

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS E INGENIERIA



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DEL PERÚ

**ESTUDIO PRELIMINAR DEL USO DEL KINECT PARA EL ANÁLISIS
DE MARCHA HUMANA EN BASE A LA CINEMÁTICA DE LA
RODILLA**

Plan de Tesis para optar el Título de Ingeniero Electrónico

Presentado por:
Ricardo Joel Dongo Herrera

Asesor:
Leslie Casas
Rocío Callupe Pérez

Lima - PERÚ
2016

Agradecimientos

Agradecimiento especial a Innóvate-Perú por apoyar en la realización de esta tesis a través del proyecto "Sistema robótico tipo exoesqueleto para las extremidades inferiores que permita reproducir patrones de movimiento en el plano sagital para marcha normal o movimientos asistidos dirigidos conforme a procedimientos médicos" con Contrato 203-FINCYT-IA-2013.



Resumen

El análisis de movimiento es una herramienta muy usada en diversas áreas, como en la medicina de rehabilitación, donde el médico usa esta herramienta para la evaluación y diagnóstico de pacientes con alguna discapacidad motora como lo son las diferentes discapacidades que se pueden presentar en el miembro inferior. Actualmente, se utilizan equipos de captura de movimiento tridimensional para el análisis de la marcha utilizando, por mencionar algunos, arreglos de cámaras infrarrojas, marcadores retroreflectantes, escáner, entre otros. Sin embargo, estos equipos elevan el costo de la evaluación y diagnóstico, haciéndolo inaccesible el alcance de esto para personas de bajos recursos, por lo que es necesario desarrollar métodos de menor costo para lograr esto.

Partiendo de lo anterior, este estudio tiene por objetivo obtener datos cualitativos y cuantitativos que permitan determinar la precisión y exactitud de un equipo Kinect para el análisis de la cinemática de la marcha humana, esto mediante una comparación entre el Kinect y un sistema de captura de movimiento Vicon, ya que este último es un sistema validado y usado para el análisis de marcha. Como resultado se obtuvo curvas de movimiento de la rodilla en una marcha humana con ambos sistemas. En el caso del Kinect, se midió la curva de movimiento de tres voluntarios, haciendo uso de dos librerías ya existentes en el caso del Kinect: OpenNI, en conjunto con Nite, y Kinect for Windows SDK, para conocer la variación del ángulo de la rodilla al caminar. En cuanto al Sistema Vicon, se usó un patrón de marcha de la rodilla obtenido de 40 voluntarios cuya marcha no presentaba ninguna patología. Las pruebas de este trabajo de tesis se realizaron con tres voluntarios para la toma de datos de la marcha humana con el equipo Kinect, realizando 5 pruebas a cada uno, analizando los 3 mejores resultados de cada pierna, esto es, los 3 mejores datos del ángulo de la rodilla durante una marcha humana. De esto se concluyó que hubo mayor precisión y exactitud con el Kinect cuando se usó la librería SDK. Además, con el uso de la librería SDK se obtuvo una forma de la curva del ángulo de la rodilla más similar al patrón obtenido mediante el sistema Vicon a comparación de las curvas de movimiento obtenidas con el uso de la librería OpenNI. En todos los casos se obtuvo la forma de una curva del ángulo de la rodilla durante la marcha lo suficientemente cercana al patrón de marcha obtenido mediante el sistema Vicon, con coeficientes de correlación superiores a 0.8, y se demuestra así que el Kinect puede ser usado para la evaluación de una marcha humana. De esta manera, se espera que este estudio sirva de línea base para futuros estudios y proyectos con el Kinect para el análisis de marcha.



TEMA DE TESIS PARA OPTAR EL TÍTULO DE INGENIERO ELECTRÓNICO

Título : Estudio preliminar del uso del Kinect para el análisis de marcha humana en base a la cinemática de la rodilla.
Área : Bioingeniería * 1213
Asesor : Leslie Casas Guido, Rocio Callupe Perez
Alumno : Ricardo Joel Dongo Herrera
Código : 20060569
Fecha : 07/05/14



Descripción y Objetivos

El análisis de marcha es un estudio hecho por muchos doctores y terapeutas físicos para el tratamiento de personas que presentan algún problema o dificultad a la hora de caminar. No obstante, los métodos normalmente usados dependen mucho del factor humano, lo que los convierte en métodos no muy confiables y no muy precisos.

La captura de movimiento para modelamiento en 3D ofrece una obtención y análisis de datos de mayor precisión y exactitud, pudiendo así ser de gran ayuda a la hora de realizar los diagnósticos y tratamientos. Sin embargo, el alto costo de los equipos utilizados hacen que el tratamiento también sea costoso. El alto costo es un problema aún mayor en los países pobres, debido al gran porcentaje de la población que no cuenta con los recursos económicos para este tratamiento.

El Kinect es una herramienta que ha demostrado ser muy útil en lo que se refiere a modelamiento en 3D y reconocimiento de objetos, llevando a que se realicen varios estudios y proyectos con este.

El fin de este estudio es comparar qué tan útil es el Kinect para el análisis de marcha en comparación con un sistema de cámaras usadas para captura de movimiento, ya que este último es el método que se usa para este tipo de tratamientos con análisis de marcha y dejar así una línea base que sirva para futuras investigaciones y proyectos con el Kinect para el análisis de marcha.

Rocio Callupe

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
SECCIÓN ELECTRICIDAD Y ELECTRÓNICA

M. Sc. Ing. MIGUEL ANGEL CATANO SANCHEZ
Coordinador de la Especialidad de Ingeniería Electrónica

[Handwritten signatures]

MÁXIMO 50 PÁGINAS



TEMA DE TESIS PARA OPTAR EL TÍTULO DE INGENIERO ELECTRÓNICO

Título : Estudio preliminar del uso del Kinect para el análisis de marcha humana en base a la cinemática de la rodilla.

Índice

Introducción

1. Descripción del análisis de marcha, de su utilidad en la terapia física y de las técnicas usadas.
2. Conceptos generales sobre el modelamiento en 3D, equipos y software utilizado.
3. Adquisición y comparación de datos para el análisis de marcha usando el Kinect y seis cámaras Vicon.
4. Análisis de los resultados.

Conclusiones

Recomendaciones

Bibliografía

Anexos

MÁXIMO 50 PÁGINAS

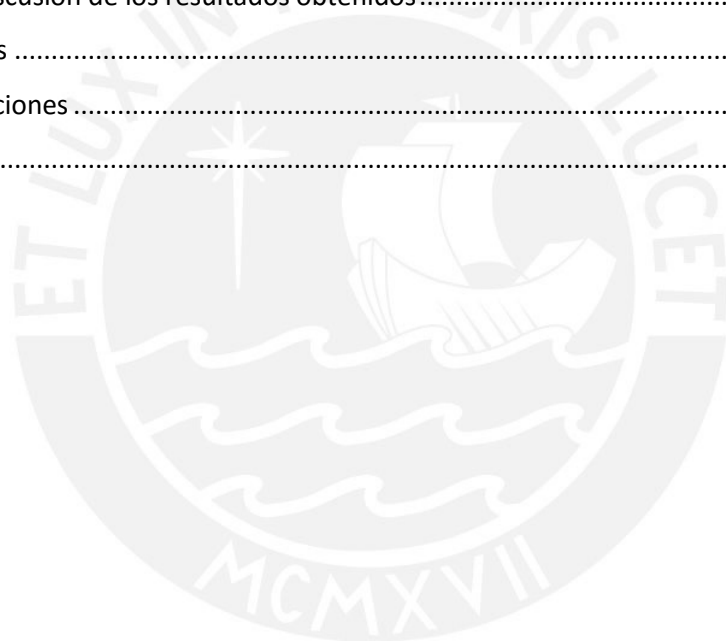
PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
SECCIÓN ELECTRICIDAD Y ELECTRÓNICA

M. Sc. Ing. MIGUEL ANGEL CATANO SÁNCHEZ
Coordinador de la Especialidad de Ingeniería Electrónica

Índice

Introducción	1
CAPITULO 1: Descripción del análisis de marcha, de su utilidad en la terapia física y de las técnicas usadas.....	3
1.1. Análisis de marcha.....	3
1.2. Fases del ciclo de la marcha	3
1.3. Adquisición de datos	4
1.4. Técnicas para realizar el análisis de marcha	5
1.5. Estado de la investigación en el análisis de marcha	7
1.6. Objetivos	8
1.6.1. Objetivos Generales	8
1.6.2 Objetivos Específicos	9
1.7. Condiciones iniciales	9
Capítulo 2: Conceptos Generales sobre el modelamiento en 3D, equipos y software utilizado.....	10
2.1. Conceptos generales sobre el modelamiento en 3D	10
2.2. Sistema de captura de movimiento en 3D Vicon V460.....	10
2.2.1. Introducción al Sistema Vicon	10
2.2.2. Arquitectura del Sistema Vicon	11
2.2.3. Especificaciones técnicas	11
2.3. Kinect.....	12
2.3.1. Introducción al Kinect	12
2.3.2. Arquitectura del Kinect	12
2.3.3. Funcionamiento del Kinect	14
2.3.4. Limitaciones del Kinect	15
2.3.5. Software utilizado con el Kinect	15
2.3.6. Estudios realizados para captura de movimiento y el Kinect	16
2.4. Software utilizado	17
2.4.1. Openni	17
2.4.2. NITE	24
2.4.3. Kinect for Windows SDK	28
Capítulo 3: Análisis y comparación de datos para el análisis de marcha usando el Kinect y seis cámaras Vicon	33

3.1.	Protocolo de adquisición.....	33
3.1.1.	Ambiente de trabajo	33
3.1.2.	Ubicación del Kinect	35
3.1.3.	Características físicas de los voluntarios	35
3.1.4.	Equipos usados	36
3.1.5.	Pruebas	37
3.2.	Resultados	40
3.2.1.	Voluntario 1	44
Capítulo 4: Análisis de los resultados.....		54
4.1.	Cuadros comparativos de los resultados obtenidos	54
4.2.	Discusión de los resultados obtenidos	55
Conclusiones		57
Recomendaciones		58
Fuentes.....		60



GLOSARIO

API	Significa Interfaz de programación de aplicaciones. Es un conjunto de funciones que ofrecen las librerías para ser usadas por otro software.
Framework	Significa marco de trabajo. Es un conjunto de bibliotecas orientadas a reutilizar componentes de software para desarrollar aplicaciones a mayor velocidad. Ejemplo: .Net. o Java.
Middleware	Software cuya función es que una aplicación pueda comunicarse con otras aplicaciones, con hardware, redes o sistemas operativos.
Aplicación	Software que permite al usuario realizar varios trabajos diferentes. Los procesadores de texto son un ejemplo de esto.
ID	Es un identificador. Se usa para etiquetar objetos con los que se esté trabajando o analizando.
Interacción Natural	Su objetivo es permitir a un usuario interactuar con una computadora como si estuviera interactuando con otra persona; es decir, a través de gestos o la voz.

INTRODUCCIÓN

El análisis de marcha es un estudio del movimiento del cuerpo humano en las fases de la marcha y puede ser realizado por muchos doctores de rehabilitación y terapeutas físicos en el campo clínico, esto para evaluar y diagnosticar pacientes que presenten problemas en sus capacidades motoras. No obstante, los métodos de evaluación tradicional de la marcha humana, normalmente usados, dependen mucho del factor humano, lo que los convierte en métodos no muy confiables y no muy precisos.

Una de las técnicas desarrolladas en los últimos 30 años para el análisis de marcha es la captura de movimiento 3D, la cual ha demostrado tener mayor precisión y exactitud en la obtención de datos, pudiendo así ser de gran ayuda a la hora de realizar los diagnósticos y tratamientos. Sin embargo, el alto costo de los equipos utilizados hace que la evaluación de la marcha también sea costosa, como por ejemplo usando los equipos de la marca Vicon. Esto es un problema aún mayor en los países pobres, debido al gran porcentaje de la población que no cuenta con los recursos económicos para una evaluación de la marcha en 3D.

Frente a esta problemática, el equipo Kinect que realiza el modelamiento en 3D y reconocimiento de objetos se presenta como una alternativa de bajo costo y de fácil acceso para el análisis de la marcha.

El objetivo principal de este trabajo de tesis es el estudio comparativo entre el Microsoft Kinect frente al sistema de captura de movimiento Vicon, con el fin de determinar la factibilidad del uso del Kinect para el análisis de marcha. Además, este estudio servirá de línea base para futuras investigaciones y proyectos en este campo de aplicación.

En el capítulo 1, se describirá el concepto de análisis de marcha, las fases del ciclo de la marcha, la adquisición de datos y las técnicas utilizadas.

En el capítulo 2, se describirá la relación del análisis de marcha y el modelamiento en 3D, el sistema de captura de movimiento en 3D Vicon, el sensor Kinect y sus librerías, OpenNI y Windows SDK.

En el capítulo 3, se describirá a detalle el protocolo de adquisición utilizado para la toma de datos con el sistema de captura de movimiento en 3D Vicon y el sensor Kinect, incluyendo dimensiones del ambiente en que se realizaron las pruebas, la vestimenta

usada por los voluntarios y como se realizaron las pruebas. También se describirá los pasos seguidos para el procesamiento de los datos del Kinect y se mostrará los resultados obtenidos en exactitud y precisión.

Y en el capítulo 4, se realizará un análisis de los resultados y se presentarán recomendaciones para futuros trabajos en este campo.



CAPITULO 1: DESCRIPCIÓN DEL ANÁLISIS DE MARCHA, DE SU UTILIDAD EN LA TERAPIA FÍSICA Y DE LAS TÉCNICAS USADAS

1.1. Análisis de marcha

La marcha puede ser definida como “el método de locomoción que requiere del uso de dos piernas, alternadamente, para proveer soporte y propulsión”. Para diferenciarla del proceso de correr, hay que agregar “por lo menos un pie en contacto con el suelo todo el tiempo”. Aunque los conceptos de marcha y caminar suelen confundirse, la diferencia está en que la marcha es la forma o estilo de caminar, y no el proceso de caminar mismo [1].

El análisis de marcha es entonces, el estudio del movimiento que realizan las piernas al caminar, tomando datos sobre todo de los huesos, músculos, articulaciones y sistema nervioso.

Se usa en el campo de aplicación de la medicina de rehabilitación, en lo que concierne al estudio de las patologías o anormalidades en la forma de caminar de las personas.

1.2. Fases del ciclo de la marcha

Las fases del ciclo de la marcha son [2]:

- **Período de apoyo:** Es la fase de la marcha en la que el pie está en contacto con el suelo. Empieza cuando el talón hace contacto con el suelo y termina cuando el ante pie deja el suelo. Representa un 60% del ciclo de la marcha.
- **Período de oscilación:** Es la fase de la marcha en la que el pie no contacta con el suelo.

Empieza después que el ante pie deja el suelo y termina cuando el talón vuelve a hacer contacto con el suelo. Representa el 40% del ciclo de la marcha.

- **Doble apoyo:** Fase en la que ambos pies contactan con el suelo.
- **Apoyo unilateral:** Fase en la que un solo pie se encuentra en contacto con el suelo, durante el período de oscilación del otro pie.

En la figura 1.1 se representa un ciclo completo de análisis de marcha.

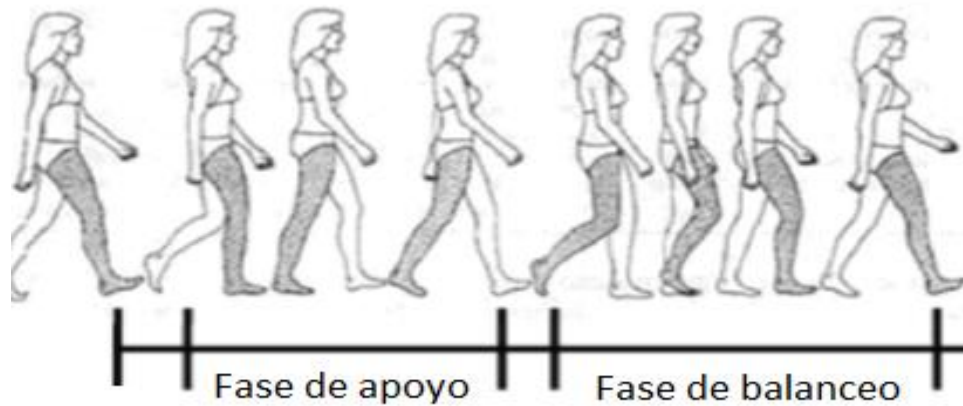


Figura 1.1. Ciclo completo de marcha de la pierna derecha (color gris). Modificada de [2]

1.3. Adquisición de datos

El análisis de marcha permite obtener información para el tratamiento de personas con anomalías de la marcha. Para esto se hace uso de diferentes técnicas para adquirir datos a partir de los cuales se obtienen curvas que puedan ser comparadas con la de personas sin patologías en la marcha. Para hacer esta comparación, se usan curvas ya existentes que sirven como patrones de marcha para determinar patologías en la marcha de una persona.

Para el análisis de marcha, se toma como referencia 3 planos: El plano sagital, el plano transversal y el plano frontal. En la figura 1.2 pueden apreciarse estos planos.

- **Plano Sagital**: Es el plano que divide al cuerpo en derecha e izquierda.
- **Plano Frontal**: Es el plano que divide al cuerpo en parte posterior y parte anterior.
- **Plano Transversal**: Es el plano que divide al cuerpo en parte superior y parte inferior.

