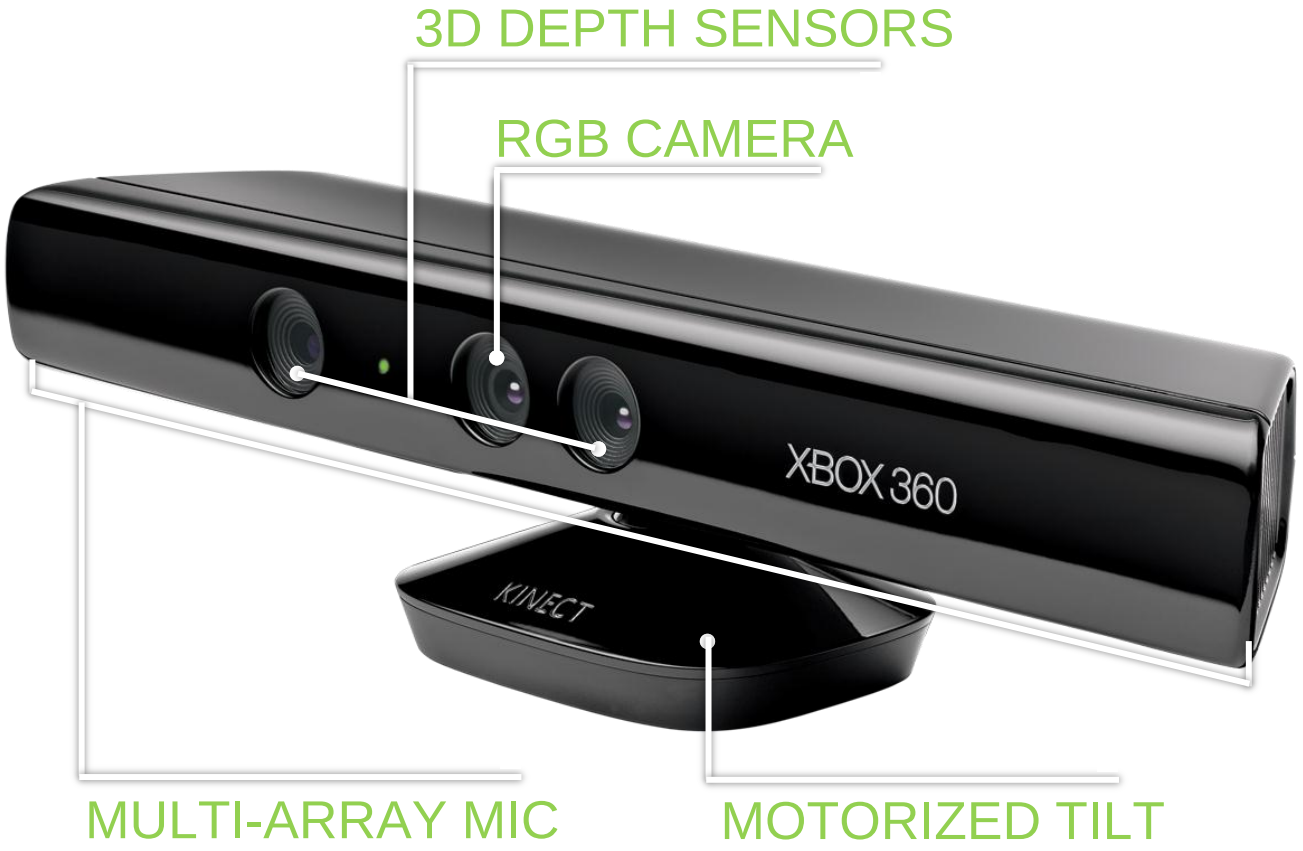
Microsoft® Research
Kinect™ for Windows® SDK beta

