

# Documentación Kinect



Jennifer Cruz Muñoz      Jose Daniel Ku Ek  
Israel Navarrete Alvarado

14 de agosto de 2012

# Índice

<b>1. Descripción del Dispositivo</b>	<b>3</b>
1.1. Capacidades técnicas. . . . .	3
<b>2. Descripción de las API's.</b>	<b>4</b>
2.1. OpenNi . . . . .	4
2.1.1. Middleware (NITE). . . . .	4
2.2. Open Kinect . . . . .	4
2.2.1. Libfreenect. . . . .	5
2.2.2. OpenKinect Analysis Library. . . . .	5
2.3. Free Kinect . . . . .	6
2.4. Kinect SDK . . . . .	6
2.4.1. Tabla comparativa de versiones. . . . .	6
2.4.2. XNA. . . . .	6
2.4.3. Microsoft Speech Platform Runtime. . . . .	7
2.4.4. Kinect Speech Language Pack. . . . .	8
2.4.5. Librería Kinect . . . . .	8
2.5. Comparativa entre APIS. . . . .	14

# 1. Descripción del Dispositivo

Kinect empezó como una propuesta de Microsoft en Junio del 2009, al principio fue conocido con el nombre "Project Natal", fue lanzado en noviembre del año siguiente teniendo el nombre con el cual lo conocemos actualmente. El dispositivo provee una NUI (Natural User Interface) que hace posible la interacción intuitiva sin ningún dispositivo intermedio, como había sido antes el control remoto. El dispositivo Kinect hace posible que una computadora  sobre sentido sobre el entorno que la rodea, lo logra usando una cámara, un  micrófono direccional y un sensor de profundidad. Gracias a este dispositivo se puede mejorar el modo de interacción humano - maquina, permitiendo que aplicaciones reaccionen a un gesto o una seña hecha por el usuario.

## 1.1. Capacidades técnicas.

Sensor.

- Lentes sensibles al color y la profundidad.
- Micrófono de voz.
- Motor para ajustar la inclinación del sensor.
- Compatible con Windows y consolas Xbox existentes.

Campo de visión.

- Campo de visión horizontal: 57 grados.
- Campo de visión vertical: 43 grados.
- Rango de inclinación física: +/- 27 grados.
- Rango de sensor de profundidad: 1.2 a 3.5 metros.

Flujo de datos.

- 320x240 profundidad de 16 bits, 30 cuadros por segundo.
- 640x480 color de 32 bits, 30 cuadros por segundo.
- 16 bits de audio, 16 KHz.

Sistema de rastreo del esqueleto.

- Rastrea hasta 6 personas, con solo 2 jugadores activos.
- Rastrea hasta 20 articulaciones por jugador activo.

- Posee la habilidad de reconocer jugadores activos con sus avatares de LIVE.

Sistema de audio.

- Party chat y chat de voz.
- Sistema de cancelación de audio que mejora la recepción de voz.
- Reconocimiento de voz.

## 2. Descripción de las API's.

### 2.1. OpenNi

Es un framework multi-language, cross-platform que define un API para desarrollo de aplicaciones, esta compuesta por un grupo de interfaces que permiten el acceso a los sensores de audio y visión, OpenNi está bajo la licencia libre, el responsable de OpenNi es PrimeSense una empresa Israelí que desarrollo la tecnología en el kinect.

#### 2.1.1. Middleware (NITE).

Para el procesamiento de la datos ~~del utiliza~~ Middleware llamado NITE, este no esta bajo licencia libre y provee las siguientes características:

- Análisis de cuerpo completo.
- Análisis de mano
- Detección de gestos
- Análisis de escena


### 2.2. Open Kinect

También conocida como free kinect es un librería bajo licencia libre cuya función es acceder a Microsoft Kinect USB camera, se comenzó desarrollarla en noviembre del 2011; esta librería, fue creada por OpenKinect que es una comunidad de programadores que según su pagina <sup>1</sup> consta con mas de 2000 miembros. Esta libreria actualmente tiene soporte para:

- RGB y sensor de profundidad.
- Motores. ?
- Acelerómetro.

<sup>1</sup>•<http://www.openkinect.org>

- ~~Led.~~ ?

- Se esta trabajando de en el acceso a audio. 

### 2.2.1. Libfreenect.

Esta librería de bajo nivel incluye todo el código necesario para activar, inicializar y establecer la comunicación de datos con el hardware kinect, incluye una API cross-platform que funciona con Windows, Linux y OS X. El API tiene adaptaciones y ~~extensiones~~ para estos lenguajes:

- C.
- C++.
- .NET (C#/VB.NET).
- Java ( JNA and JNI).
- Python.
- C Synchronous Interface.

### 2.2.2. ~~OpenKinect Analysis Library.~~

Esta bibliotecas se comunica con el API de OpenKinect y analiza la información en bruto en abstracciones más útiles.