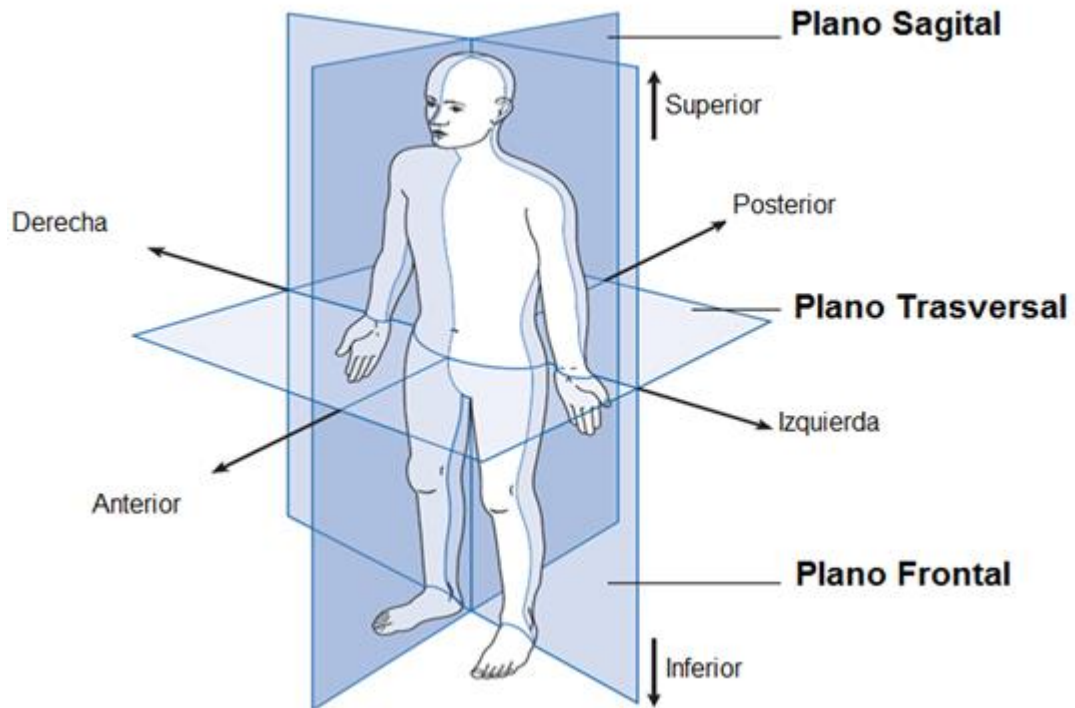


Figura 1.2. La posición anatómica, con tres planos de referencia y seis direcciones fundamentales. Modificado de [1]

Los posibles movimientos que puede realizar una articulación son los siguientes [1]:

- **Flexión y Extensión:** Toman lugar en el plano sagital. En el caso del tobillo se llaman dorsiflexión y plantarflexión.
- **Abducción y Aducción:** Toman lugar en el plano frontal.
- **Rotación interna y externa:** Toman lugar en el plano transversal. También son llamados rotación media y lateral respectivamente.

1.4. Técnicas para realizar el análisis de marcha

Hay diversas técnicas utilizadas por los especialistas en rehabilitación física para identificar patologías en la marcha de una persona, de estas las más utilizadas son las que se mencionan a continuación:

- **Análisis visual:** Esta técnica se utiliza para obtener parámetros básicos de la marcha como ciclo de tiempo (o ritmo), longitud del paso y velocidad para realizar los diagnósticos. Este método puede realizarse viendo directamente al paciente caminar varias veces o usar una cámara para grabarlo y observar la grabación posteriormente. El uso de los ojos, aunque puede ser útil para casos simples, es de hecho impreciso, ya que depende mucho de la habilidad del que observa. Con el uso de cámaras, es posible llevar un registro, además de que no es necesario que el paciente repita el proceso de caminar tantas veces, ya que el que analiza puede repetir el video todas las veces necesarias para observar las anomalías. Este segundo método es más preciso, más aún si la observación del video se realiza con un grupo de personas, para así poder llegar a mejores conclusiones; sin embargo, no ofrece datos cuantitativos ni es tan exacto como otros métodos.
- **Cinética:** Permite medir la fuerza ejercida al realizarse los movimientos necesarios para caminar. Con esto se puede realizar curvas como fuerza y momento en la articulación versus porcentaje de paso. Este tipo de curvas permite ver como se distribuye el peso al momento de caminar entre las piernas, la cadera, el tronco y las articulaciones.
- **Cinemática:** Estudia la trayectoria de los cuerpos en el tiempo sin considerar las fuerzas que lo originan. Para el análisis de marcha, se estudia la variación del ángulo de la articulación de interés en el tiempo. Como el ciclo de marcha es repetitivo, es posible analizar su trayectoria para detectar anomalías que pudieran presentarse.
- **Cinética/Cinemática:** Es una combinación de los dos métodos anteriores, proporcionando así casi todos los datos necesarios de la forma de caminar de una persona. Este método es el más completo.

Para este estudio se medirá la cinemática de la rodilla en el plano sagital. En la figura 1.3 se muestra la curva teórica de la variación del ángulo de la rodilla en un ciclo de marcha:

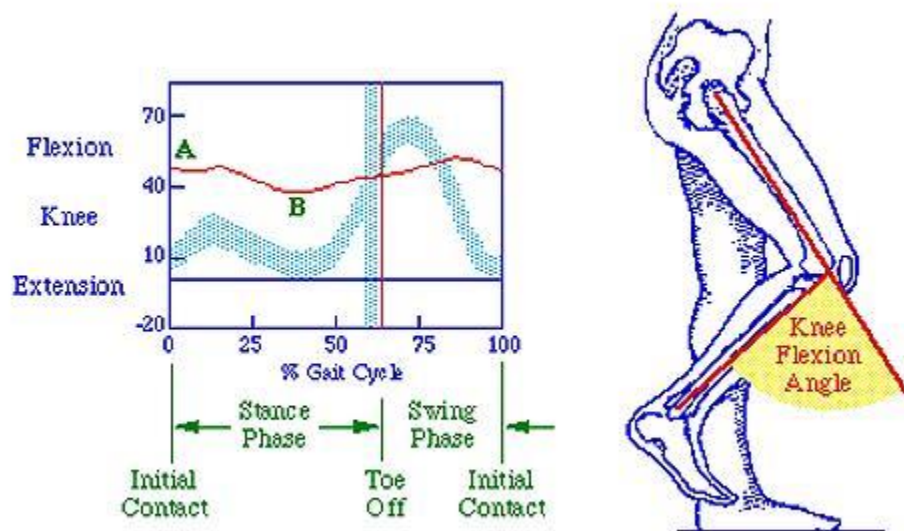


Figura 1.3. Variación del ángulo de la rodilla en un ciclo de marcha [3]

1.5. Estado de la investigación en el análisis de marcha

Varios estudios se han realizado sobre la utilidad y confiabilidad del análisis de marcha en el tratamiento de los pacientes que presentan problemas o dificultades a la hora de caminar. En el artículo hecho por Tishya A.L. Wren y su equipo [4], se analiza la eficacia del análisis de marcha para el tratamiento de personas con problemas o dificultades para caminar. En este artículo se mencionan varios estudios realizados entre Enero del 2000 hasta Setiembre del 2009 (junto con publicaciones anteriores relevantes) concluyen que el análisis de marcha permite obtener datos suficientemente precisos como para cambiar el tratamiento que se le sigue a un paciente. Por ejemplo, los resultados de los tratamientos en los pacientes han sido mejores si se ha seguido las recomendaciones del análisis de marcha, en especial si se ha realizado una cirugía siguiendo las recomendaciones de este.

En el artículo, hecho por el equipo de Jennifer L. McGinley [5], se evaluó los resultados de varios informes para evaluar la confiabilidad de los datos obtenidos del modelamiento en 3D para el análisis de marcha. En este artículo se concluyó que la mejor confiabilidad provenía de los datos obtenidos de la cadera y rodilla en el plano sagital, mientras que la menor confiabilidad provenía de los datos obtenidos de la cadera y rodilla en el plano transversal.

En otro artículo, hecho por el equipo de Catia Miyuki Kawamura [6], se hizo una comparación entre el método visual y la captura de movimiento en 3D para el análisis de marcha en pacientes con parálisis cerebral. Para este estudio, cuatro observadores vieron videos del caminar de diferentes pacientes, analizando diez puntos de interés en el ciclo de cada paso: flexión de la cadera en posición terminal, flexión de la rodilla en el contacto inicial, extensión de la rodilla en posición terminal, flexión de la rodilla en balanceo inicial, dorsiflexión del tobillo en contacto inicial, oblicuidad de la pelvis en posición media, aducción de la cadera en respuesta de carga, rotación de la pelvis, rotación de la cadera en posición media y progresión del ángulo del pie, en relación con el miembro inferior y la posición media. Los datos que obtuvieron fueron comparados con datos obtenidos de un análisis 3D de la cinemática del paciente. Los datos que obtuvieron los observadores eran muy similares entre ellos, pero cuando estos fueron comparados con los datos del análisis 3D de la cinemática del paciente, solo los datos de dos puntos fueron similares, los de la flexión de la rodilla en contacto inicial y la oblicuidad de la pelvis, concluyendo así que solo estos datos del ciclo del paso de un paciente son confiables si se analizan con el método visual, mientras que los otros ocho puntos de interés analizados del ciclo de paso en este estudio no lo son.

Se han realizado varios estudios donde se utiliza el sistema de captura de movimiento en 3D Vicon como referencia para comparar y validar sistemas de captura de movimiento desarrollado por investigadores. Por ejemplo, en el estudio realizado por Kaiyu Tong [16], se busca utilizar un giroscopio de un eje para evaluar el ángulo de la rodilla, velocidad angular e inclinación de la pantorrilla (shank) y del muslo (thigh); en este estudio se hace uso del sistema Vicon como referencia para validar los datos obtenidos por el giroscopio. Algo similar se hace en el estudio de Ruth E. Mayagoitia [17], donde se desarrolla un sistema de 4 acelerómetros de un eje y un giroscopio, donde también se usa el sistema Vicon como referencia para validar este sistema.

1.6. Objetivos

1.6.1. Objetivos Generales

Realizar un estudio comparativo entre el Microsoft Kinect y un sistema de captura de movimiento en 3D Vicon para determinar la precisión y exactitud del Kinect para el análisis de marcha, haciendo uso de librerías y algoritmos ya existentes. De esta manera

se espera que este estudio sirva de línea base para futuros estudios y proyectos con el Kinect para el análisis de marcha.

1.6.2 Objetivos Específicos

- a) Obtener los ángulos de las articulaciones de las rodillas en el proceso de marcha usando el Kinect mediante 2 librerías diferentes, con al menos 3 personas por cada uno.
- b) Obtener, mediante el sistema Vicon, un patrón de la variación del ángulo de la rodilla en el proceso de marcha válido para una persona de nacionalidad peruana, el cuál sirva como referencia para el análisis de la exactitud y precisión del Kinect.
- c) Hacer una comparación entre los datos obtenidos con el Kinect y mediante el sistema de captura de movimiento en 3D Vicon.
- d) Realizar conclusiones sobre los resultados obtenidos en este trabajo para que sirvan como línea base para futuros estudios y aplicaciones con el Kinect para el análisis de marcha.

1.7. Condiciones iniciales

Para determinar si es factible realizar análisis de marcha con el Kinect, las curvas obtenidas con el Kinect deben tener una trayectoria semejante a las curvas obtenidas mediante el sistema de captura de movimiento en 3D Vicon y encontrarse dentro del rango del patrón de marcha obtenido con este.

CAPÍTULO 2: CONCEPTOS GENERALES SOBRE EL MODELAMIENTO EN 3D, EQUIPOS Y SOFTWARE UTILIZADO

2.1. Conceptos generales sobre el modelamiento en 3D

El análisis de marcha es una herramienta muy usada en diversas áreas, por ejemplo, en la medicina de rehabilitación. Actualmente se utilizan equipos de captura tridimensional, utilizando arreglos de cámaras infrarrojas, marcadores retroreflectantes, escáner, entre otros. Sin embargo, estos equipos elevan el costo del tratamiento, haciéndolo inaccesible para personas de bajos recursos, por lo que es necesario desarrollar métodos de menor costo para lograr esto.

El Kinect ha demostrado ser una alternativa para sistemas de captura de menor costo, como en la obtención de datos para el procesamiento de imágenes, lo cual ha llevado a realizar varios estudios sobre sus aplicaciones para el reconocimiento de personas y objetos. Además, es portable, comercial y de fácil acceso.

2.2. Sistema de captura de movimiento en 3D Vicon V460

2.2.1. Introducción al Sistema Vicon

El sistema de captura de movimiento en 3D Vicon V460 es un sistema de fotogrametría, el cual permite capturar el movimiento de una persona en tiempo real, y se genera un modelo en 3D usando el software del sistema (Workstation), del cual se puede conocer la cinemática de la persona. Para esto se coloca marcadores (esferas pequeñas de 25mm de diámetro y 1gr. de peso) sobre la persona, los cuales son vistos por las 6 cámaras Pulnix TM6701 del sistema, lo cual permite determinar la posición de estas en el espacio, lo que servirá para la generación del modelo en 3D [2].

Se debe evitar el uso de ropa suelta por parte de los voluntarios, ya que esto causa que los marcadores que utiliza el sistema Vicon se muevan mucho durante la toma de datos, generando errores. El cabello largo también puede ser una fuente de error, ya que puede tapar los marcadores. Tanto cabello largo como ropa suelta también es una fuente de error para el NITE, como se explicará más adelante en el punto 2.4.2.



Figura 2.1. Laboratorio Multifuncional ubicado en la Sección de Ingeniería Mecánica de la Pontificia Universidad Católica del Perú [20]

2.2.2. Arquitectura del Sistema Vicon

- *Estación base*: Hardware que captura los datos enviados por las seis cámaras y transforma en 3D.
- Seis cámaras de video: 6 cámaras Pulnix 6710 de alta velocidad con obturación electrónica. Posee un rango de frecuencia de 60 hasta 200 Hz, y a través de una interfaz de video digitaliza la señal analógica.

2.2.3. Especificaciones técnicas

- Sensor CCD con obturación electrónica.
- Estroboscopio infrarrojo.
- Frecuencia de muestreo de hasta 250 Hz.
- 6 Visible Strobe front S/S: Permiten regular la cantidad de emisión de luz de la cámara.
- Seis cables de 12W SYnc 12M: Medio de conexión entre la cámara y la interfaz.
- Dos cables de 14W Straight 25M: Medio de conexión entre la interfaz y la estación base.
- Cable de alimentación.

- Marcadores: Esferas cubiertas de un material retroreflectante, su diámetro es 25mm y permiten capturar los movimientos de grandes volúmenes.
- Interfaz de vídeo: permite la conexión entre las cámaras y la estación base.
- Estación de trabajo (*Workstation*): el equipo es una computadora, y es en ella donde se instala el software y se configura el protocolo.
- Capacidad de captura de datos: 24 horas.
- Reconstrucción de las trayectorias de los marcadores en 3D.
- Permite la visualización del movimiento de los marcadores en varias ventanas en el mismo instante.

Conexión de la red entre la PC (*Workstation*) y la Estación base

- Para conectarlos se usa un cable crossover, y utiliza el protocolo de red TCP/IP.
- La estación base del sistema Vicon tiene los siguientes parámetros de red [2]:
 - IP: 192.168.0.11
 - máscara de subred: 255.255.255.0

2.3. Kinect

2.3.1. Introducción al Kinect

El Kinect es categorizado como un sensor de profundidad. Las cámaras de profundidad están cambiando la visión de las máquinas y la percepción de los robots alrededor del mundo reemplazando las cámaras de monovisión, estereovisión y los sensores de rango, como los sensores que hacen uso de láser, sensores de ultrasonido y radares. El Kinect está siendo usado mucho en diferentes proyectos, reemplazando así a los sensores anteriormente mencionados, debido a que puede producir imágenes de calidad decente (resolución de 640x480) e información de profundidad del ambiente en el que se encuentra, todo esto a un costo más bajo.

2.3.2. Arquitectura del Kinect

El Kinect está formado por una cámara de monovisión (RGB) que puede ser procesada como en cualquier sistema de procesamiento de imágenes, y un sensor de profundidad que consiste en un proyector infrarrojo combinado con un sensor monocromo CMOS, que captura datos de profundidad bajo cualquier condición de luz ambiental. Usando los dos

juntos, se puede estimar la profundidad de un objeto con los datos que provienen del Kinect.

El Kinect también cuenta, como se puede ver en la figura 2.2, con un procesador que captura y muestra las imágenes apropiadamente del sensor infrarrojo y de la cámara de monovisión (RGB) a través del puerto USB.

