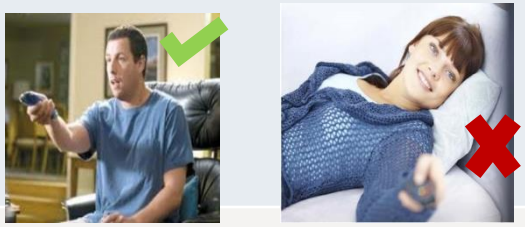
Scene Complexity



Raw Depth Map

Sample Vision Gesture Issues

Technology	Issues	Directions
Standard 2D sensors	Robust hand pose detection Low light operation High resolution tracking Fast hand tracking	Machine learning RGB+IR sensors plus illumination and Ultrasensitive sensors Use of DSP and GPU for acceleration Random forest + motion flow + …
3D TOF Sensors - Near Range	Robust hand skeletonization	Machine learning Iterative depth image rendering and comparison Self organizing maps
3D Hybrids	Partial hand occlusion Overlapping hands	Machine learning Feature identification
3D Sensors - Far Range	Hand control from “laying down on the couch” position	Machine learning



When? The State of the Gesture Market.

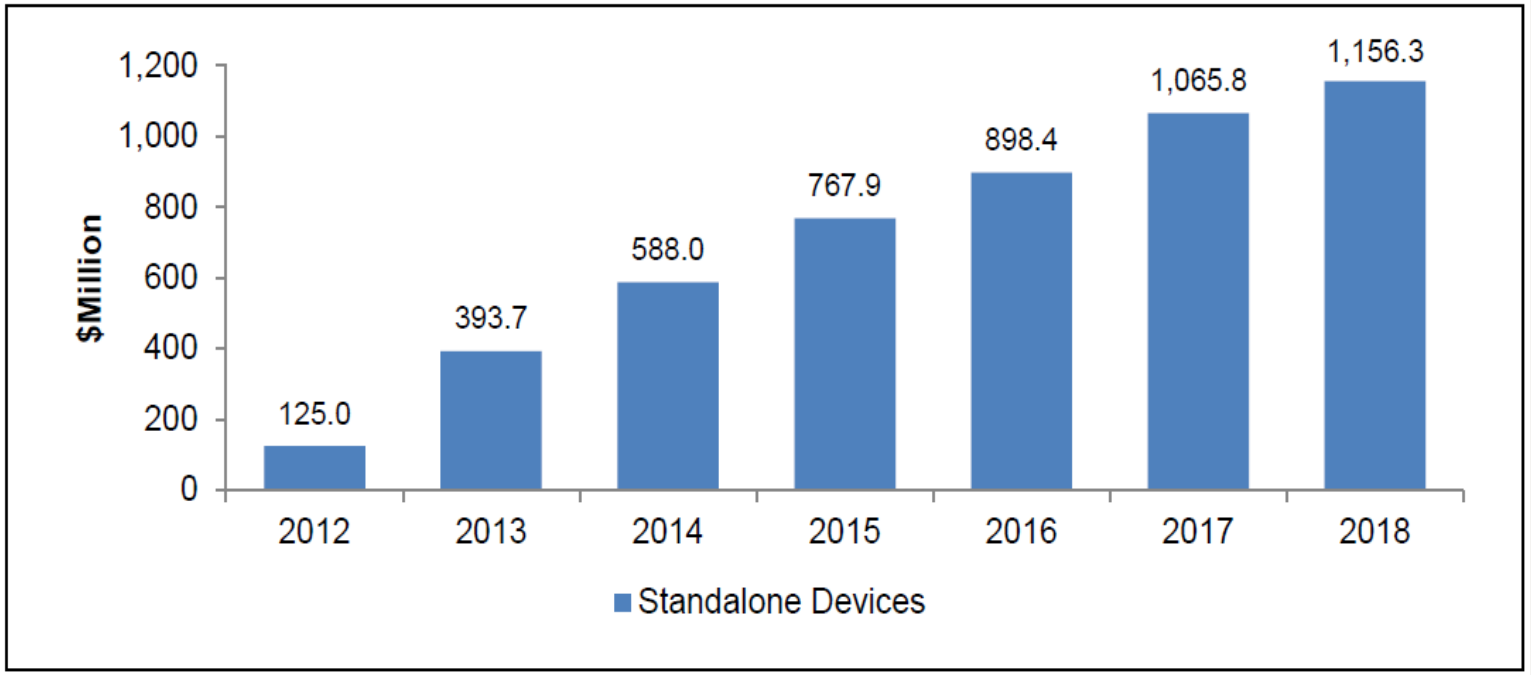
Global Gesture Recognition Market Opportunity, 2012 - 2018