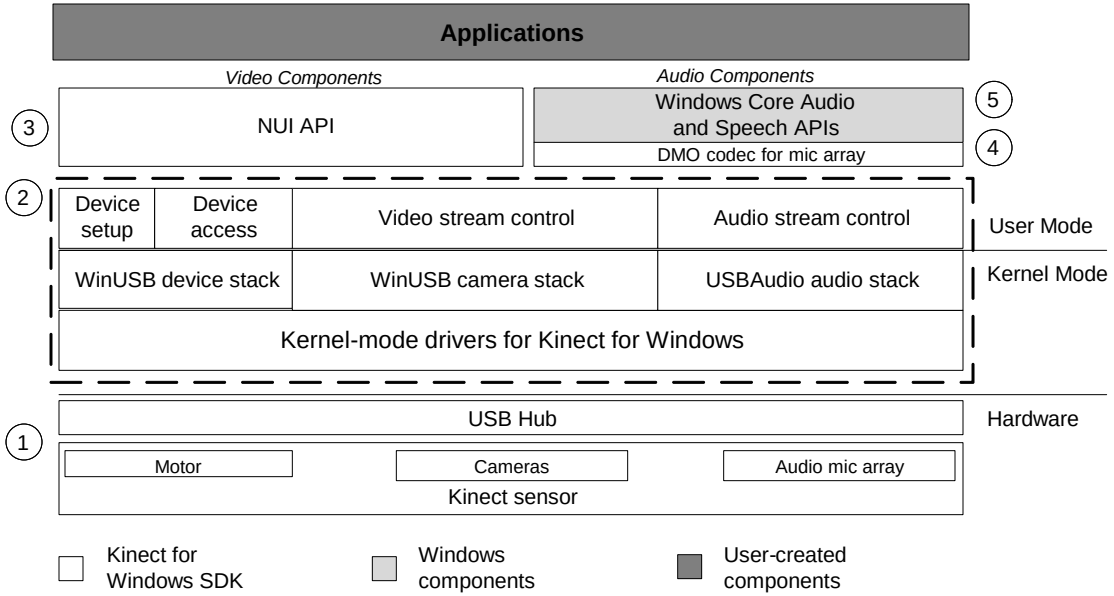
Programming with the Kinect for Windows SDK

What we'll cover

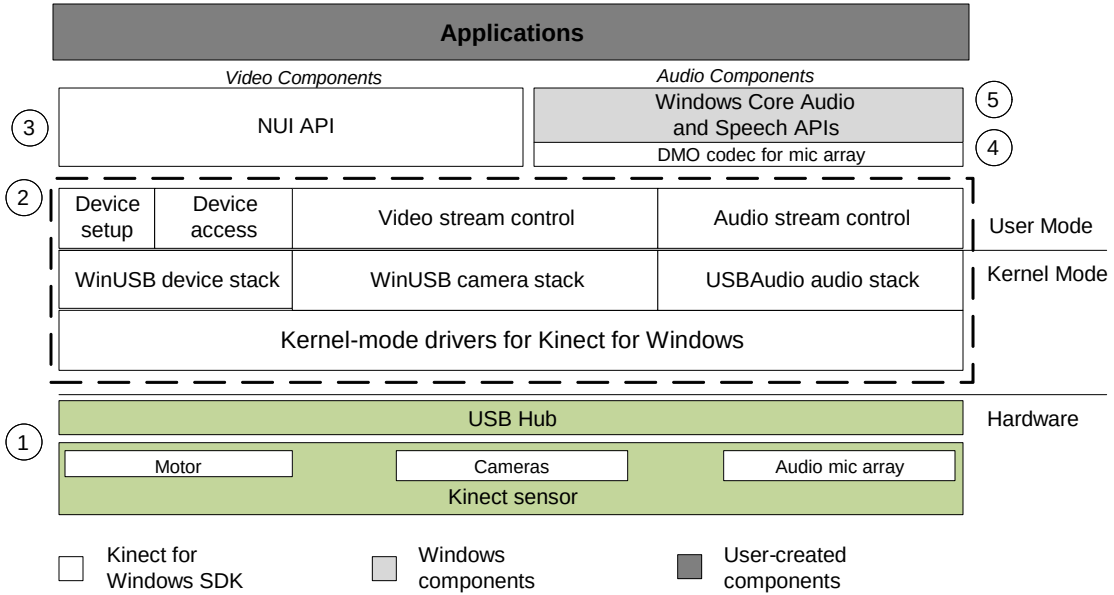
- Kinect Sensor
- Using Cameras
- Understanding Depth Data
- Skeletal Tracking
- Audio