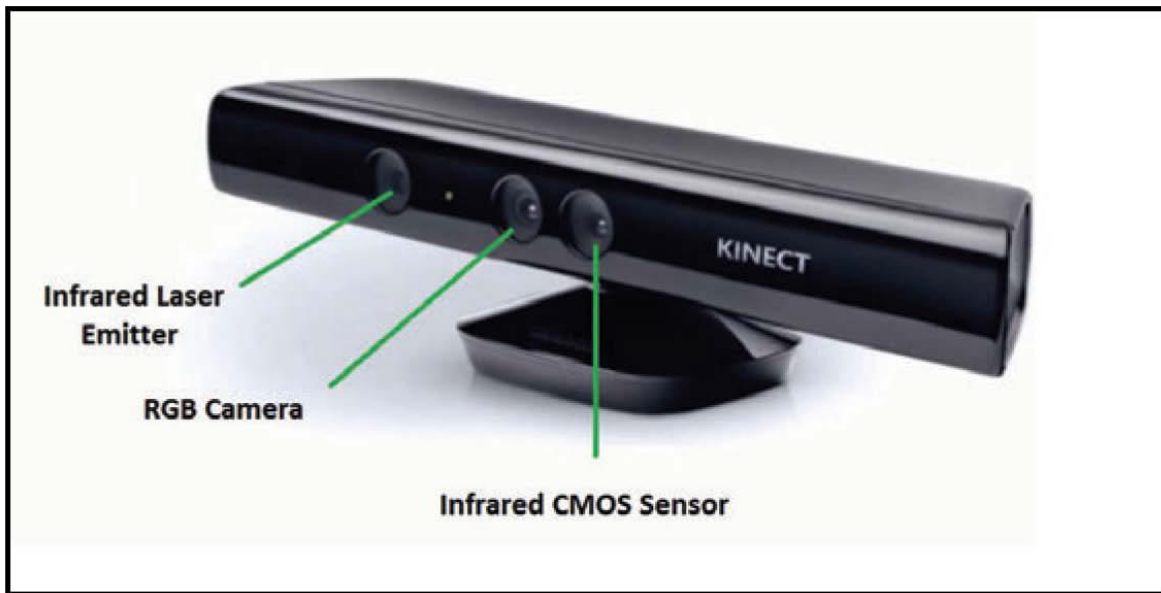


Figura 2.2. Kinect [7]

Las especificaciones técnicas del Kinect son las siguientes [7]:

- a) Sensor:
 - i. Lentes de color y detección de profundidad.
 - ii. Arreglo de micrófonos de voz.
 - iii. Motor de inclinación para ajuste del sensado.

- b) Campo de visión:
 - i. Campo horizontal de visión: 57 grados.
 - ii. Campo vertical de visión: 43 grados.
 - iii. Rango de inclinación: +/- 27 grados.
 - iv. Rango del sensor de profundidad: 1.2 – 3.5 metros.

- c) Flujo de información:
 - i. 320x240 16-bit profundidad@ 30 tramas/segundo.
 - ii. 640x480 32-bit color@ 30 tramas/segundo.
 - iii. 16-bit audio@ 16 KHz.

- d) Sistema de rastreo de esqueleto:
 - i. Rastreo de hasta 6 personas, 2 activos al mismo tiempo.
 - ii. Rastreo de 20 articulaciones por persona activa con SDK y 14 con OpenNI.

En la **figura 2.3** se puede observar la arquitectura interna del Kinect.

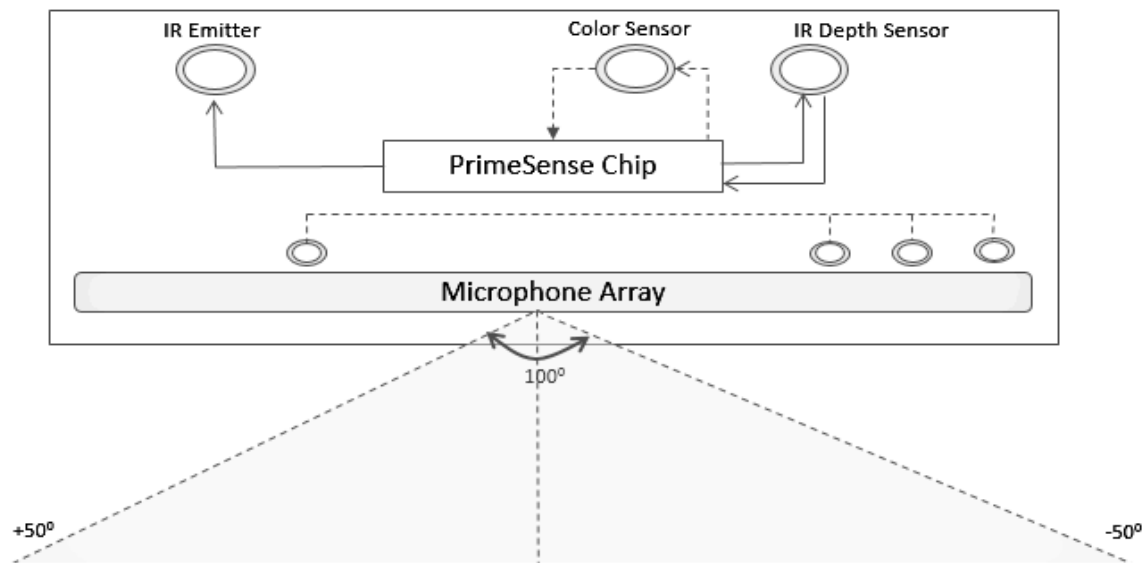


Figura 2.3. Arquitectura del Kinect [8]

2.3.3. Funcionamiento del Kinect

La cámara de color (**color sensor en la figura 2.3**) es la encargada de capturar la data de color. Detecta los colores rojo, azul y verde de la fuente y envía una trama de imágenes de color.

El emisor y la cámara infrarroja (**IR emitter e IR depth sensor en la figura 2.3 respectivamente**) son los encargados de proporcionar la data de profundidad. El emisor infrarrojo emite constantemente luz infrarroja en forma de puntos sobre todo dentro de su

área de visión. Estos puntos de luz se reflejan de todos los objetos enfrente del Kinect y son detectados por la cámara infrarroja, la cual los convierte en data de profundidad al medir la distancia entre el Kinect y el objeto de donde fue reflejado el punto de luz.

2.3.4. Limitaciones del Kinect

El Kinect ha probado tener algunas limitaciones, como se menciona en el artículo hecho por Riyad A. El-laithy y su equipo [9], donde se menciona que la luz intensa ciega al sensor infrarrojo del Kinect, debido a la excesiva cantidad de luz infrarroja, notándose que se obtienen mejores datos si se trabaja de noche o en un lugar cerrado. Además, se menciona que el Kinect no puede detectar vidrio o cualquier objeto transparente, debido a que la luz infrarroja del Kinect se refracta y no rebota al Kinect.

2.3.5. Software utilizado con el Kinect

Para realizar el esqueleto virtual con el Kinect, del cual se extraerá la información de las articulaciones, existe una variedad de software y librerías que pueden ser usadas. Si se desea trabajar en Windows, existe una librería desarrollado por Microsoft como Kinect for Windows Software Development Kit (Kinect SDK) que puede ser descargada de forma gratuita. En lo que se refiere a Linux, también existen librerías y software desarrollados por programadores e investigadores que pueden ser descargados gratuitamente que permiten obtener los datos del Kinect y realizar el esqueleto virtual. Entre las librerías disponibles se encuentran: OpenCV, OpenNI, Nite, PCL, Libfreenect, entre otras. Notar que estas librerías mencionadas también pueden ser instaladas en Windows. De las librerías desarrolladas por programadores e investigadores, la más completa y la que más ha sido usada en los estudios consultados, es la combinación OpenNI-Nite, por lo que será uno de los métodos usados en este estudio; el Kinect for Windows SDK será el otro método, ya que esta fue desarrollada los mismos creadores del Kinect.

Cada librería utiliza su propio algoritmo para ordenar la nube de puntos de luz captados por la cámara infrarroja y obtener la imagen de profundidad con la que se trabajará. Además, las librerías son capaces de reconocer las diferentes partes del cuerpo del individuo, lo cual será necesario para la realización del esqueleto virtual.

2.3.6. Estudios realizados para captura de movimiento y el Kinect

En lo que se refiere a captura de movimiento, se han realizado varios estudios sobre como estimar la pose de una persona y de cómo obtener el esqueleto virtual de formas más eficientes usando imágenes de profundidad, como la que usa el Kinect.

Un ejemplo, es el artículo realizado por Loren Arthur Schwarz [9], que realizó el estudio sobre un método para estimación de pose de todo el cuerpo usando cámaras ToF (Time of Flight), las cuales miden el “tiempo de vuelo” de la señal de luz entre la cámara y la persona en cada punto de la imagen, para detectar puntos conocidos del cuerpo en la data 3D y fijar un esqueleto virtual usando cinemática inversa. En vez de tratar de identificar poses conocidas, este método consiste en construir una representación gráfica de la información de las cámaras ToF que permite medir las distancias geodésicas entre las partes del cuerpo, ya que la distancia geodésica entre dos partes del cuerpo es independiente del movimiento y los resultados indicaron que este método permite detectar las poses del cuerpo de forma más precisa.

Otro ejemplo, es el artículo de Jamie Shotton [10], quien propuso un método para estimar rápidamente la posición de las articulaciones de una imagen de profundidad, sin usar información temporal. El primero utiliza el método de reconocimiento de objetos, luego se realiza una representación intermedia de las partes del cuerpo (se reconoce cada parte del cuerpo y se le da un color diferente a cada parte), convirtiendo el problema de estimación de pose a un problema más sencillo de clasificación de píxeles, para finalmente generar un modelo donde están marcadas cada articulación del usuario (representada por un punto) y el resultado obtenido fue un gran aumento en la precisión de la detección de las articulaciones de una persona.

En lo que se refiere al Kinect, este ha demostrado ser una herramienta valiosa en lo que se refiere a la obtención de datos para el procesamiento en 3D, como se menciona en el artículo hecho por Zhengyou Zhang [11], donde se indica que el Kinect ofrece muchas posibilidades, como por ejemplo reconocimiento de postura de manos, reconocimiento de actividad humana, estimación de la biométrica del cuerpo, reconocimiento de la superficie del cuerpo, y aplicaciones para la salud.

En otro artículo, hecho por Mark Draelos [12], se realiza un estudio donde se busca que el Kinect pueda reconocer objetos a una distancia más corta. Para lograr esto, se utilizó

unos lentes amplificadores Nyko (Nyko Zoom), el cuál encaja sobre los tres lentes del Kinect y se utilizó un algoritmo de calibración diferente. Los resultados fueron que al usar los lentes Nyko y modificando el software se logró reducir el rango mínimo de visión del Kinect en 30%, demostrando que es posible para el Kinect reconocer objetos a distancias más cortas.

Asimismo, el artículo hecho por Antonio Padilha Lanari Bó [13], se investigó el uso combinado del Kinect y de sensores inerciales (girómetro y acelerómetro) para la obtención de datos de las articulaciones de una persona al moverse, esto para mejorar sus deficiencias individuales, mejorando la visualización del movimiento y la precisión. Los resultados indicaron que el Kinect permite simplificar los procedimientos de inicialización, mejor visualización de los ángulos estimados, y la posibilidad de calibrar los sensores en tiempo real.

Además, en lo que respecta al uso del Kinect para el análisis de la marcha, se han realizado estudios para verificar su precisión y poder realizar su validación. En el estudio de Ross A. Clark [18], se buscaba evaluar la capacidad del Kinect para medir variables espacio-temporales (Tiempo de paso, longitud del paso, velocidad de la marcha, tiempo de un ciclo de la marcha, longitud de un ciclo de marcha, velocidad de balanceo del pie), comparando las mediciones realizadas con este sistema con el de un sistema Vicon, determinando que la similitud entre los datos obtenidos por el sistema Vicon y el Kinect va desde regular hasta buena. Alexandra Pfister [19] realizó una evaluación similar a la que se va a realizar en este estudio, evaluando la flexión y extensión de la rodilla y de la cadera con el Kinect y comparándolo con el sistema Vicon para validación, llegando a la conclusión de que el Kinect tiene una precisión muy variable. Cabe mencionar que ella coloca el Kinect con una inclinación de 45° con respecto a una rueda de andar sobre la cuál camina los voluntarios; además de que su análisis es cuantitativo y no incluye la parte cualitativa.

2.4. Software utilizado

2.4.1. Openni

El Openni (que en inglés significa “*Open Natural Interaction*”) es un conjunto de librerías, o ‘*framework*’ que define APIs para aplicaciones utilizando *Interacciones Naturales*. Su

propósito es formar un API estándar que permita la comunicación entre sensores de audio y visión, y el software que permita analizar dicha data de audio y video y comprenderla.

El API estándar del OpenNI permite a los desarrolladores de aplicaciones de interacción natural rastrear imágenes utilizando tipos de datos (data types) que son calculados de la información obtenida de un sensor, por ejemplo, la representación completa de un cuerpo, representación de la localización de una mano, un arreglo de píxeles en un mapa de profundidad, etc.

A continuación, se muestra un gráfico que representa el concepto de funcionamiento del OpenNI:

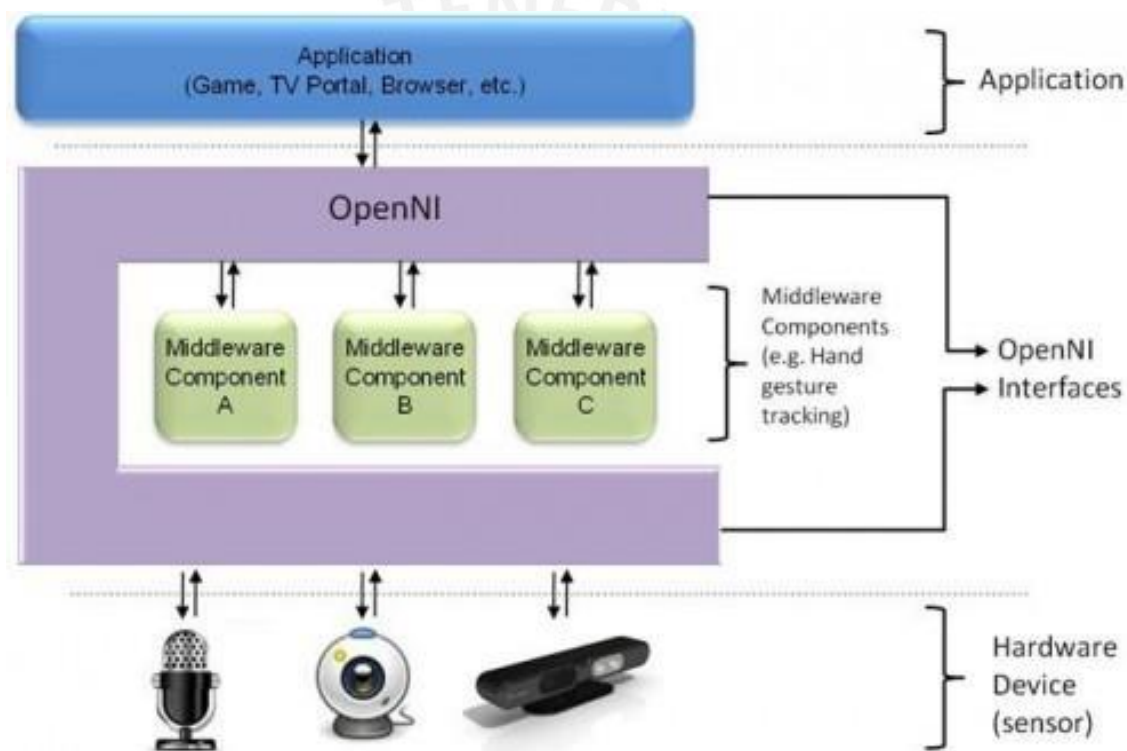


Figura 2.4. Diagrama de funcionamiento del Openni [14]

La parte de arriba representa el software que implementa aplicaciones de interacción natural sobre el Openni. El medio representa al Openni, que proporciona interfaces de comunicación que interactúan con el sensor y los componentes del middleware, analizando la data del sensor. La parte de abajo representa el hardware usado para capturar la información de video y audio.

Componentes y funciones del OpenNI

A continuación, se describirá los componentes y funciones más relevantes de la librería OpenNI.

Módulos

El API del Openni permite que varios componentes sean registrados en las librerías del Openni. A estos componentes se les refiere como módulos, y son usados para producir y procesar la data del sensor.

Estos módulos son:

- Sensor 3D
- Cámara RGB
- Cámara IR (infrarroja)
- Dispositivo de audio (micrófono o arreglo de micrófonos)

Componentes de Middleware:

- **Middleware de análisis de cuerpo:** componente de software que procesa la data del sensor y genera información relacionada al cuerpo (articulaciones, orientación, centro de masa, etc.).
- **Middleware de análisis de punto de la mano:** componente de software que procesa la data del sensor y genera la localización del punto que representa la mano (ver figura 2.5).

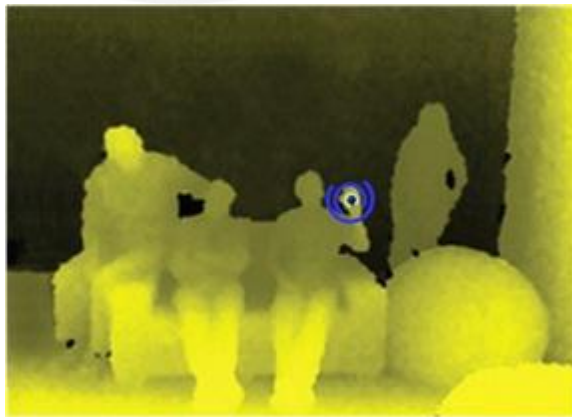


Figura 2.5. Detección del punto de la mano [14]

- **Middleware de detección de gestos:** componente de software que identifica gestos predefinidos (por ejemplo, saludar con la mano) y alerta a la aplicación.
- **Middleware de análisis de escena:** componente de software que analiza la escena y proporciona información como:
 - Separación entre los objetos o personas y el fondo.
 - Coordenadas
 - La identificación individual de personas u objetos en la escena.