Particulars	2012	2013	2014	2015	2016	2017	2018	CAGR% (2013-2018)
Revenue ($Million)	322	918	1,639	2,667	3,834	5,335	7,152	50.77
Shipments (Million Units)	10	110	309	604	985	1,471	2,049	79.22

- Global gesture recognition market revenue expected to reach $7.15B by end of 2018, increasing at a CAGR of 50.77% from 2013 till 2018.
- Market shipments expected to reach 2.04B units by 2018, growing at a CAGR of 79.22%, calculated from 2013 till 2018.
- 2012 shipments include Kinect, early mobile adopters, and Samsung TVs among others
- 2013 shipments include Samsung S4 and other mobile adopters as well as follow-on TVs and Leap Motion

Source: MarketsandMarkets Analysis, April 2013

Global Standalone Devices for Gesture Recognition



Standalone gesture recognition products from startups like Leap Motion, Thalmic labs, and others –

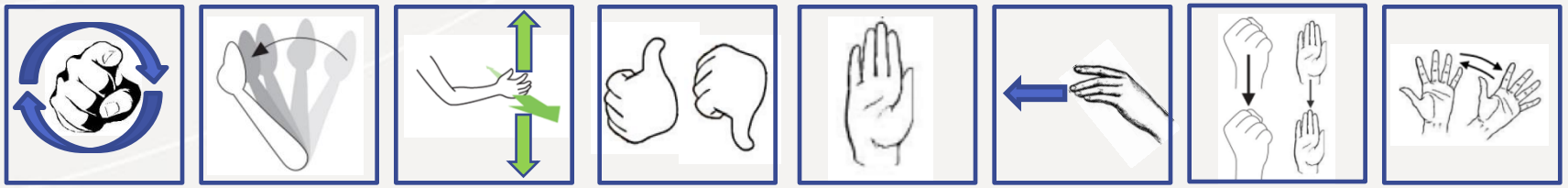
- Expected to grow from $393.70 million in 2013 to $1.15 billion by the end of 2018
- Increasing at a healthy CAGR of 24.05% from 2013 till 2018.

Source: MarketsandMarkets Analysis, April 2013

Future Use Cases: Embedded Automotive

- Voice plus additional interaction methods desired by car companies
- Identification of users provides added benefits

Use Case	Description
Infotainment System Control	Control GPS, music, bluetooth devices directly using no-look swipes, hand poses, thumb up-down, etc. Detection of passenger vs driver interaction.
Sunroof 	No-look control of the sunroof while driving. Swipe backward to open, forward to close, push up to tilt, pull down to close.
Recognition	Identity can be used for custom settings of seat position, music selection, temperature, etc.

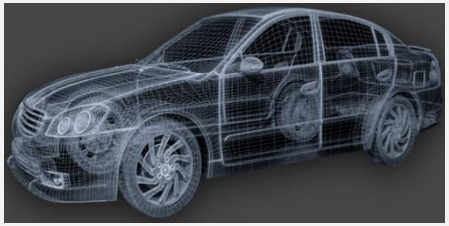


The Future: Gesture Enabled

- Intel's Perceptual Computing initiative plans to make gestures as ubiquitous as the mouse and keyboard for laptops and desktops

- Asus, Dell, HP, and Lenovo will place 3D-depth cameras right inside the screen bezel of laptops, starting in the second half of 2014

- Qualcomm is releasing embedded gesture into all of its upcoming chipsets, fueling handsets and tablets and TVs worldwide, with cars and other devices to follow



In context, a final word…

- It took “touch” over 20 years to mature to the point that the technology was robust enough to track more than one finger at a time in a consumer device
- Once “touch” matured to “multi-touch”, the UX of the iPhone emerged to leverage it and change an industry
- Gesture technology is rapidly maturing from simple swipes to robust full 3D hand skeleton tracking on consumer devices
- The utilization of high fidelity Gesture will be the basis of the next UX revolution

Thank you!

Contact: francism@qti.qualcomm.com

Resources

http://www.embedded-vision.com/platinum-members/embedded-vision-alliance/embedded-vision-training/documents/pages/gesture-interface-compelling
http://www.embedded-vision.com/platinum-members/embedded-vision-alliance/embedded-vision-training/videos/pages/april-2013-summit-omek-interactive
http://www.embedded-vision.com/industry-analysis/video-interviews-demos/2013/01/27/january-2013-consumer-electronics-show-product-d
http://www.embedded-vision.com/industry-analysis/video-interviews-demos/2013/01/24/january-2013-consumer-electronics-show-product-d