3D DEPTH SENSORS
RGB CAMERA
XBOX 360
KINECT
MULTI-ARRAY MIC
MOTORIZED TILT

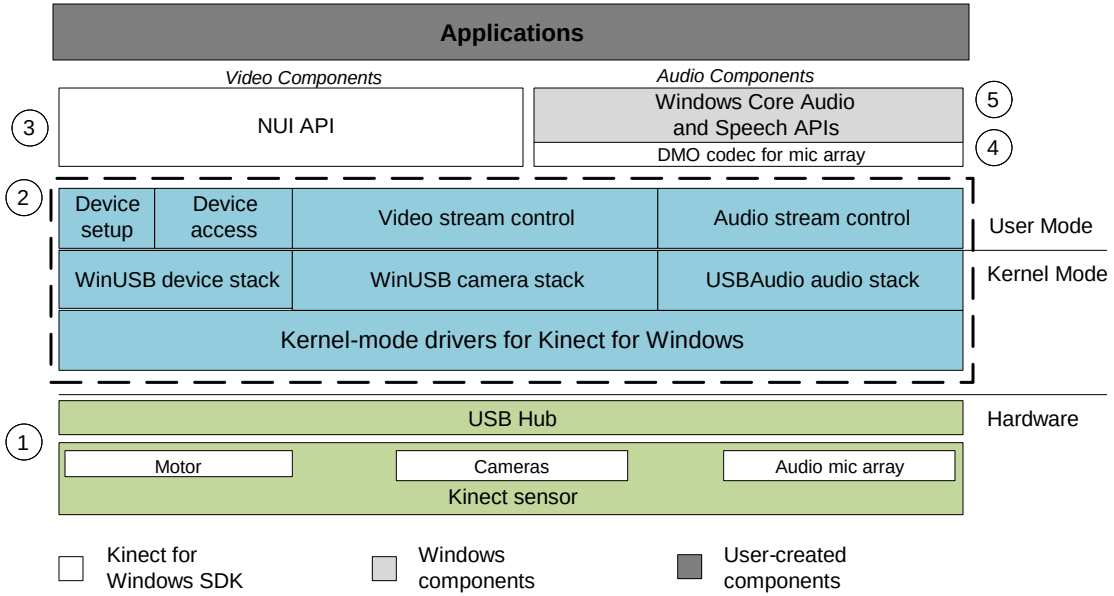
SDK Architecture



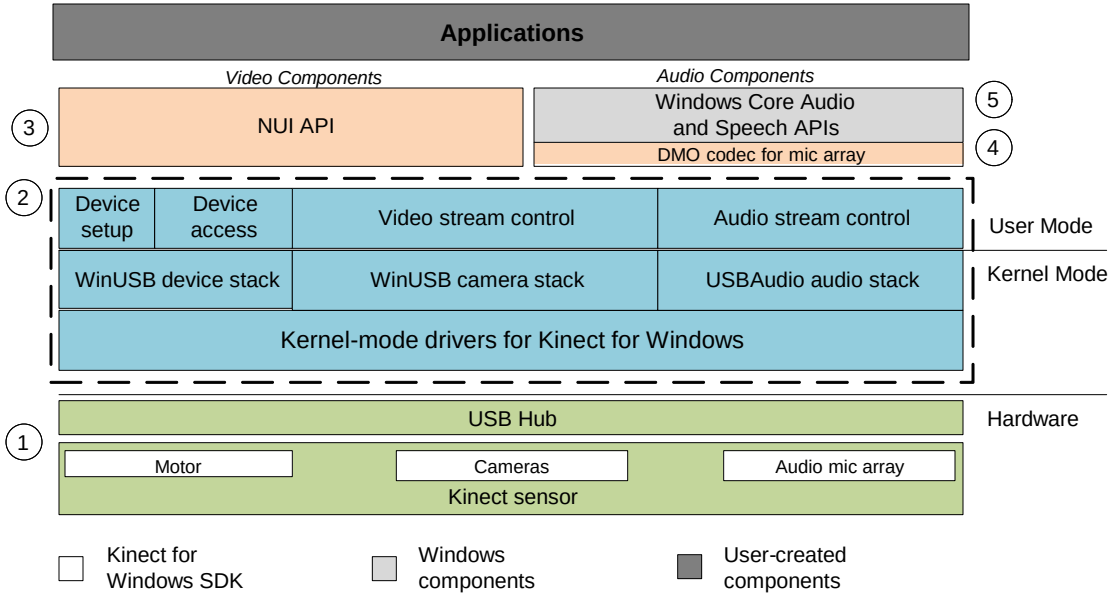
SDK Architecture



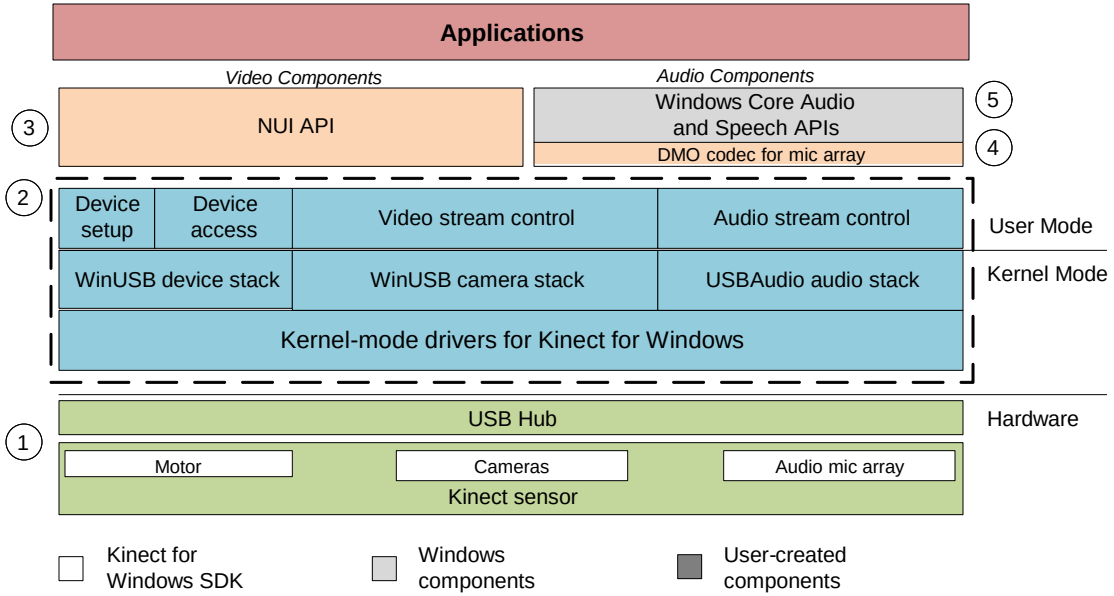
SDK Architecture



SDK Architecture



SDK Architecture



Using Cameras

Demos

Understanding Depth Data

- ImageFrame.Image.Bits
- Array of bytes - `public byte[] Bits;`
- Array
  - Starts at top left of image
  - Moves left to right, then top to bottom
  - Represents distance for pixel in millimeters