Nodos de producción

Los nodos de producción son un conjunto de componentes que tienen como función la creación de un tipo de data específico. Los nodos son independientes unos de otros, y están ordenados en niveles. Mientras más bajo sea el nivel del nodo trabajará con un nivel de data de menor nivel (por ejemplo, la data que llega del Kinect); en cambio, un nodo de mayor nivel trabajará data de mayor nivel (por ejemplo, la data para generar una imagen de profundidad). Los nodos siempre usan otros nodos que representan un nivel menor de data para producir un nivel de data más alto.

Por ejemplo, si se quiere rastrear el movimiento de una persona en una imagen 3D, se necesita un nodo de producción que proporcione data del cuerpo para la aplicación (**generación de usuario**). Para esto se necesitará otro nodo (**generador de profundidad**) que le proporcionará la data al primero. La data del **generador de profundidad** proviene del sensor de profundidad.

Tipos de nodos de producción

Cada nodo de producción del Openni pertenece a uno de los siguientes tipos:

Nodos de producción relacionados al sensor

- **Dispositivo:** Nodo que representa a un dispositivo físico (sensor de profundidad, cámara RGB, etc.). Su propósito es permitir la configuración del dispositivo.
- **Generador de profundidad:** Nodo que genera mapas de profundidad.
- **Generador de imagen:** Nodo que genera mapas de imagen de color.
- **Generador IR:** Nodo que genera mapas de imagen infrarroja.
- **Generador de audio:** Nodo que genera una trama de audio.

Nodos de producción relacionados a Middleware

- **Generador de alerta de gestos:** Llama a la aplicación cuando determinados gestos son detectados (por ejemplo, saludar con la mano).
- **Analizador de la escena:** Analiza la escena, separa las personas u objetos del fondo y detecta el plano del suelo. Proporciona un mapa de profundidad, donde cada pixel tiene una etiqueta que indica si representa una persona u objeto, o si es parte del fondo.
- **Generador del punto de mano:** Este nodo llama a la aplicación cuando el punto que representa la mano es detectado, y cuando este punto cambia de posición cuando está siendo rastreado.
- **Generador de usuario:** Genera una representación de un cuerpo en una escena 3D.

Para poder grabar, también están los siguientes tipos de nodos:

- **Grabador:** Implementa grabación de data.
- **Reproductor:** Lee data de la grabación y la reproduce.
- **Codec:** Comprime y descomprime data en la grabación

Cadenas de producción

En la sección “Nodos de producción” se presentó un ejemplo relacionando la dependencia del Generador de usuario con el generador de profundidad, y este a su vez con la data extraída del sensor. Esto es a lo que se llama cadena de producción.

Normalmente, en una aplicación solo importa el nodo superior de una cadena de producción. Openni permite a la aplicación usar un solo nodo, sin tener conocimiento de los otros nodos de la cadena que están debajo de este. Para ajustes avanzados, hay una opción para acceder a esta cadena y modificar los otros nodos.

Capacidades

Las capacidades son extensiones opcionales del Openni API, y proporcionan funcionalidades adicionales. Openni se produce con un juego específico de capacidades, con la opción de adicionar capacidades en el futuro.

Las capacidades admitidas son las siguientes:

- **Vista alternativa:** Permite a cualquier generador de mapa (profundidad, imagen, IR) transformar su data, de tal forma que pareciera que el sensor está ubicado en otra posición.
- **Recortar (Cropping):** Permite a un generador de mapa generar una sección del área de la trama en vez de la trama completa, haciendo que el tamaño del mapa generado sea reducido para ajustarse a una menor resolución (menos pixeles).
- **Sincronización de trama:** Permite a dos sensores que producen tramas de datos (por ejemplo, profundidad e imagen) sincronizar sus tramas para que lleguen al mismo tiempo.
- **Espejo:** Refleja invertir la trama producida por el generador; por lo que, por ejemplo, la mano derecha aparecerá como la izquierda en la imagen.
- **Detección de pose:** Permite a un generador de usuario reconocer cuando un usuario se encuentra en una posición específica.
- **Esqueleto:** Permite a un generador de usuario producir la data del esqueleto de un usuario (ver Figura 2.6). Esta data incluye la localización de las articulaciones, rastrear la posición del esqueleto y capacidades de calibración del usuario.

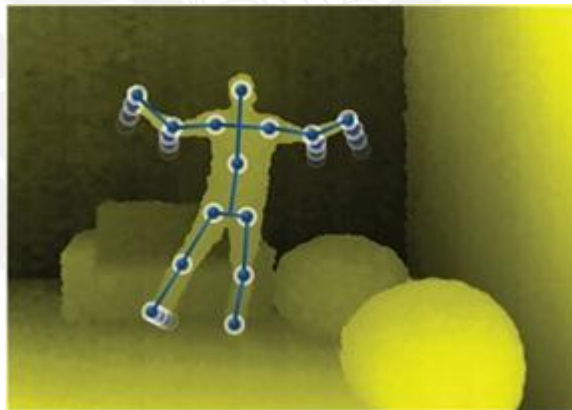


Figura 2.6. Generación de esqueleto [14]

- **Posición del usuario:** Permite a un generador de profundidad optimizar el mapa de profundidad para un área específica de la escena.
- **Estado de error:** Permite a un nodo reportar estado de error, lo que significa que el nodo podría no estar funcionando bien.

Generar y leer data

Generar data

Los nodos de producción generan alguna trama de información, ya sea de video o audio, son llamados generadores. Cuando son creados, no generan data inmediatamente, para así permitir a la aplicación poner la configuración requerida. La función **xn::Generator::StartGenerating()** es usada para comenzar la generación. Para detener la generación de data sin destruir el nodo, se puede usar la función **xn::Generator::StopGenerating**.

Leer Data

El generador siempre está recibiendo nueva data; sin embargo, la aplicación puede estar aun usando la data anterior. Es por esto que todo generador debe poder almacenar nueva data, hasta que se le requiera que actualice la data. Para esto se puede usar la función **UpdateData**. Si se quiere actualizar la data apenas haya nueva data disponible, se puede usar la función **xn::Generator::WaitAndUpdateData()**.

Cuando se tiene varios nodos, se puede usar las siguientes funciones:

- **xn::Context::WaitAnyUpdateAll()**: Espera que cualquier nodo tenga nueva data, luego actualiza todos los nodos.
- **xn::Context::WaitOneUpdateAll()**: Espera un nodo en específico tenga nueva data, luego actualiza todos los nodos.
- **xn::Context::WaitNoneUpdateAll()**: No espera nada. Actualiza inmediatamente todos los nodos.
- **xn::Context::WaitAndUpdateAll()**: Espera que todos los nodos tengan nueva data, luego actualiza todos los nodos.

Nodos simulados

Son usados usualmente por los nodos de reproducción para simular nodos reales cuando se lee data de una grabación.

Grabado

Permite captura total de la data y la habilidad de reproducirlo luego, permitiendo que la aplicación pueda simular la misma situación.

Estado de error de nodo de producción

Esta capacidad permite al nodo de producción cambiar su estado de error si un error se produce. Un nodo que no tiene implementada esta capacidad siempre tiene su estado de error en "Ok" [14].

2.4.2. NITE

NITE es un *Middleware* que procesa la data extraída del Kinect. La función más común que se le da es el control de dispositivos usando los movimientos de extremidades del cuerpo, como por ejemplo las manos.

El NITE también tiene la capacidad de generar un esqueleto virtual, identificando las partes del cuerpo del usuario. Esta es la función del NITE que será de interés.

- **Segmentación del Usuario**

Descripción General

La segmentación del usuario consiste en rastrear e identificar usuarios dentro del rango de visión. A cada usuario dentro de la escena se le asigna un *ID*, esto dándole a cada pixel de la imagen del usuario el ID de este. Esto es necesario para la generación del esqueleto virtual, por lo que cualquier error generado aquí se manifestará en errores en la detección de posición.

Problemas conocidos

- El ambiente donde se encuentre el usuario puede ser una causa de errores ya sea por los objetos presentes o por otras personas. Por ejemplo:
 - Cuando un usuario se mueve muy cerca de algún objeto, por ejemplo, sillas.
 - Cuando dos usuarios se tocan.
- Cambiar el sensor de posición mientras la segmentación se está llevando a cabo.

- Cuando hay oclusiones o cuando el usuario deja el campo de visión, ya que su ID se puede perder o se le puede asignar incorrectamente a otro usuario.
 - Si el usuario no es visible por más de 10 segundos, se le considera perdido, por lo que se es detectado de nuevo se le asignará un nuevo ID, o lo ya mencionado de que se le puede asignar incorrectamente su ID a otro usuario.
- **Rastreo de Esqueleto (Skeleton Tracking)**

Condiciones Básicas

- La parte superior del cuerpo del usuario está en su mayoría dentro del área de visión. Tratar de evitar oclusiones de objetos u otros usuarios.
- La distancia ideal es alrededor de 2.5 m.
- Para mejores resultados, el usuario debe usar ropa ceñida al cuerpo. Además, el usuario no debe tener el cabello largo porque puede bajar la calidad del rastreo.

Calibración

La calibración es usada para ajustar el esqueleto a las proporciones del cuerpo del usuario debido a que es necesario para inicializar el rastreo de posición. En caso no se calibre, entonces no se va a obtener ninguna información del sensor. La calibración demora aproximadamente menos de 3 segundos y para que esta se realice, el usuario debe mirar al sensor como se muestra en la figura 2.7:

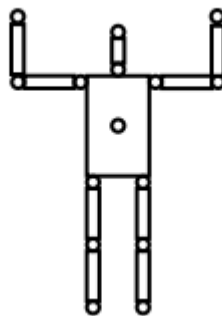


Figura 2.7. Pose de calibración [15]

- La cabeza y los brazos del usuario deben ser completamente visibles durante la calibración.
- El usuario no debe doblar las rodillas ni el torso durante la calibración.

Salida del Rastreo del Esqueleto

El **API** devuelve posiciones y orientaciones de las articulaciones del esqueleto. Hay que notar que las longitudes de los segmentos del cuerpo pueden variar durante el rastreo.

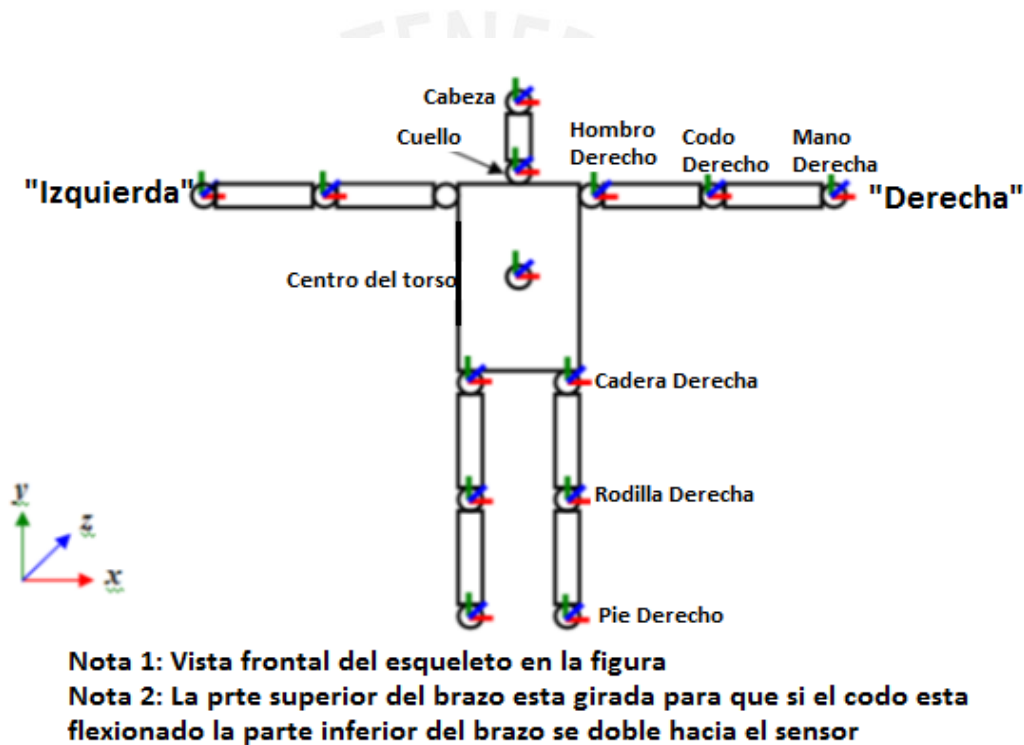


Figura 2.8. Esqueleto virtual del NITE. Modificada de [15]

La Figura 2.8 ilustra el sistema de coordenadas y la representación del esqueleto cuando está mirando al sensor. A continuación, se explicará la definición de "Izquierda" y "Derecha" de este esqueleto.

La imagen de profundidad generada con la información extraída del sensor, actuará como un espejo cuando se esté al frente del sensor. Por ejemplo, cuando se mueve el brazo

derecho, se moverá el brazo del lado “Derecho” de la imagen de profundidad, que para la imagen sería su brazo izquierdo, al igual que nuestro reflejo en un espejo.

Las posiciones y orientaciones de las articulaciones están dadas en un sistema de coordenadas del mundo real, siendo el origen del sistema el propio sensor. +X puntos a la derecha, +Y puntos arriba, +Z puntos a mayor profundidad.

Las posiciones de las articulaciones (X, Y y Z) se miden en milímetros.

Problemas Conocidos

- El rastreo del brazo es menos estable cuando este está cerca de otras partes del cuerpo, especialmente el torso. Si los brazos están cerca del torso o cruzados pueden mezclarse.
- El rastreo de la pierna es aún algo inestable y ruidoso. Funciona mejor cuando el usuario está de pie con las piernas separadas. Movimientos rápidos o agacharse pueden hacer que falle el rastreo.
- Rastreo de pose también puede volverse algo inestable si la cabeza no es visible.
- Brazos y piernas extremadamente estirados (cerca del máximo posible que puede el cuerpo) pueden perderse del rastreo.
- En general, movimientos rápidos pueden causar fallas en el rastreo.

Recomendaciones

- Si el esqueleto se queda congelado en una pose errada, ponerse en una pose simple (brazos estirados y piernas separados) debería arreglar el problema.
- En algunos casos, el rastreo en general podría estar mal. Recalibrar al usuario podría arreglar esto [15].

2.4.3. Kinect for Windows SDK

El Kinect for Windows SDK (en adelante será referido solo como SDK) procesa la data obtenida del sensor Kinect, permitiendo a los usuarios realizar aplicaciones como detección de gestos, generar imágenes de color y de profundidad, generar un esqueleto virtual, reconocimiento de palabras usando la data de audio, entre otros.

El SDK incluye *drivers* para el Kinect que permiten controlar la cámara, el sensor de profundidad, el arreglo de micrófonos y el motor de inclinación del Kinect. La data pasa del sensor a una aplicación y viceversa como tramas de datos de los siguientes tipos:

- Trama de data de color
- Trama de data de profundidad
- Trama de data de audio

El siguiente diagrama muestra la forma de trabajar del SDK (ver figura 2.9):

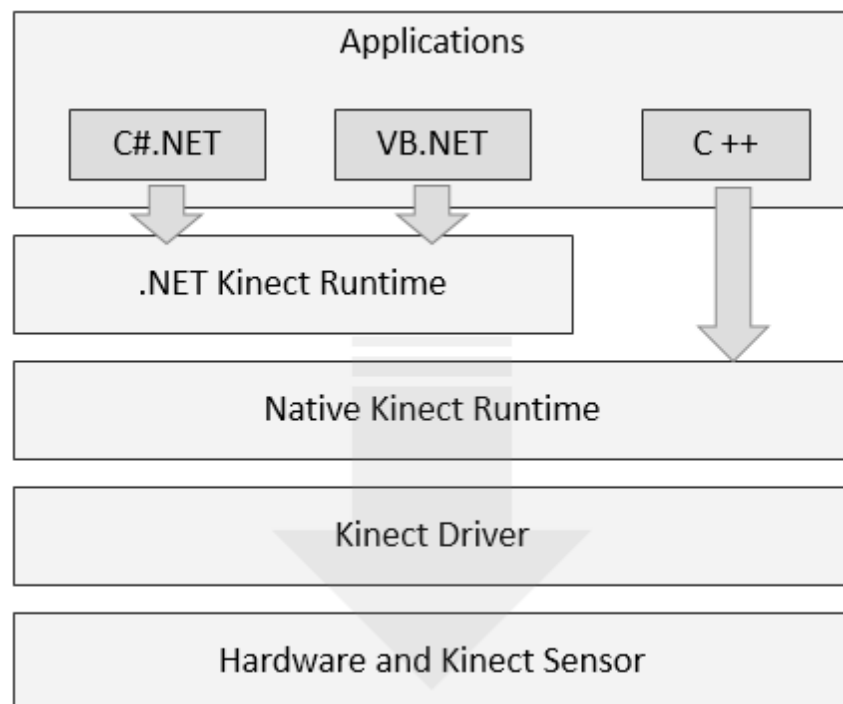


Figura 2.9. Arquitectura del SDK [8]

Características del SDK

El SDK proporciona librerías que permiten interactuar con los diferentes componentes del Kinect, como su cámara, sensores, micrófonos y motor de inclinación; permitiendo así desarrollar aplicaciones que detecten gestos basándose en el movimiento del cuerpo o aplicaciones que realicen el reconocimiento de palabras. Las operaciones que se pueden realizar con el SDK son las siguientes:

- Capturar y procesar la data de la trama de imagen de color
- Procesar la data de la trama de la imagen de profundidad
- Capturar la trama infrarroja
- Rastrear Un esqueleto virtual y movimientos de las articulaciones
- Reconocimiento de gestos
- Capturar la trama de audio
- Permitir reconocimiento de palabras
- Ajustar el ángulo de inclinación del Kinect
- Obtener data del acelerómetro

Capturar y procesar la data de la trama de imagen de color

El SDK puede capturar la trama de imagen a diferentes resoluciones. La data de color, a la cual nos estamos refiriendo como trama de imagen, es como una sucesión de imágenes de color enviadas por el sensor. El SDK es capaz de enviar una trama de imagen del sensor cuando se le solicita. Hay que tomar en cuenta que la resolución afecta las tramas por segundo que envía el Kinect, ya que con una resolución de 640 x 480 envía 30 tramas por segundo, mientras que a una resolución de 1280 x 960 solo envía 10 tramas por segundo.

