

Microsoft® Research

# Kinect™ for Windows® SDK beta



Microsoft® Research

Kinect™ for Windows® SDK beta

# Community Update and Next Steps

**Stewart Tansley, PhD**

Senior Research Program Manager  
& Product Manager (acting)



*Special Guest:*

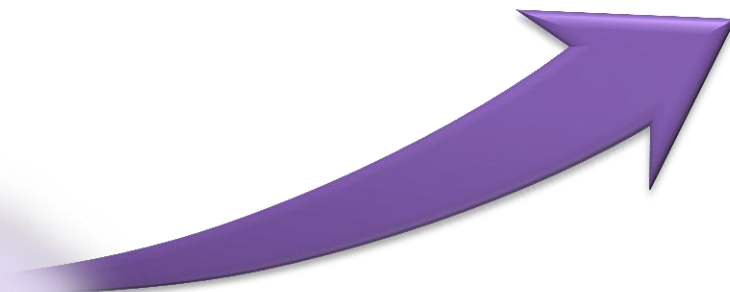
**Anoop Gupta, PhD**

Distinguished Scientist



**Project  
Natal**

**KINECT™**  
for  **XBOX 360.**



# Origins: Project Natal

- Named after the Brazilian city, meaning “relating to birth” (Alex Kipman)
- *“The birth of the next generation of home entertainment”*
- Not just the device. The sensor provides eyes and ears,  
→ **but it needs a brain**
- Raw data from that sensor is just *“a whole bunch of noise that someone needs to take and turn into signal”*  
→ **that is what our software does: find the signal**



Source: Wikipedia

# Natal → Kinect

- You know this: decades of research in computer vision...
- Xbox called up MSR in September 2008
- First announced June 1, 2009 at E3
- Launched in North America on November 4, 2010 (then EU, Japan, Australia...)
- 10 million sold (as of March 9, 2011)  
*Guinness world record: fastest selling consumer electronics device of*



# The Problem

- Find the people in the scene, ignore background
- Find their limbs and joints, which person is which
- Find and track their gestures
- Map the gestures to meaning and commands
  
- *Also, recognize faces*
- *Also, recognize voices and commands*

**PS. And play the game!**

# Software Magic!

- Machine Learning
- Effectively...
  - Evaluate trillions of the possible body configurations of 32 body (skeletal) segments
  - Every video frame
  - 30 times a second
  - On <10% of the CPU



# Behind the Magic

- Decades of computer vision research between industry and academia, including our own at Microsoft Research and Xbox
- State of the art in human body tracking in 2007 had the ability to track a wide range of motion  
→ **but with limited agility and not in real-time**
- Xbox's requirement:  
***all motions, all agilities, 10x real-time, for multiple bodies!***
- But they did have a low-cost 3D camera...



# Vision Algorithm (Paper)



- CVPR 2011 Best Paper:

## ***Real-Time Human Pose Recognition in Parts from a Single Depth Image***

Jamie Shotton, Andrew Fitzgibbon, Mat Cook, Toby Sharp,  
Mark Finocchio, Richard Moore, Alex Kipman, Andrew Blake

<http://research.microsoft.com/apps/pubs/default.aspx?id=145347>



[Paper](#)



[Supplementary](#)



[Video](#)

<http://cvpr2011.org>

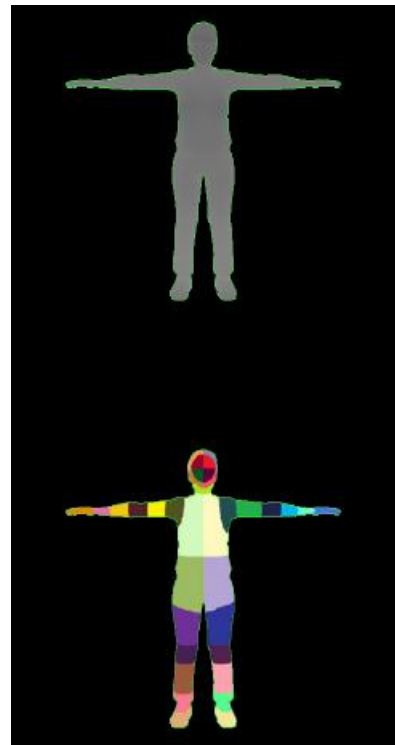
# Vision Algorithm (Summary)