- Hand tracking
- Skeleton tracking
- Depth processing
- 3D audio isolation coning?
- Audio cancellation (perhaps a driver feature?)
- Point cloud generation
- Inertial movement tracking with the built in accelerometer or an attached WiiMote
- 3d reconstruction
- GPU acceleration for any of the above

Sin embargo esta biblioteca no esta disponible aun. Ademas de estas bibliotecas el proyecto ofrece **OpenKinect Apps** que es código ejemplo, y aplicaciones

### ~~2.3. Free Kinect~~

### 2.4. Kinect SDK

Fue lanzado por Microsoft en Junio del 2011 para ser usado con Windows 7, el paquete de desarrollo SDK ofrece la posibilidad de explotar las capacidades del sensor Kinect dentro del entorno de trabajo de Windows. Incluye los drivers para el funcionamiento del dispositivo Kinect, estos pueden usarse para la programación de aplicaciones con C++, C-Sharp o Visual Basic, usando Microsoft Visual Studio 2010. Al descargar el SDK de la pagina oficial están incluidos ejemplos de códigos en los tres lenguajes mencionados.

#### 2.4.1. Tabla comparativa de versiones.

Característica	Beta 1	Beta 2	Versión 1	Versión 1.5
Robust Skeletal Tracking	•	•	•	•
API de Reconocimiento de voz	•	•	•	•
Cancelación de eco	•	•	•	•
Camara XYZ (depth)	•	•	•	•
Windows 7	•	•	•	•
Windows 8		•	•	•
Cámara a colores		•	•	•
Acceso a datos en bruto del sensor depth		•	•	•
Acceso a datos en bruto de la cámara		•	•	•
Acceso a datos en bruto del micrófono		•	•	•
Multithreading al usar Skeletal Tracking		•	•	•
Detección y gestión del estado de Kinect		•	•	•
Construcción de apps de 64 bits		•	•	•
Supresión de ruido de fondo		•	•	•
Near Mode			•	•
Soporte de 4 sensores kinect al mismo PC			•	•
Control de detección de usuario en S.T.			•	•
Microsoft Speech V.11 incluido			•	•
Acustica mejorada con far-talk			•	•
Soporte para idioma Español				•
Kinect Studio				•
Human Interface Guidelines (HIG)				•
Face Tracking SDK				•
Seat Skeletal Tracking				•
Skeletal Tracking Near Mode				•

#### 2.4.2. XNA.

XNA es un conjunto de tecnologías para video juegos de Microsoft, entre las cuales estan incluidas las siguientes: ?

- **DiectX.** Se trata de una librería de C++ para programar gráficos.
- **XNA Game Studio.** Son complementos de Visual Studio para la canalización del contenido y el despliegue de funciones usadas con el framework. Esta es la herramienta usada al hacer un juego para Xbox y enviar los datos a este.
- **Framework de XNA.** Es el conjunto de bibliotecas .NET construidas.

### 2.4.3. Microsoft Speech Platform Runtime.

Microsoft Speech Platform SDK v11 ofrece una nueva funcionalidad en las herramientas de desarrollo de Microsoft de gramática para ayudar a validar, depurar, probar y optimizar las gramáticas de las aplicaciones de voz. Entre sus herramientas se encuentran las siguientes:

- **CheckPhrase.exe.** Esta nueva herramienta tiene una frase y una gramática como entradas y salidas. Check Phrase emite frases personalizadas asociadas con la frase que se encuentra.
- **Confusability.exe.** Esta nueva herramienta identifica frases en una gramática que son fonéticamente similares. La herramienta puede ayudar a detectar las frases que a la larga pueden causar a los usuarios tener una mala experiencia, de tal manera que una frase en la gramática es falsamente reconocido por otra frase que también está en la gramática. La herramienta Confusability acepta varios archivos de gramática como de entrada y realiza su análisis en el conjunto de archivos de entrada especificados.
- **GrammarValidator.exe** Valida la sintaxis del elemento gramatical. Emite una serie de advertencias al validar una gramática, que incluye la detección de la repetición y las reglas inalcanzables.
- **PhraseGenerator.exe** Genera un subconjunto de frases sobre la base de los pesos en la gramática, y si lo desea, puede optar por no ampliar las referencias específicas de la regla.
- **Simulator.exe** Usted puede elegir si desea o no volver a utilizar el estado del motor en la estructura de la EMMA (Extensible Multi-Modal Annotation) de entrada, en lugar de una opción en la línea de comandos. Es decir, el estado motor de reconocimiento es ahora automáticamente reutilizado sólo cuando las expresiones están contenidas dentro de un emma en un bloque de secuencia. Además, el /ReuseEngineState de línea de comandos se ha eliminado.
- **SimulatorResultsAnalyzer.exe** Proporciona información adicional para el ajuste de sus gramáticas, incluyendo detección de frases .out

of grammar7 nuevos tipos de categorías de errores para darle un mejor análisis y métricas para el reconocimiento simulado. Añade nuevos elementos XML (AudioType, RecoResultSemantics y TranscriptSemantics) a la salida para hacer más fácil analizar los resultados sin necesidad de referirse de nuevo al archivo de salida del simulador.

#### **2.4.4. Kinect Speech Language Pack.**

Los paquetes de lenguaje de Kinect para Windows son complementos para el Kinect en tiempo de ejecución que usan el reconocimiento de voz para habilitar la forma en que cierto idioma se habla en diferentes lugares.