Procesar la data de la trama de la imagen de profundidad

En la trama de profundidad, cada pixel en la data representa la distancia entre el objeto y el sensor. Los APIs del SDK pueden soportar data de tramas de imágenes de profundidad de resoluciones de 640 x 480, 320 x 240, y 80 x 60 pixeles.

Capturar la trama infrarroja

Se puede capturar data en lugares con poca iluminación, capturando la trama infrarroja del Kinect. Hay que tener en cuenta que no se puede leer data de color e infrarrojas al

mismo tiempo, aunque si data de profundidad e infrarroja. Esto es porque la trama infrarroja es capturada como parte de un formato de la imagen de color.

Rastrear Un esqueleto virtual y movimientos de las articulaciones

El SDK provee rastreo del esqueleto virtual de una persona (ver figura 2.10), pudiendo rastrear 20 articulaciones en un solo esqueleto (a diferencia del OpenNI que solo puede rastrear hasta 14). Puede rastrear hasta 6 esqueletos al mismo tiempo, pero solo puede detectar las posiciones de las articulaciones (X, Y y Z) de dos, las cuales se miden en metros. También es posible para el SDK rastrear el esqueleto de una persona sentada, pero en este modo solo detectará 10 articulaciones, que serán la de la parte superior del cuerpo.

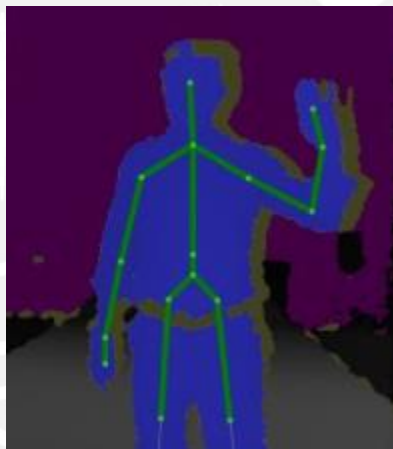


Figura 2.10. Esqueleto virtual obtenido con SDK [8]

Reconocimiento de gestos

Se debe entender por gesto a una acción realizada por una persona para expresar algo. El SDK no provee directamente un API para reconocer y tratar con los gestos de una persona. Sin embargo, usando en conjunto el rastreo del esqueleto y el procesamiento de la data de profundidad, es posible para un usuario desarrollar su propio API para reconocimiento de gestos que pueda interactuar con una aplicación.

Capturar la trama de audio

El SDK provee de capacidades de procesamiento de audio de alta calidad. Además, aparte de permitir capturar data de audio, es posible procesar esta data, suprimiendo el ruido y cancelando los ecos.

Permitir reconocimiento de palabras

Se puede hacer uso de los micrófonos del Kinect y los APIs de reconocimiento de palabras de Windows para reconocer la voz y desarrollar aplicaciones con ello. Por ejemplo, se puede desarrollar una serie de comandos de voz para controlar diferentes aplicaciones.

Ajustar el ángulo de inclinación del Kinect

El SDK permite controlar el motor de inclinación del Kinect. Con esto, es posible ajustar el ángulo de visión del Kinect en el eje vertical acorde a la necesidad. El rango del ángulo de inclinación del Kinect es de -27° hasta 27° (ver figura 2.11).

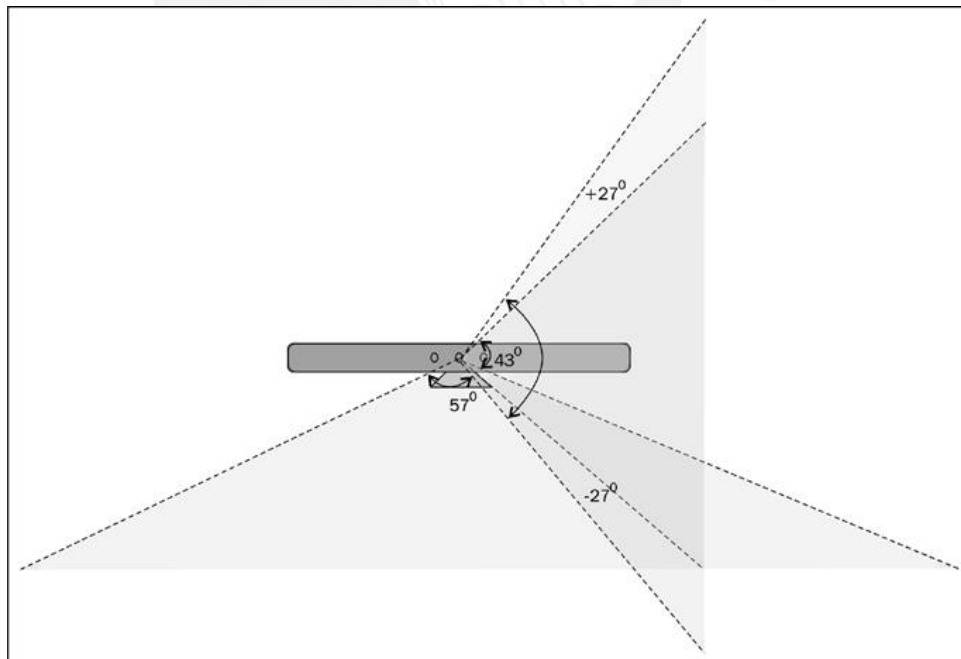


Figura 2.11. Ángulos de visión del Kinect [8]

Obtener data del acelerómetro

Para el Kinect, el ángulo de inclinación es relativo a la gravedad y no a su base, por lo que usa sus acelerómetros para controlar la rotación. El SDK provee APIs para leer la data de los acelerómetros del sensor, lo que permite, por ejemplo, detectar la orientación del sensor al leer esta data [8].



CAPÍTULO 3: ANÁLISIS Y COMPARACIÓN DE DATOS PARA EL ANÁLISIS DE MARCHA USANDO EL KINECT Y SEIS CÁMARAS VICON

3.1. Protocolo de adquisición

Para realizar las pruebas, primero fue necesario definir un protocolo de adquisición de datos, para ser utilizado en futuras investigaciones en este tema. El protocolo que se va a definir a continuación se realizó tomando en cuenta las recomendaciones de médicos expertos en terapia física, las limitaciones que posee el sensor Kinect y problemas y recomendaciones que se señalan en los manuales de las librerías OpenNI, NITE y SDK.

3.1.1. Ambiente de trabajo

- **Dimensiones**

Como se puede ver en la figura 3.1, las pruebas se realizaron en un ambiente de 7.7 x 6.8 metros, donde fue instalado el sistema de captura de movimiento en 3D Vicon. Es necesario un espacio de estas dimensiones, ya que las cámaras Pulnix, como cualquier otra cámara, necesitan estar a una distancia apropiada de la superficie de trabajo para poder enfocarla correctamente.

La superficie de trabajo es un área de 3 x 1 m², donde los voluntarios realizaron las pruebas de marcha tanto con el sistema Vicon como con el Kinect. Cuando se realizó la captura de movimiento de los voluntarios con el sistema Vicon, los voluntarios caminaban casi todo el largo del área de trabajo, la cual era de 6 metros aproximadamente; sin embargo, la data más confiable era la capturada dentro de la superficie de trabajo. Por otro lado, cuando se realizó la captura de movimiento de los voluntarios con el Kinect, al ser su rango de visión más reducido que el del sistema Vicon (va desde 1.2 hasta 3.5 metros, dando un rango total de trabajo de 2.3 metros), los voluntarios caminaban los 3 metros de largo de la superficie de trabajo, pero solo se registraba información hasta que se acercaran a 1.2 metros del Kinect.

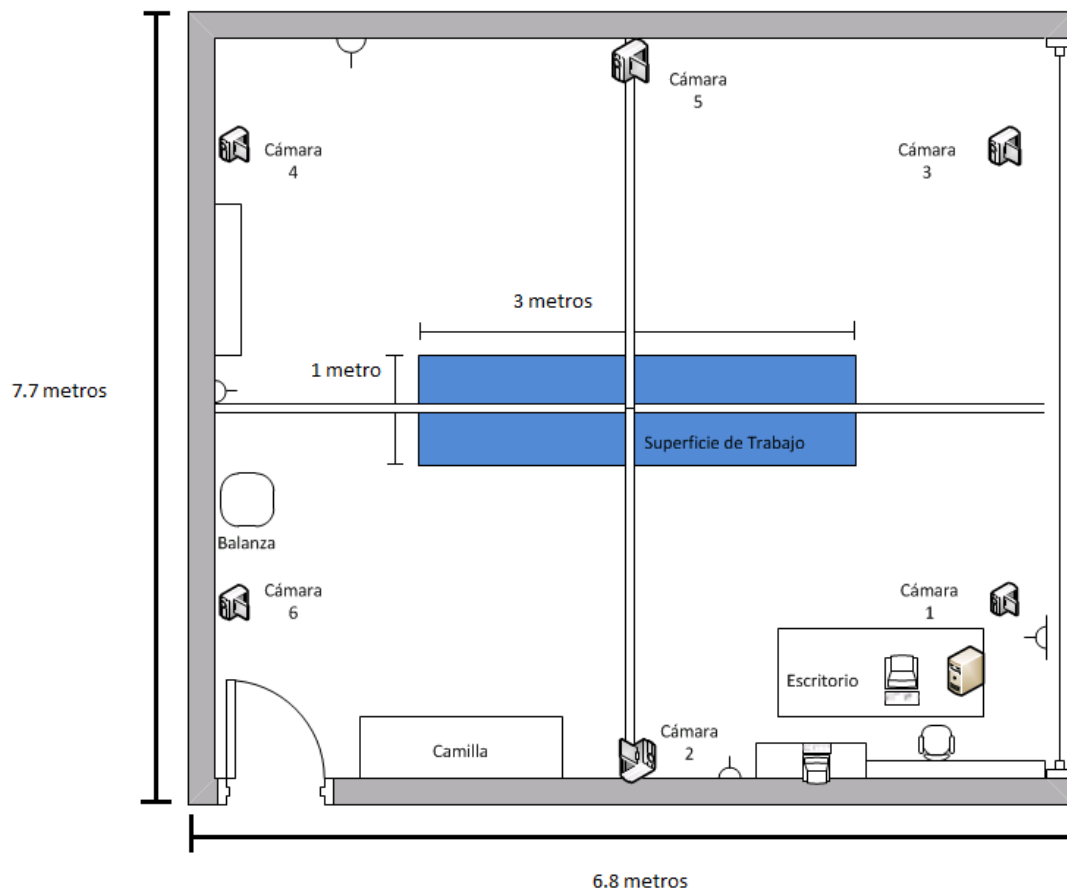


Figura 3.1. Dimensiones del ambiente de trabajo [20]

- **Condiciones del ambiente**

Las pruebas se realizaron en un espacio despejado (ver **figura 3.2**), para evitar incomodidades a los voluntarios al realizar las pruebas y, en el caso del Kinect, para evitar errores en el rastreo del esqueleto virtual (como se mencionó en la descripción del NITE).

El espacio donde se trabajará también debe tener una iluminación leve para no introducir errores en el rastreo del usuario, ya que tanto el sistema Vicon como el Kinect no trabajan bien en lugares con mucha iluminación como ya se mencionó anteriormente.



Figura 3.2. Condiciones del área de trabajo [20]

3.1.2. Ubicación del Kinect

El Kinect estará ubicado sobre una mesa pequeña de 75 cm de altura, donde se ajustará el ángulo de inclinación de este para poder dar preferencia a la captura de las articulaciones de las piernas, pero sin estar demasiado inclinado como para no captar la parte superior del cuerpo, ya que como se mencionó anteriormente, es necesario que la parte superior del cuerpo de los voluntarios sea visible para que el rastreo se realice. El ángulo de inclinación variará dependiendo de la altura de los voluntarios.

Además, el Kinect estará cerca del borde de la mesa para que así la luz infrarroja que el Kinect emite no se vea obstruida por la mesa (en la imagen obtenida del Kinect se observa zonas borrosas en la parte inferior si esto pasa).

3.1.3. Características físicas de los voluntarios

Como ya se mencionó antes, se busca que los voluntarios sean personas con una forma

de caminar dentro de lo considerado como normal por médicos especialistas en rehabilitación; por ello, es necesario que se encuentren dentro de determinados rangos de edad, peso y altura, ya que estos factores afectan la marcha de una persona.

Se ha considerado los siguientes parámetros de inclusión para los voluntarios:

- Edad: entre 17 y 35 años (adulto joven)
- Talla:
 - talla promedio de mujeres en el Perú es 1.57 metros
 - talla promedio de varones en el Perú es 1.69 metros
 - para el estudio se considera valores en mujeres entre 1.50 m – 1.57 metros
 - para el estudio se considera valores en varones entre 1.60 -1.80 metros
- Índice de masa corporal según la OMS:
 - IMC delgadez leve: 17 -18.49
 - IMC normal: 18.5 – 24.99
 - IMC sobrepeso leve: 25 – 27.49
 - Para el estudio se considera IMC entre 17 – 27.49

3.1.4. Equipos usados

Para el desarrollo de este estudio se hizo uso del sistema de captura de movimiento Vicon y del Kinect, que ya fueron descritos anteriormente. Para poder utilizar el software Workstation que permite manejar el sistema de captura de movimiento en 3D Vicon y procesar sus datos, se usó una computadora personal (PC) Intel Core 2 Duo con sistema operativo Windows XP, sistema operativo adecuado para el modelo Vicon 460. En esta misma PC, se instaló el sistema operativo basado en Linux Ubuntu versión 12.10, en el cuál fueron instaladas las librerías Openni y Nite, usadas para procesar la información del Kinect, donde los códigos usados para el estudio fueron ejecutados desde el terminal de Ubuntu. Aparte de esto también fue necesario usar una PC Intel Core i7 con sistema operativo Windows 7 para poder instalar los siguientes programas: “Kinect for Windows SDK v1.8” y “Kinect for Windows Developer Toolkit Browser v1.8”, ambos software libre proporcionados por Windows. Estos softwares son necesarios para desarrollar aplicaciones con el SDK; además solo pueden ser instalados en sistemas operativos Windows de la versión 7 en adelante.

Para la evaluación clínica realizada por un médico especialista en medicina de rehabilitación, rol que fue desempeñado por la Dra. Jesica Pajaya, se usó un calibrador (vernier) de 400mm-16" quijada 100mm W PAVICH y un centímetro de longitud de 150 centímetros para la medición de la altura de los voluntarios y las dimensiones de sus piernas.

3.1.5. Pruebas

- **Preparación del sujeto**

Primero, se les pide a los voluntarios que se cambien ropa pegada al cuerpo con la que se cuenta para las pruebas, y si tienen cabello largo, que se lo amarren. Lo siguiente que se hace es tomar datos clínicos de los voluntarios, para que después se le realice la evaluación médica a cargo del médico especialista para determinar alguna anomalía en su marcha. Si no se encuentra ninguna anomalía se procede con la siguiente etapa que consiste en la toma de datos de los voluntarios con ambos equipos (Kinect y sistema Vicon).

Los datos que se toman son los siguientes:

- Estatura
- Peso
- Longitud de pierna derecha e izquierda
- Ancho de rodilla derecha e izquierda
- Ancho de tobillo derecho e izquierdo

- **Evaluación Clínica**

La evaluación clínica consiste en una serie de pruebas realizadas por el médico especialista para determinar si la existe anomalías en la marcha de los voluntarios y si su marcha se encuentra dentro de los parámetros considerados normales. Las condiciones que los voluntarios deben cumplir para que su forma de caminar sea considerada normal y libre de anomalías son las siguientes:

- Ausencia de acortamiento muscular
- Ausencia de alteraciones torsionales

- Ausencia de alteraciones angulares
- Fuerza muscular normal
- Rango articular normal

- **Pruebas en Vicon**

Al voluntario se le colocaron 16 marcadores en las piernas en posiciones definidas en los manuales proporcionados por la empresa Vicon (ver figura 3.3) y colocados por un médico especialista, como se mencionó anteriormente esto fue realizado por la Dra. Jessica Pajaya. Luego se hizo caminar al voluntario sobre un volumen de trabajo (ver figura 3.2) para registrar 3 pasos completos por el sistema Vicon, labor que fue realizada por un especialista en el manejo de este sistema, en este caso fue la Ing. Rocío Callupe, además de generar los archivos .csv con la data obtenida. Al final, se registró 3 veces la marcha de cada voluntario, siendo un total de 40 voluntarios.