Calculating Distance

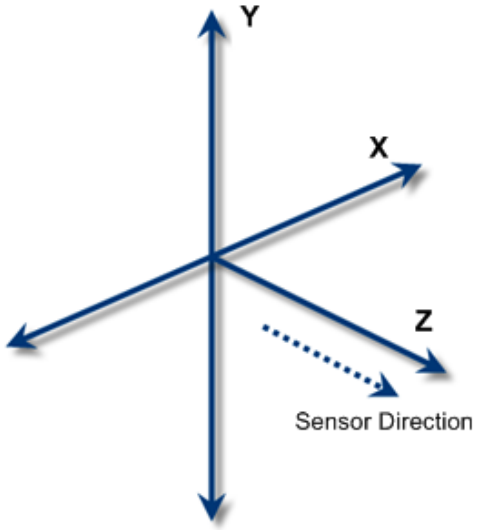
- 2 bytes per pixel (16 bits)
- Depth – Distance per pixel
  - Bitshift **second byte by 8**
  - Distance (0,0) = `(int)(Bits[0] | Bits[1] << 8);`
  - VB `(int)(CInt(Bits(0)) Or CInt(Bits(1)) << 8);`
- DepthAndPlayer Index – Includes Player index
  - Bitshift by **3 first byte** (player index), **5 second byte**
  - Distance (0,0) =`(int)(Bits[0] >> 3 | Bits[1] << 5);`
  - VB:`(int)(CInt(Bits(0)) >> 3 Or CInt(Bits(1)) << 5);`

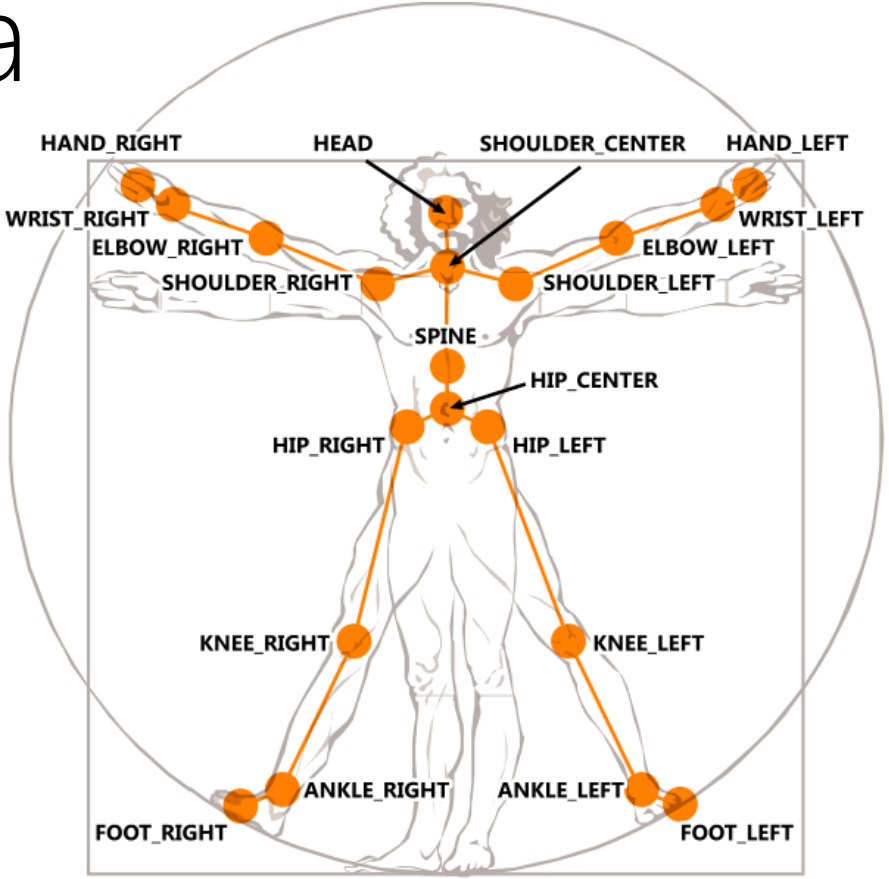
Depth Reference

- Distance Range: 850 mm to 4000 mm range
- Depth value 0 means unknown
  - Shadows, low reflectivity, and high reflectivity among the few reasons
- Player Index
  - 0 – No player
  - 1 – Skeleton 0
  - 2 – Skeleton 1
  - …