- Quickly and accurately predict 3D positions of body joints.
- From a single depth image, using no temporal information.
- **Object recognition approach.**
- Intermediate body parts representation that maps the difficult pose estimation problem into a simpler **per-pixel classification problem**.
- Large and highly varied training dataset allows the classifier to estimate body parts invariant to pose, body shape, clothing, etc.
- Generate confidence-scored 3D proposals of several body joints by re-projecting the classification result and finding local modes.
- System runs at 200 frames per second on consumer hardware.
- Evaluation shows high accuracy on both synthetic and real test sets.
- State of the art accuracy in comparison with related work and improved generalization over exact whole-skeleton nearest neighbor matching.



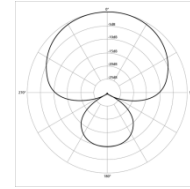
# In Practice

- **Collect training data – thousands of visits to global households, filming real users, the Hollywood motion capture studio generated billions of images**
- **Apply state-of-the-art object recognition research**
- **Apply state-of-the-art real-time semantic segmentation**
- **Build a training set – classify each pixel's probability of being in any of 32 body segments, determine probabilistic cluster of body configurations consistent with those, present the most probable**
- **Millions of training images → Millions of classifier parameters**
- **Hard to parallelize → New algorithm for distributed decision-tree training**
- **Major use of DryadLINQ (large-scale distributed cluster computing)**



# Don't Forget the Audio!

- 4 supercardioid microphone array in Kinect
- See: 1hr MIX presentation by Ivan Tashev
- <http://channel9.msdn.com/events/MIX/MIX11/RES01>
- “The talk will cover the overall architecture and algorithmic building blocks of the Kinect device, **especially the audio pipeline**. We will present the opportunities it opens for building better human-machine interfaces, new user experiences, and other potential applications. No specialized signal processing background is required. **The presenter is the creator of most of the audio algorithms in the Kinect pipeline.**”



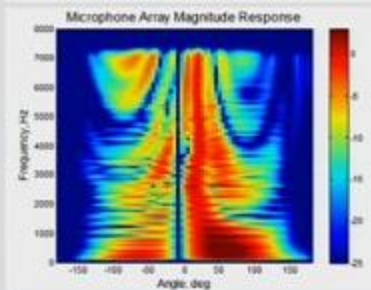
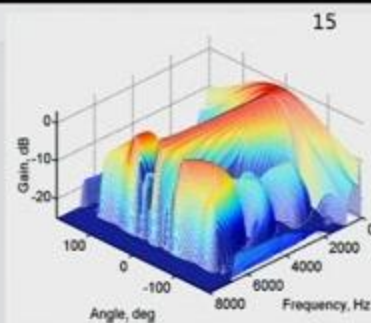
Source: Wikipedia





## Adaptive beamforming

- 1 On the fly computation of the weights
- 2 Higher CPU requirements
  - Does null steering
- 3 MVDR beamformer
  - $$\mathbf{w}_{MVDR}(f) = \frac{\mathbf{D}_e^H(f)\Phi_{NN}^{-1}(f)}{\mathbf{D}_e^H(f)\Phi_{NN}^{-1}(f)\mathbf{D}_e(f)}$$
- 4 Nulls can be enforced if known
- 5 Two microphone array demos



# Preparing for a Windows SDK

- SDK conversations through 2010 (personal: ~1yr)
- Retail entertainment launch, November 2010
- SDK statement of intent, February 21, 2011
  - Don Mattrick & Craig Mundie
  - "Available Spring 2011"
  - "Non-commercial use" – research/academic, enthusiasts
  - "Free download"
- SDK website "Coming Spring 2011", MIX & Paris, 4/2011  
→ **Launch, June 16, 2011 – [http://\[rmc\]/kinectsdk](http://[rmc]/kinectsdk)**




Take control.