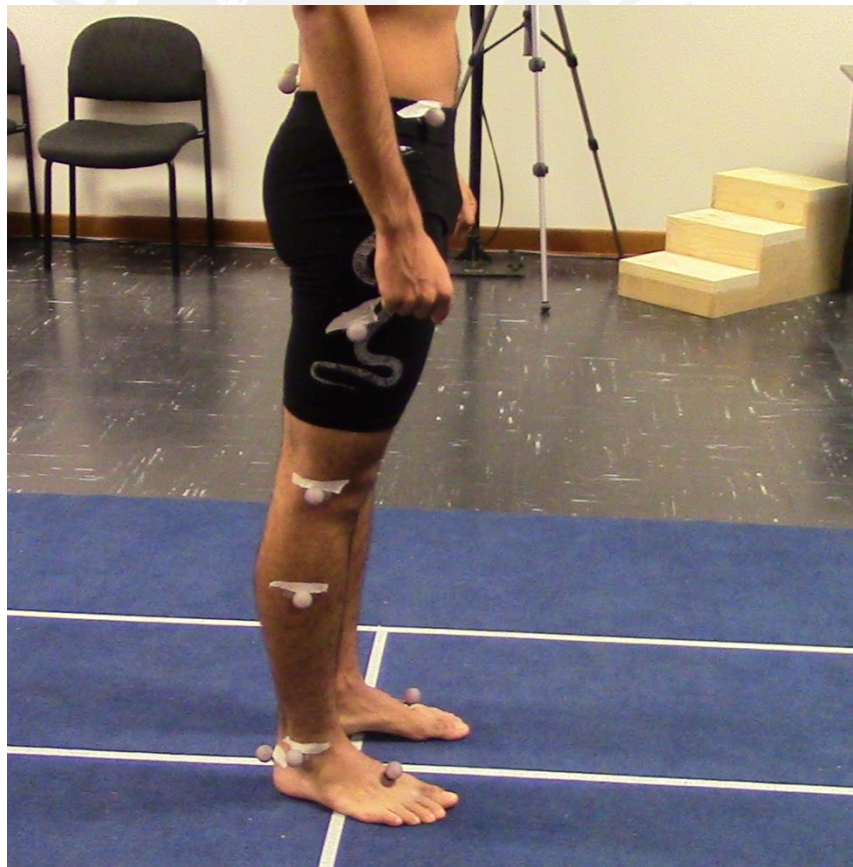


Figura 3.3. Colocación de 16 marcadores reflectantes de 25mm de diámetro [20]

- **Pruebas en Kinect**

Estas pruebas se realizaron haciendo uso de dos librerías diferentes: OpenNI, en conjunto con NITE, y SDK. Estas librerías fueron escogidas ya que en los estudios que se revisó (ver estado del arte), la mayoría usaban una de estas dos librerías. Así, se podrá saber con cuál de estas dos librerías se obtiene mejores resultados para el análisis de marcha.

En este estudio se medirá la cinemática de la rodilla. Se escogió la rodilla porque lo que se busca es analizar la exactitud y precisión del Kinect para el análisis de la marcha, en un caso de análisis del ángulo de la articulación de la rodilla; por ello, las partes del cuerpo que son de interés son las piernas, ya que es en ellas donde se presentan las anomalías al caminar.

Las articulaciones que se pueden encontrar en las piernas (sin contar la de los dedos) son las de la cadera, las rodillas y los tobillos. El esqueleto virtual representa cada articulación y algunas partes del cuerpo como un punto (ver fig. 9 11 y 13), por lo que se pueden encontrar 3 puntos en cada pierna: cadera, rodilla y pie (en el Microsoft SDK son 4, ya que el tobillo y el resto del pie son puntos distintos). Para obtener la cinemática de una articulación se requiere de las coordenadas de tres puntos conocidos para formar dos vectores (esto se explicará mejor en la siguiente sección “Análisis de los Resultados”).

La cadera fue descartada porque los puntos necesarios para obtener la variación del ángulo de la cadera son la rodilla, la cadera y puntos del tronco. Sin embargo, el tronco es representado en el esqueleto virtual por un solo punto, lo que haría inexacta la información extraída de la cadera.

El tobillo fue descartado ya que como se puede observar en la figura 9 el NITE representa a todo el pie como un solo punto (a diferencia del SDK), y los puntos que se necesitan para la medición del tobillo son rodilla, tobillo y pie; es decir, el NITE no hace distinción del tobillo del resto del pie, faltando así un punto, impidiendo así la medición del tobillo.

Para la medición de la rodilla se necesita tener los puntos de la cadera, la rodilla y el tobillo. El tobillo y el pie son un solo punto en el NITE, pero este punto (el del pie) también se puede usar, por lo que se tendrían 3 puntos, es por esto que se medirá la rodilla.

Las pruebas (con el sistema Vicon y con el Kinect) no se hicieron en simultáneo ya que ambos sistemas trabajan con luz infrarroja, lo que convertiría a cada sistema en una fuente de ruido para el otro.

Las pruebas se realizaron en el siguiente orden por cada uno de los métodos (Openni y SDK) con un voluntario a la vez:

- Primero, el voluntario se paró en frente del Kinect a una distancia de 3.5 metros con los brazos separados y se quedó inmóvil hasta que la calibración se completó.
- Segundo, el voluntario caminó hacia el Kinect en línea recta de la forma más natural posible, sin cruzar los brazos en ningún momento, pero sin caminar demasiado rápido, ya que esto podría haber creado errores en el rastreo del esqueleto. Es importante que al menos se registren tres pasos completos, ya que esto significaría que hay un ciclo de paso completo en cada pierna.
- Tercero, el procedimiento se repitió 4 veces más, ya que una sola prueba no brinda información suficiente para conocer la exactitud y precisión del Kinect, especialmente si tomamos en cuenta su corto rango de trabajo (2.3 metros).

3.2. Resultados

Las pruebas se realizaron con tres voluntarios, como fue descrito anteriormente, a los que en adelante se referirá a ellos como voluntario 1, voluntario 2 y voluntario 3. Para el procesamiento de los resultados se siguió los siguientes pasos:

- Se obtuvo los datos de los voluntarios mediante el Kinect por ambas librerías (OpenNI y SDK) descargando ejemplos de rastreo de esqueleto (Skeleton Tracking) proporcionados por los creadores de estas librerías (ambos en lenguaje C++) y se les realizó las modificaciones necesarias para poder extraer las coordenadas (x, y, z) de los puntos de interés: caderas, rodillas y tobillos en el caso de SDK y caderas, rodillas y pies en el caso de OpenNI (recordar que el OpenNI no hace distinción entre pie y tobillo). Los datos adquiridos mediante el sistema Vicon fueron procesados en el trabajo "*Sistema robótico tipo exoesqueleto para las extremidades inferiores que permita reproducir patrones de movimiento en el plano sagital para marcha normal o movimientos asistidos dirigidos conforme a procedimientos médicos*", realizado por Callupe [20], en la etapa de captura de movimiento de la marcha humana. Estos datos del movimiento de la marcha humana son los más óptimos por ser datos de una misma población a la de este trabajo por lo que se toma como referencia.
- Se calculó la variación del ángulo de la rodilla en la marcha de los voluntarios. Para esto, se restó el punto del tobillo (o pie en el caso del OpenNI) con el punto de la rodilla para formar el primer vector y el punto de la rodilla con el punto de la

cadena para formar el segundo vector; todo esto usando solo las coordenadas Y (que representa la altura) y Z (que representa la distancia al Kinect), mediante el programa Matlab. Esto debido a que con esas dos coordenadas se puede obtener la variación del ángulo de la rodilla en el plano sagital, que es el de interés. La fórmula usada para el cálculo del ángulo fue la siguiente:

$$\cos \emptyset = \frac{u \cdot v}{|u||v|} \quad (1)$$

- Para un mejor análisis y apreciación de las curvas obtenidas, se aplicó un filtro butterworth a la señal para disminuir el ruido presente. El filtro fue de segundo orden y, para hallar la frecuencia de muestreo, se obtuvo el espectro de las gráficas obtenidas para conocer la frecuencia de corte, la cual se variaba entre 6 y 9 Hz dependiendo de cada curva. A continuación, un ejemplo:

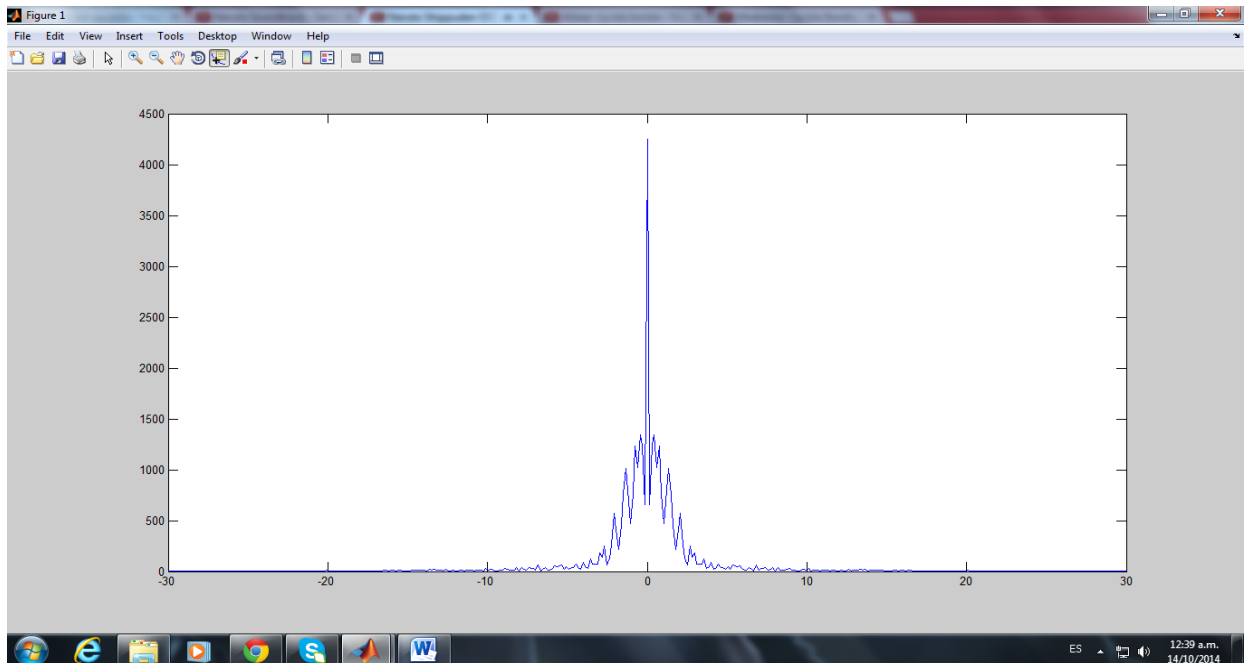


Figura 3.4. Frecuencia de corte: 6hz (de -6 a 6 hz).

- Manualmente se seleccionó las tramas de las gráficas obtenidas que se acercaran más a las curvas teóricas, tomando las tres mejores por cada pierna.
- Como el sistema Vicon trabaja al doble de frecuencia del Kinect y además de que cada voluntario realiza su ciclo de marcha en un intervalo de tiempo distinto, se realizó una estandarización de los valores obtenidos, realizando una interpolación cúbica de los datos del Kinect y del sistema Vicon para que todos los ciclos de marcha tuvieran el mismo número de muestras.
- Una vez realizadas las gráficas con los patrones de marcha, se procedió a calcular el **coeficiente de correlación** entre las curvas de movimiento de la rodilla resultantes de promediar las 40 curvas obtenidas con el sistema Vicon y las 3 mejores curvas obtenidas con Kinect. La fórmula para hallar el coeficiente de correlación es la siguiente:

$$r = \frac{N * \sum x * y - \sum x \sum y}{\sqrt{(N * \sum x^2 - (\sum x)^2) * (N * \sum y^2 - (\sum y)^2)}} \quad (2)$$

- Luego se procedió a calcular la precisión y exactitud. Para el cálculo de la precisión, se calculó **la desviación estándar de la media las tres gráficas del Kinect**, usando la fórmula de desviación estándar:

$$S = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}} \quad (3)$$

Finalmente, para el cálculo de la exactitud se realizó un cálculo de “**Cohen’s d**” **entre las curvas de movimiento de la rodilla resultantes de promediar las 40 curvas obtenidas con el sistema Vicon y las 3 mejores curvas obtenidas con Kinect**. La fórmula de “Cohen’s” d es la siguiente:

$$d = \frac{m_1 - m_2}{S_{pooled}} \quad (4)$$

Donde m_1 y m_2 son las medias y S_{pooled} es la desviación estándar promedio.

A continuación, se muestran los resultados obtenidos. En las gráficas 3.5 y 3.7 se pueden ver los resultados de la rodilla derecha obtenidos mediante SDK y OpenNI respectivamente, de la misma manera en las gráficas 3.9 y 3.11 para la rodilla izquierda. Las gráficas 3.6, 3.8, 3.10 y 3.12 muestran la correlación entre las curvas de movimiento obtenidas con Kinect y las obtenidas con Vicon, mientras las gráficas 3.13 y 3.14 muestran un diagrama de barras de la precisión y exactitud, respectivamente, obtenidas de cada rodilla y por ambos métodos. Todas las gráficas mencionadas anteriormente corresponden al voluntario 1, las correspondientes a los voluntarios 2 y 3 pueden ser vistas en los anexos.



3.2.1. Voluntario 1

- **Rodilla derecha**

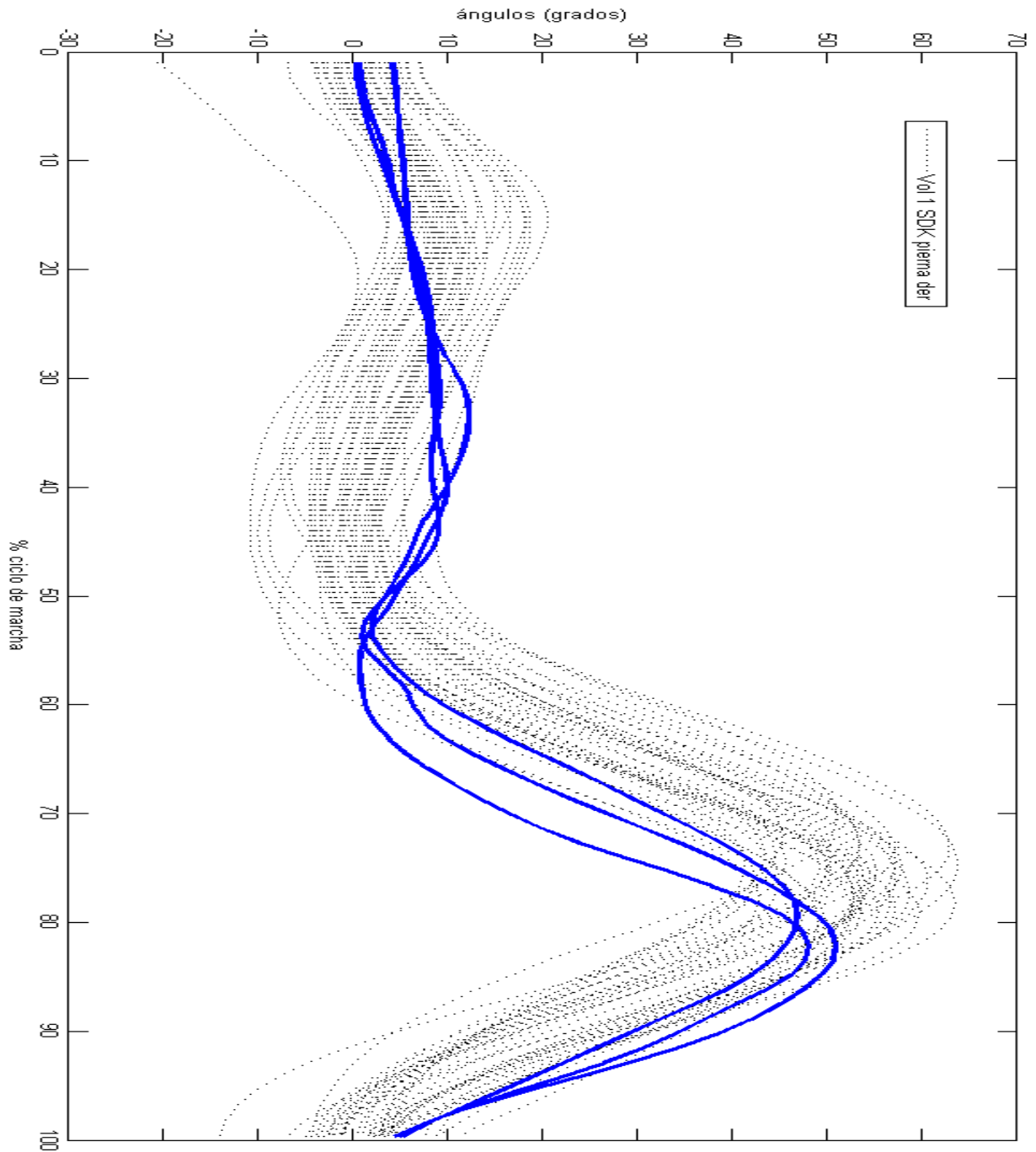


Figura 3.5. Curvas de movimiento (SDK) en la rodilla derecha del voluntario 1

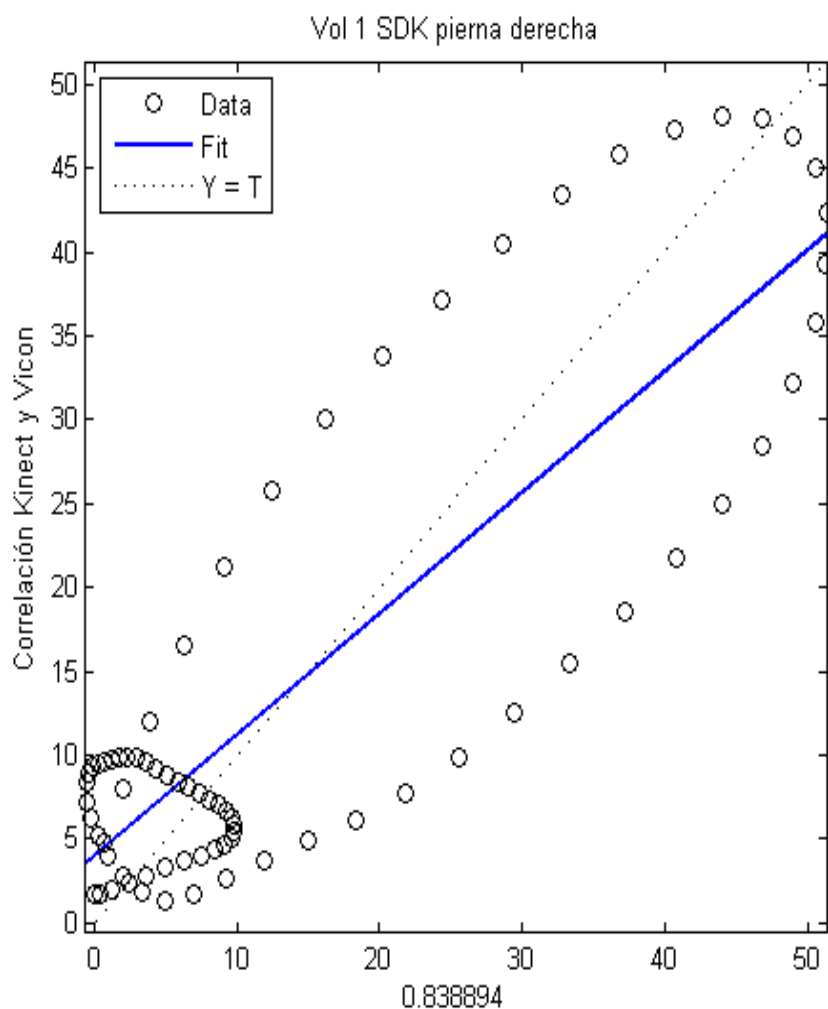


Figura 3.6. Correlación entre las curvas de movimiento la rodilla derecha del voluntario 1 obtenidas con Kinect (SDK) y Vicon