Demos

Skeleton Data

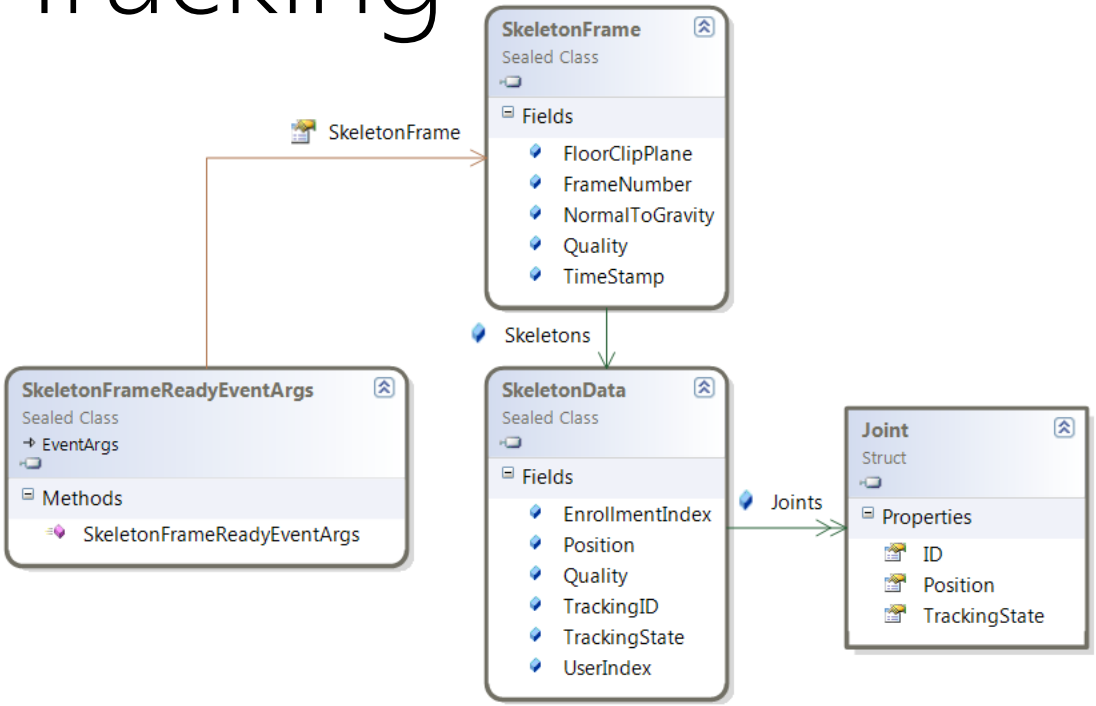




Joints

- Maximum two players tracked at once
  - Six player proposals
- Each player with set of <x, y, z> joints in meters
- Each joint has associated state
  - Tracked, Not tracked, or Inferred
- Inferred - Occluded, clipped, or low confidence joints
- Not Tracked - Rare, but your code must check for this state