#### **2.4.5. Librería Kinect**

- KinectSensor
  - KinectSensors
  - Start
  - Stop
  - ColorStream
  - ColorFrameReady
  - SkeletonStream
  - InstanceIndex
  - DeviceConnectionId
  - NuiCamera
  - MaxElevationAngle
  - MinElevationAngle
  - ElevationAngle
  - Dispose
  - AllFramesReady: boolean
  - UniqueKinectId: boolean
  - MapDepthFrameToColorFrame: boolean
  - MapDepthToColorImagePoint: boolean
  - MapDepthToSkeletonPoint: boolean
  - MapSkeletonPointToDepth: boolean
  - MapSkeletonPointToColor: boolean
- RuntimeOptions
  - UseDepthAndPlayerIndex



- UseSkeletalTracking
  - UseColor
  - HasFlag
- KinectDeviceCollection
  - NotifyUserKinectsStatusChanged
  - Dispose: boolean
- StatusChangedEventArgs
  - Sensor
- DepthImageFrame
  - MapToColorPixel
- KinectStatus
  - Undefined: boolean
  - Initializing: boolean
  - DeviceNotGenuine: boolean
  - DeviceNotSupported: boolean
  - InsufficientBandwidth: boolean
- ColorImageStream
  - OpenNextFrame
  - Enable
  - Disable: boolean
  - StreamType: boolean
  - Resolution: boolean
  - Type: boolean
  - CreateCompatibleImageFrame: boolean
  - FrameHeight: int
  - FrameWidth: int
- ImageFrameReadyEventArgs
- ColorImageFrameReadyEventArgs
  - OpenColorImageFrame
- ColorImageFrame

- CopyPixelDataTo
- Type: boolean
- Resolution: boolean
- Image: boolean
- ViewArea: boolean
- ImageFrame: boolean
- Type: boolean
- Sourcestream: boolean
- Dispose: boolean
- BytesPerPixel: int
- Height: int
- Width: int
- PixelData: int
- DepthImageStream
  - OpenNextFrame
  - Enable
  - StreamType
  - Resolution
  - Type
  - CreateCompatibleImageFrame
  - IsTooFarRangeEnabled: boolean
  - Disable: boolean
  - FrameHeight: int
  - FrameWidth: int
- DepthImageFrameReadyEventArgs
  - OpenDepthImageFrame
- DepthImageFrame
  - MapToColorImage
  - MapFromSkeletonPoint
  - MapToSkeletonPoint
  - PlanarImage
  - ImageDigitalZoom
  - ImageStreamType

- ImageType
- ImageResolution
- ImageViewArea
- DepthImageToSkeleton
- SkeletonToDepthImageFrame
- PlayerIndexBitmask: boolean
- PlayerIndexBitmaskWidth: boolean
- DepthImageFormat: boolean
- DepthRange: boolean
- DepthImageFormat: boolean
- DepthRange: boolean
- DepthImagePoint: boolean
- DepthImagePointFloat: boolean
- SkeletonStream
  - OpenNextFrame
  - Disable: boolean
  - Enable: boolean
  - ChooseSkeletons: boolean
- SkeletonFrameReadyEventArgs
  - OpenSkeletonFrame
- SkeletonFrame
  - Timestamp
  - CopySkeletonDataTo
  - Quality
  - Dispose: boolean
- Skeleton
  - TrackingId
  - Parent
  - JointCollection
  - SkeletonPoint
  - Count: boolean
  - JointType: boolean

- Joint: int
  - Position: int
- FrameEdges
  - SkeletonFrameQuality
  - None: boolean
  - Left: int
  - Top: int
  - Right: int
  - Bottom: int
- KinectAudioSource
- FindCaptureDevices
- RetrieveTsStats
- QualityMetrics
- DevicePairGuid
- KinectAudioSource
- AutomaticGainControlEnabled
- SoundsSourceAngle
- SoundSourceAngleConfidence
- BeamAngleMode
- ManualBeamAngle
- BeamAngleChanged
- EchoCancellationSpeakerIndex
- AcousticEchoSuppression
- CenterClip
- EchoLength
- FeatureMode
- FrameSize
- GainBoulder

- MicArrayPreprocess
- MicrophoneIndex
- NoiseFill
- SourceMode
- VoiceActivityDetector
- EchoCancellationMode
- Dispose
- MicArrayMode
- SystemMode
- BeamAngleChangedEventArgs
- SoundSourceAngleChangedEventArgs: boolean
- EchoCancellationMode: boolean
- BeamAngleMode: boolean

## 2.5. Comparativa entre APIS.

En la siguiente tabla se muestra una comparación entre las APIS.

API	OpenNI	OpenKinect	Kinect SDK
Multi plataforma	•	•	
Multi lenguaje	•	•	•
Acceso alto nivel	•		•
Licencia	GPL y propietario	GPL	Propietario
Comentarios	Tiene acceso a alto nivel pero en modo Middleware (NITE)	No soporta skeleton tracking.	Solo funciona con Windows 7 o superior.