En la gráfica 3.5 se puede apreciar que la similitud que existe entre las curvas de movimiento de la pierna derecha obtenidas mediante el Kinect, haciendo uso de la librería Kinect SDK, y las obtenidas mediante el sistema de captura de movimiento Vicon. El índice de correlación obtenido del promedio de las curvas medidas mediante Kinect y el sistema Vicon, como se muestra en la figura 3.6, es de 0.838894, un valor superior a 0.8, por lo que se puede considerar que existe una buena semejanza entre estas dos gráficas (si las 2 curvas fueran idénticas, el índice de correlación sería 1).

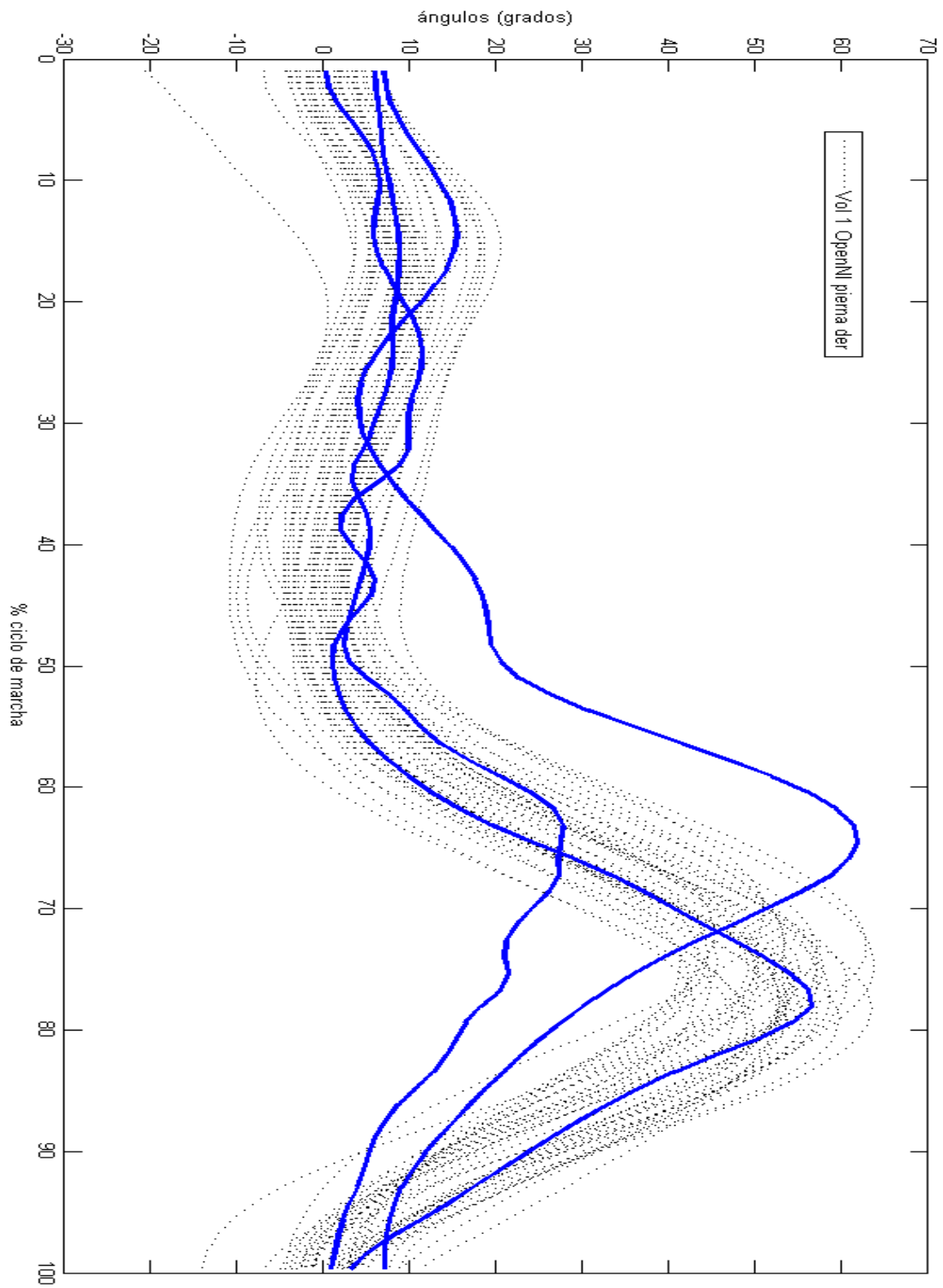


Figura 3.7. Curvas de movimiento (Openni) en la rodilla derecha del voluntario 1

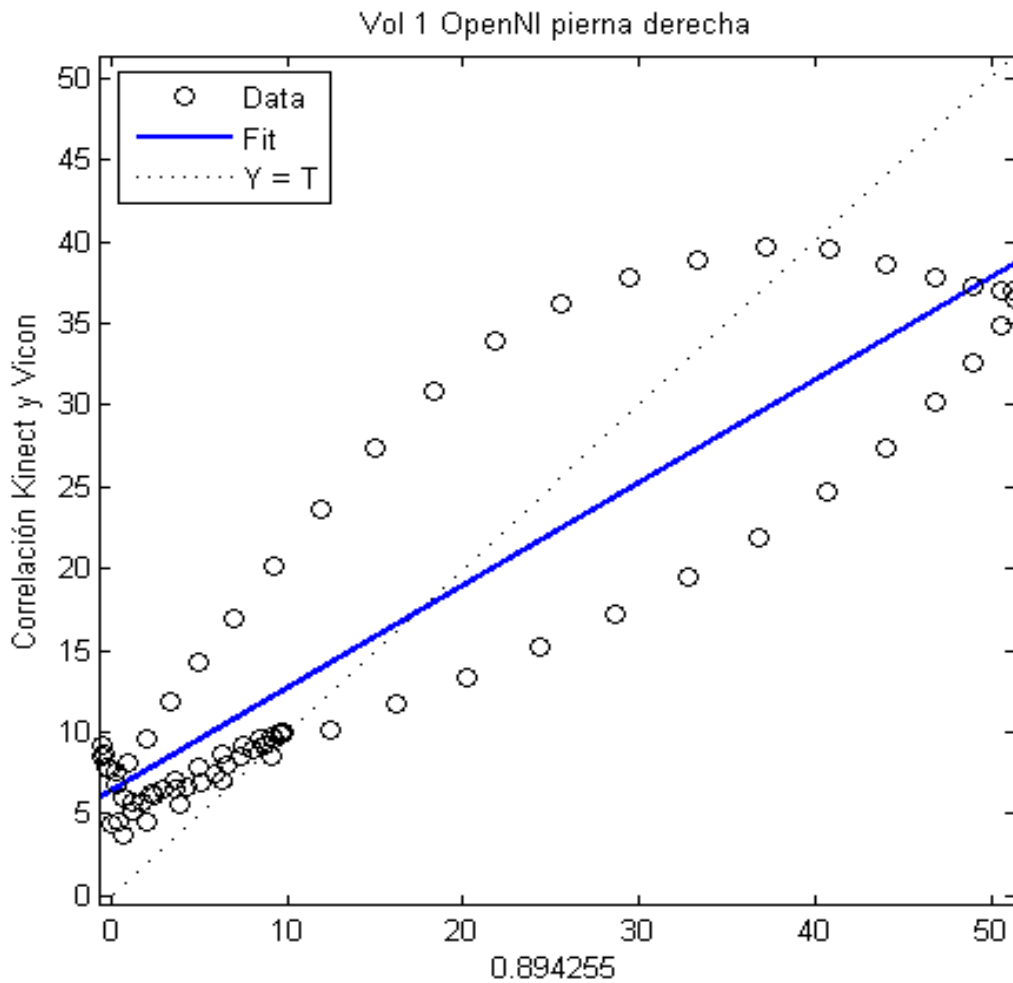


Figura 3.8. Correlación entre las curvas de movimiento la rodilla derecha del voluntario 1 obtenidas con Kinect (OpenNI) y Vicon

En la gráfica 3.7 se puede ver las curvas de movimiento de la pierna derecha obtenidas mediante Kinect, esta vez obtenidas con la librería OpenNI, y las obtenidas mediante Vicon. Se puede apreciar mayor diferencia, o mayor desfase, entre las curvas obtenidas con Kinect y las curvas obtenidas con Vicon, así como entre sí mismas. El coeficiente de correlación obtenido del promedio de las curvas medidas mediante Kinect y Vicon, mostrado en la figura 3.8, también es superior a 0.8 al igual que lo fue en la figura 3.6, por lo que se puede decir que existe similitud significativa entre estas.

- **Rodilla Izquierda**

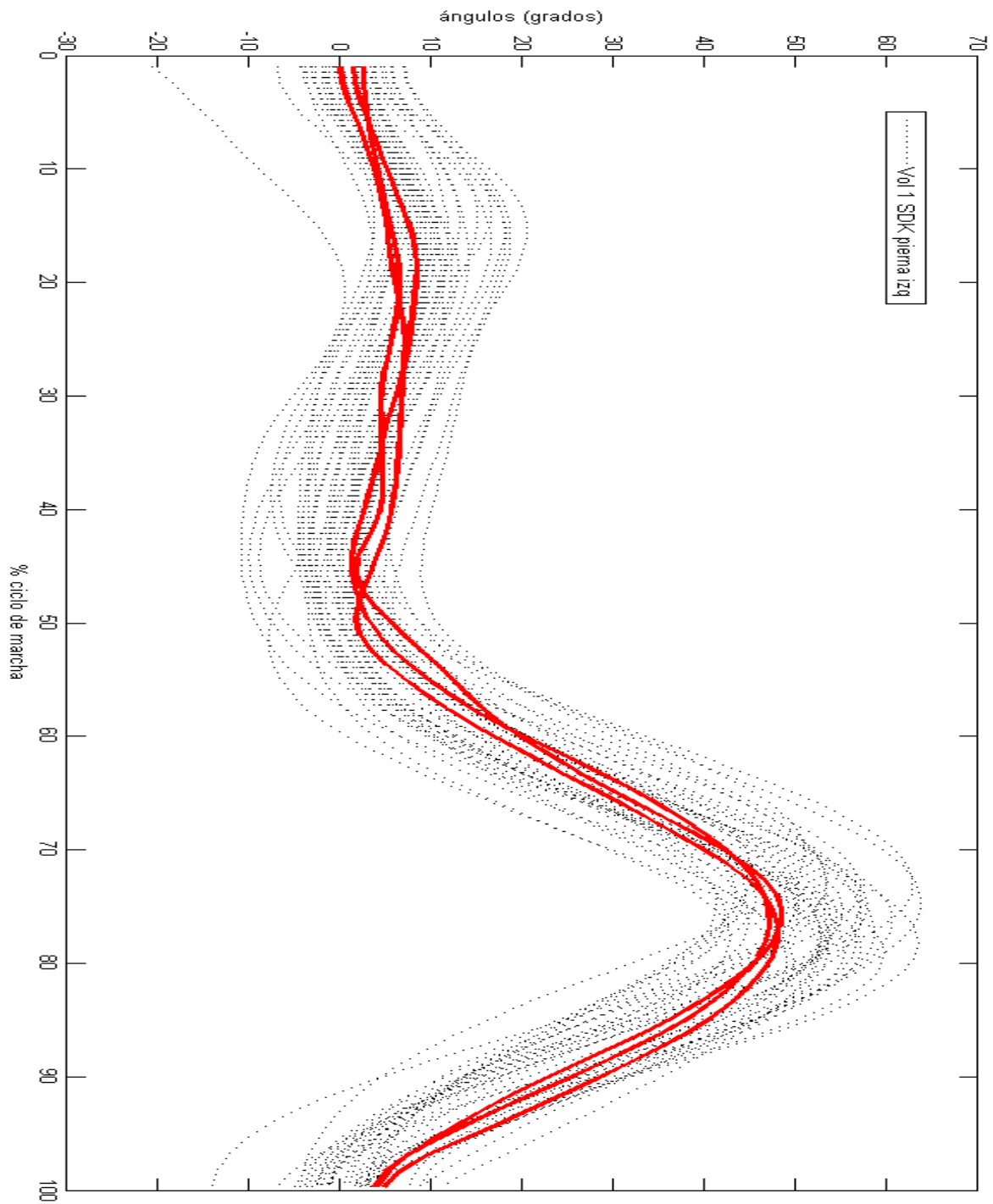


Figura 3.9. Curvas de movimiento (SDK) en la rodilla izquierda del voluntario 1

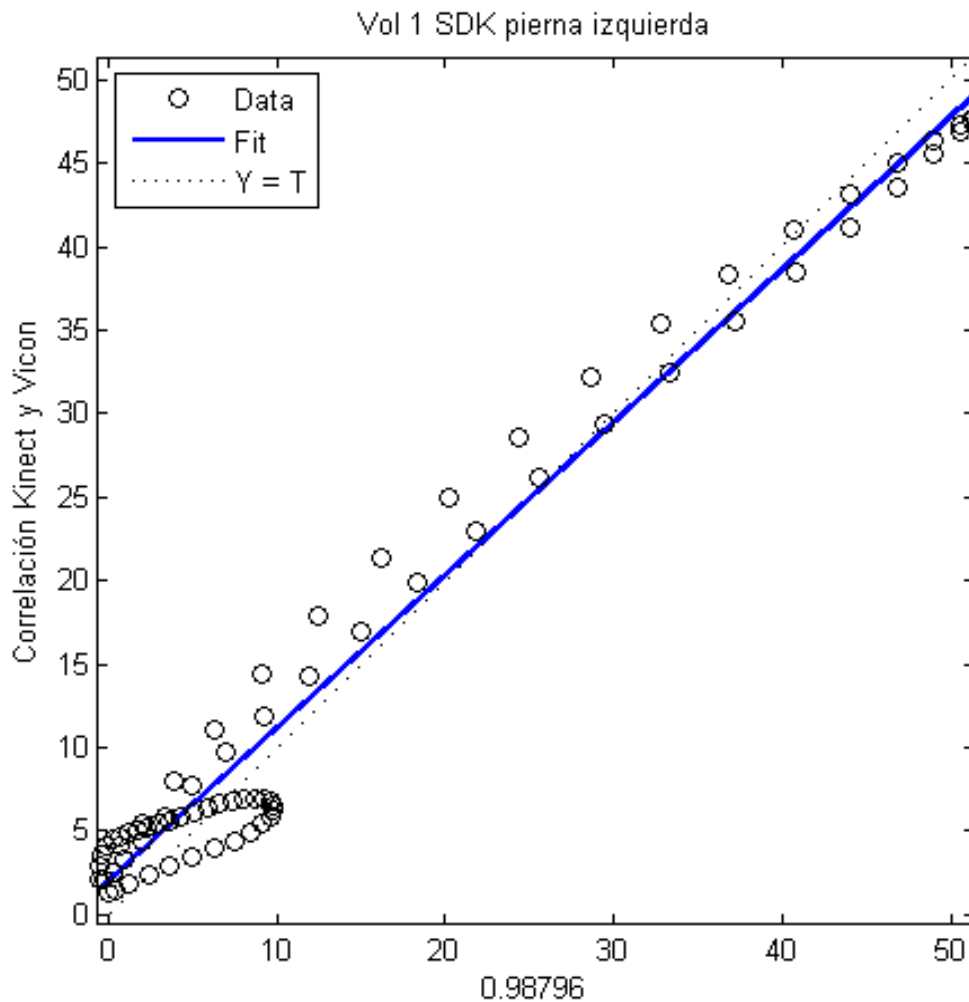


Figura 3.10. Correlación entre las curvas de movimiento la rodilla izquierda del voluntario 1 obtenidas con Kinect (SDK) y Vicon

Al igual que con la pierna derecha, en la figura 3.9 se muestra las curvas de movimiento de la pierna izquierda obtenidas con Kinect, usando la librería SDK, y el sistema Vicon, y de la misma manera las curvas obtenidas mediante Kinect muestran más similitud y menos desfase entre ellas y con respecto a las curvas medidas con el sistema Vicon. El coeficiente de correlación mostrado en la figura 3.10 vuelve a ser mayor a 0.8.