## Kinect for Windows SDK beta

Coming later this spring, the Kinect for Windows SDK is a programming toolkit that will enable researchers and enthusiasts easy access to the capabilities offered by the Microsoft Kinect device connected to computers running Microsoft Windows 7.

Get release announcements, updates, news, and more.

 [Subscribe](#)

With this SDK, you'll be able to take advantage of:

- The latest advances in audio processing, which include a four-element microphone array with sophisticated acoustic noise and echo cancellation for crystal clear audio.
- Sound source localization for beamforming, which enables the determination of a sound's spatial location, enhancing reliability when integrated with the Microsoft speech recognition API.
- Depth data, which provides the distance of an object from the Kinect camera, as well as the raw audio and image data, which together open up opportunities for creating richer natural user interface experiences.
- Highly performant and robust skeletal tracking capabilities for determining the body positions of one or two persons moving within the Kinect field of view.
- Documentation for the APIs and a description of the SDK architecture.
- Sample code that demonstrates how to use the functionality in the SDK.

This SDK is intended for non-commercial use to enable experimentation in the world of natural user interface experiences, with new state-of-the-art features planned for future releases that will continue to provide new ways to experiment.

Be part of the movement.



Kinect™ for Windows® SDK beta

[Download »](#)

Download  
the SDK



Learn  
about the features



Watch  
the launch event





# What's in the SDK?

- **Raw sensor streams**
  - Access to raw data streams from the depth sensor, color camera sensor, and four-element microphone array enables developers to build upon the low-level streams that are generated by the Kinect sensor.
- **Skeletal tracking**
  - The capability to track the skeleton image of one or two people moving within the Kinect field of view make it easy to create gesture-driven applications.
- **Advanced audio capabilities**
  - Audio processing capabilities include sophisticated acoustic noise suppression and echo cancellation, beam formation to identify the current sound source, and integration with the Windows speech recognition API.
- **Sample code and documentation**
  - The SDK includes more than 100 pages of technical documentation. In addition to built-in help files, the documentation includes detailed walkthroughs for most samples provided with the SDK.
- **Easy installation**
  - The SDK installs quickly, requires no complex configuration, and the complete installer size is less than 100 MB. Developers can get up and running in just a few minutes with a standard standalone Kinect sensor unit (widely available at retail outlets).
- Designed for non-commercial purposes; a commercial version is expected later.
- Windows 7 – C++, C#, or Visual Basic in Microsoft Visual Studio 2010.





Channel 9

BROWSE

FORUMS

CODING4FUN

EVENTS

Search this site



Be part of the movement.



Microsoft Research  
Kinect for Windows SDK beta

Launch Event | June 16

Events | Kinect for Windows SDK (Beta)

## Kinect for Windows SDK Beta Launch

Subscribe to this event

Redmond, WA | June 16, 2011

Tweet 32

Share 12

Speakers at Kinect for Windows SDK Beta Launch

Channel 9 and Microsoft Research are excited to announce the Kinect for Windows SDK Beta [available to download](#)

To mark the launch we hosted a 24 hour CodeCamp on the Microsoft campus in Redmond, Washington. Fueled by coffee and pizza, 30 lucky developers were given a Kinect, the SDK Beta bits and 24 hours to create something awesome.

On June 16th 2011 we broadcasted live from the Channel 9 studios on campus demonstrating the SDK, taking your questions and showing off the projects of our CodeCamp attendees. This page is the place to find all the on-demand recording of those sessions, posted as soon as they are available.

### Keynote



#### Kinect for Windows SDK Beta Launch Announcement with Anoop Gupta & Jeff Sanquist

Jeff Sanquist, Anoop Gupta | Kinect, Ch9Live, Microsoft Research | 100  
C9L001 | June 16, 2011 from 9:30AM to 10:00AM

★★★★★ 0

### Event Homepage

Get the SDK.