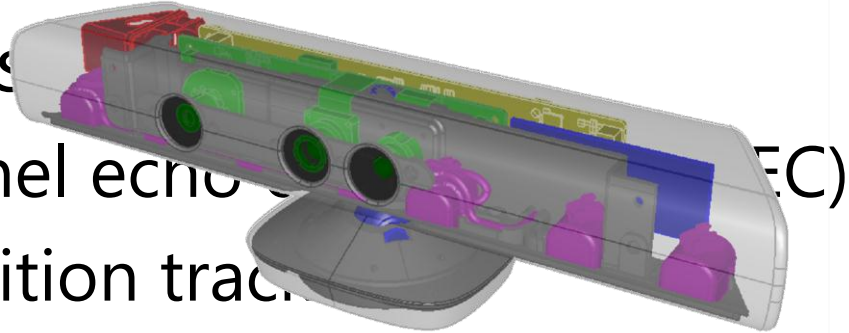
Skeletal Tracking



Demos

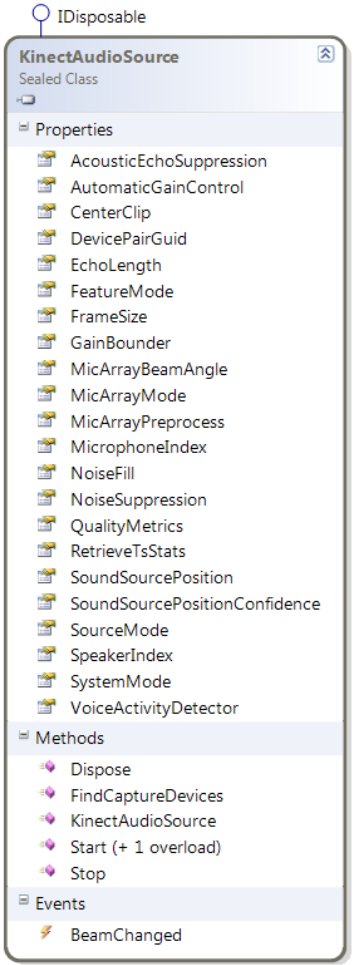
Audio Going Inside of Kinect

- Four microphone array with hardware-based audio proces
  - Multichannel echo [illegible]EC)
  - Sound position trac[illegible]
  - Other digital signal processing (noise suppression and reduction)



Audio

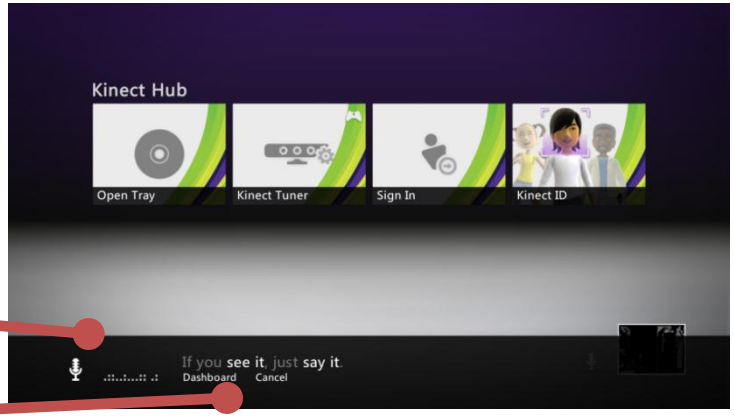
- Kinect as a microphone
- Kinect for Speech Recognition



Multi-modal Feedback

- Engagement model
  - Mic indicator for speech-enabled menus
  - Keyword to engage

- Feedback and confirmation (both passive and active)



Speech Recognition

- Kinect Grammar available to download
- Grammar – What we are listening for
  - Code – GrammarBuilder, Choices
  - Speech Recognition Grammar Specification (SRGS)
    - C:\Program Files (x86)\Microsoft Speech Platform SDK\Samples\Sample Grammars\

Grammar

```
<!-- Confirmation_YesNo._value: string
["Yes", "No"] -->
<rule id="Confirmation_YesNo"
scope="public">
  <example> yes </example>
  <example> no </example>
  <one-of>
    <item>
      <ruleref uri="#Confirmation_Yes" />
    </item>
    <item>
      <ruleref uri="#Confirmation_No" />
    </item>
  </one-of>
  <tag> out = rules.latest() </tag>
</rule>
</rule>
```

```
<!-- Confirmation_Yes._value: string ["Yes"]
-->
<rule id="Confirmation_Yes" scope="public">
  <example> yes </example>
  <example> yes please </example>
  <one-of>
    <item> yes </item>
    <item> yeah </item>
    <item> yep </item>
    <item> ok </item>
  </one-of>
  <item repeat="0-1"> please </item>
  <tag> out._value = "Yes";</tag>
```

Demos

Resources

- Kinect Programming Walkthroughs
  - http://research.microsoft.com/kinectsdk/
- Coding4Fun Kinect Toolkit
  - http://c4fkinect.codeplex.com
- Kinect SDK Quickstarts
  - http://channel9.msdn.com/series/KinectSDKQuickstarts

© 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.