### Filter Sessions

Term:

Tag:

- Ch9Live
- Kinect
- Microsoft Research



# CODING4FUN

Subscribe to this blog

15 Posts

Tweet

19

Share

19

Home

Blog

Projects

Articles

Show

Kinect

About

Recent

Most Viewed

Top Rated

@channel9



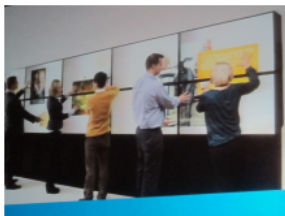
## Intro to the Kinect SDK–Drawing Joints in XNA

Coding4Fun Kinect Projects

8 hours ago 0



When getting up to speed on something as new as the Kinect for Windows SDK, multiple point of views can be a life saver. What is explained in one place might have you scratching your head, whereas the same thing explained by someone else can trigger an "ah ha!" moment. Today's project takes an...



## "I have the Minority Report UI.."

Coding4Fun Kinect Projects

1 day ago 4



Today's project is the start of doing what we all want to see done, putting a NUI (Natural User Interface) on Windows. You know this is coming, that it's just a matter of time, that touch support for desktop/notebooks is just a stopgap until we get real NUI/gesture support. Do you really want to...

Be part of the movement.



Microsoft® Research

Kinect™ for Windows® SDK beta

[research.microsoft.com/kinectsdk](http://research.microsoft.com/kinectsdk)

# Community Update

- Launch
  - Codecamp 24hr pre-launch event
- First Month
  - Seattle (UW)
  - UK
  - France
  - Australia
  - New York (Imagine Cup)
  - ...



# CodeCamp Showcase

36 campers

12 teams

24 hours!



## [Kinect for Windows SDK Beta Launch CodeCamp Demos #01](#)

June 16, 2011 from 10:00AM to 10:15AM

## [Kinect for Windows SDK Beta Launch CodeCamp Demos #02](#)

June 16, 2011 from 11:15AM to 11:45AM

## [Kinect for Windows SDK Beta Launch CodeCamp Demos #03](#)

June 16, 2011 from 1:30PM to 1:45PM

[http://www.flickr.com/photos/msr\\_redmond/sets/72157626971787454/show/](http://www.flickr.com/photos/msr_redmond/sets/72157626971787454/show/)

### **Universities:**

Seattle University  
Oregon State University  
Lewis & Clark College  
University of Victoria  
Simon Fraser University  
Washington State University  
UC Santa Cruz  
University of British Columbia (UBC)  
University of Washington  
University of Maryland  
Georgia Tech  
McGill University  
UCLA, MIT

### **Businesses**

Cynergy Systems  
IdentityMine  
InfoStrat Advanced Technology Group  
Developer Express  
Wire Stone  
Pixel Lab, ZAAZ, KEXP

Microsoft® Research

Kinect™ for Windows® SDK beta



# Next Steps

- Contests (proposed)
  - Undergraduate (Imagine Cup)
  - Research
  - Open (all-comers)
- Training Workshops
  - Locations in planning
- Research Workshop(s)
  - Later;  
Let's do some (more) work first!





# Kinect SDK at Faculty Summit 2011

- **Monday**
  - 13:30-15:00 **Community Update & Next Steps**: you are here 😊
  - 16:30-19:30 **DemoFest: Kinect SDK Showcase**
- **Tuesday**
  - 9:00-10:30 **Tutorial #1: Introduction and Overview**
  - 11:00-12:30 **Tutorial #2: Deep Dive**
  - 13:30-15:00 Panel: NUI—The Road Ahead  
Mark Bolas, University of Southern California; Justine Cassell, Carnegie Mellon University;  
Mary Czerwinski, MSR; Daniel Wigdor, University of Toronto + Kristin Tolle, MSR
  - 16:00-17:00 Plenary: Vision-based NUI  
Rick Szeliski, MSR

Be part of the movement.



Microsoft® Research

Kinect™ for Windows® SDK beta

[research.microsoft.com/kinectsdk](http://research.microsoft.com/kinectsdk)

# *Microsoft*<sup>®</sup>

© 2011 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation.  
MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.