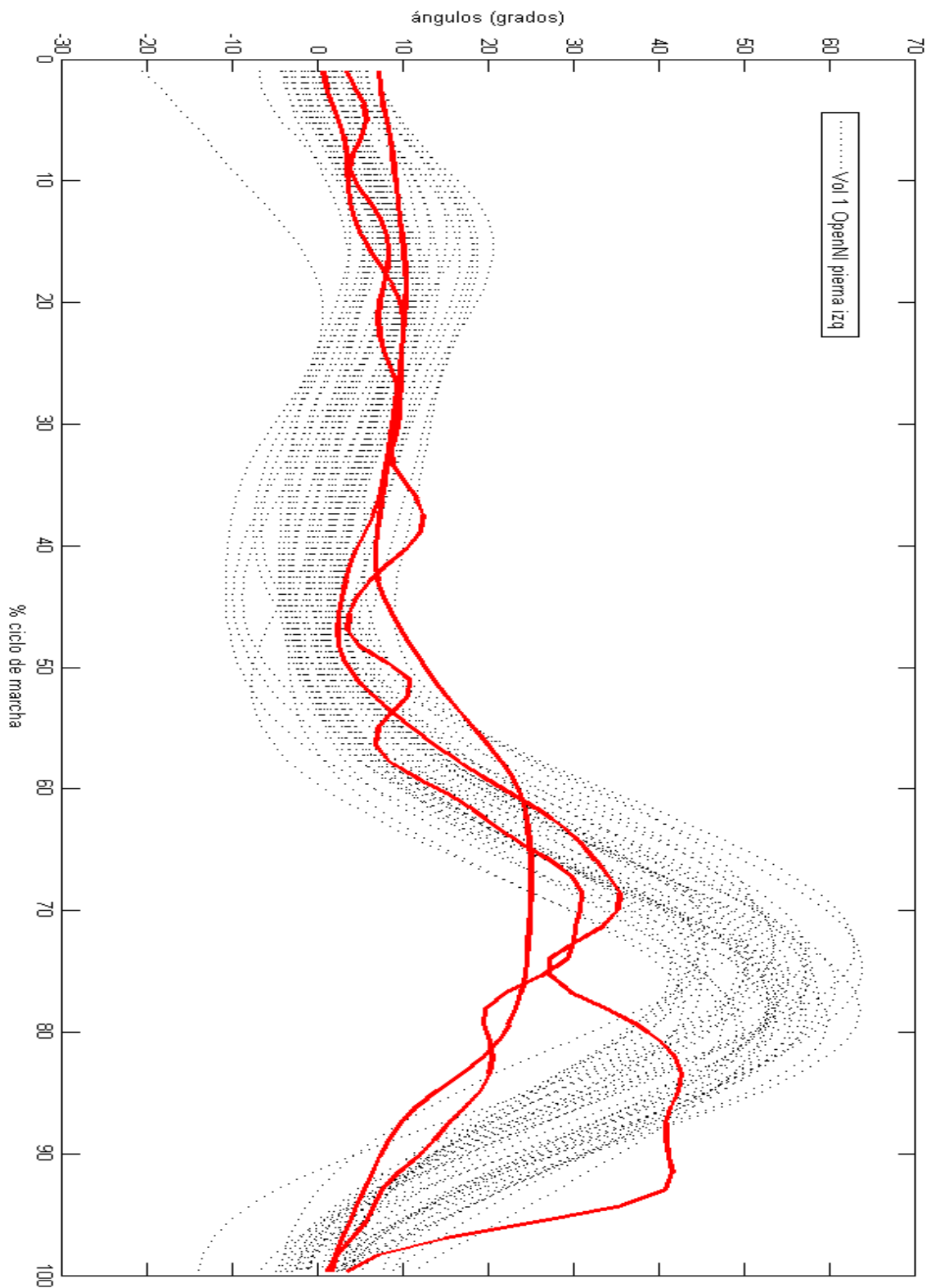


Figura 3.11. Curvas de movimiento (Openni) en la rodilla izquierda del voluntario 1

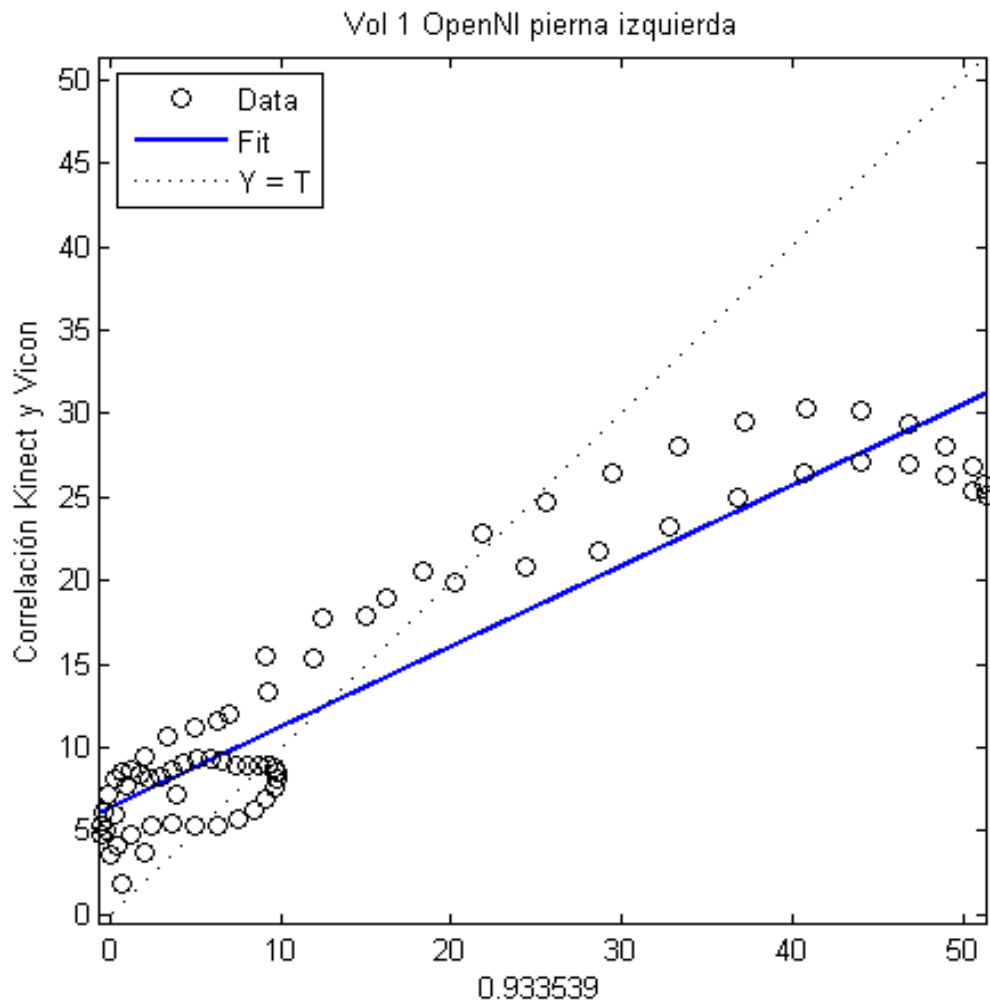


Figura 3.12. Correlación entre las curvas de movimiento la rodilla izquierda del voluntario 1 obtenidas con Kinect (OpenNI) y Vicon

En la figura 3.11 se muestra las curvas obtenidas mediante Kinect, haciendo uso de la librería OpenNI, y el sistema Vicon y al igual que en la figura 3.7, muestra menor similitud y mayor desfase con las curvas medidas con el sistema Vicon en comparación con las medidas haciendo uso de las librerías de Kinect SDK. El coeficiente de correlación de los promedios obtenidos por ambos sistemas es, sin embargo, mayor a 0.8, lo que señala que existe una similitud significativa entre las curvas.

- **Precisión y Exactitud**

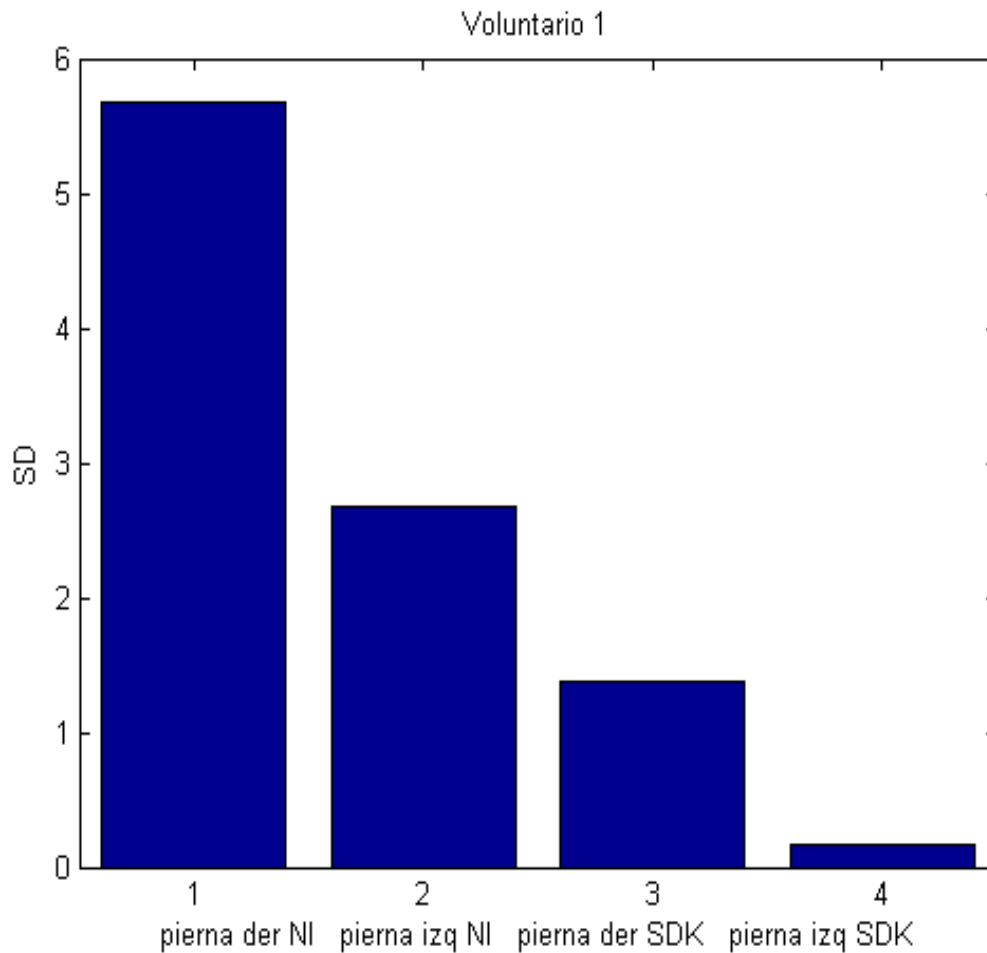


Figura 3.13. Precisión del Kinect medida de las pruebas del Voluntario 1

En la figura 3.13 se puede apreciar las precisiones obtenidas de la rodilla derecha y de la izquierda con el Kinect habiendo hecho uso de las librerías OPenNI y Kinect SDK, calculando estas precisiones con la fórmula 3. Las curvas de movimiento obtenidas mediante la librería SDK muestran una precisión mayor que las obtenidas mediante OPenNI (una menor desviación estándar representará una mayor precisión).

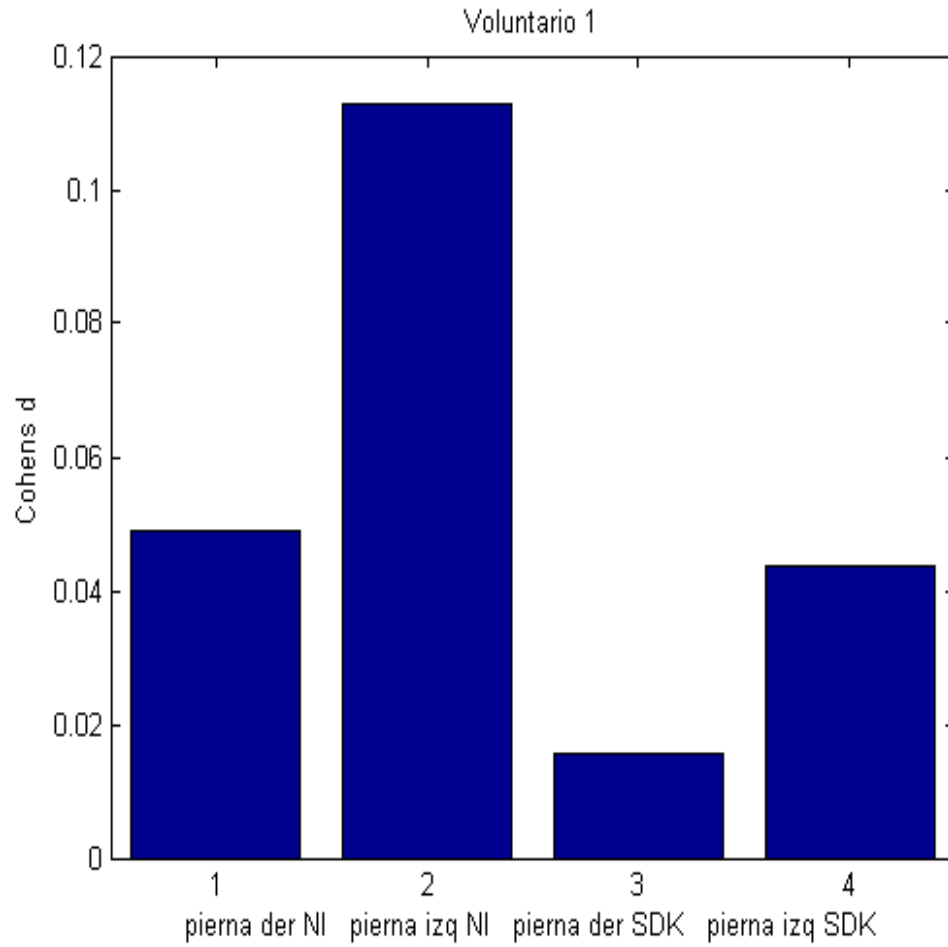


Figura 3.14. Exactitud del Kinect medida de las pruebas del Voluntario 1

En la figura 3.14 se muestran las exactitudes entre el promedio de las curvas medidas con el Kinect y el promedio de las curvas medidas con el sistema Vicon, haciendo uso de las 2 librerías anteriormente mencionadas, OpenNI y Kinect SDK. Las curvas medidas con SDK muestran mayor exactitud que las medidas con OpenNI (menor coeficiente de Cohen's d, mayor exactitud), aunque la diferencia no es tan marcada como en la precisión

CAPÍTULO 4: ANÁLISIS DE LOS RESULTADOS

4.1. Cuadros comparativos de los resultados obtenidos

En las siguientes tablas se muestran los valores de precisión, exactitud y el coeficiente de correlación de cada voluntario mostrado en los gráficos en el capítulo 3.

Tabla 4.1. Valores de voluntario 1

Voluntario 1	Rodilla Derecha	Openni-Nite	Precisión	5,68139771
			Exactitud	0,049019125
			Coef. Corr.	0.894255
		SDK	Precisión	1,379930105
			Exactitud	0,015549852
			Coef. Corr.	0.838894
	Rodilla Izquierda	Openni-Nite	Precisión	2,668793644
			Exactitud	0,112830526
			Coef. Corr.	0.933539
		SDK	Precisión	0,16282359
			Exactitud	0,043834684
			Coef. Corr.	0.98796

Tabla 4.2. Valores de voluntario 2

Voluntario 2	Rodilla Derecha	Openni-Nite	Precisión	7,343420099
			Exactitud	0,504831178
			Coef. Corr.	0.927138
		SDK	Precisión	0,333445101
			Exactitud	0,145944699
			Coef. Corr.	0.97366
	Rodilla Izquierda	Openni-Nite	Precisión	2,549384269
			Exactitud	0,320275059
			Coef. Corr.	0.916485
		SDK	Precisión	1,384525465
			Exactitud	0,18707944
			Coef. Corr.	0.935162

Tabla 4.3. Valores de voluntario 3

Voluntario 3	Rodilla Derecha	Openni-Nite	Precisión	3,248982944
			Exactitud	0,024880664
			Coef. Corr.	0.899807
		SDK	Precisión	1,542234206
			Exactitud	0,182649189
			Coef. Corr.	0.973843
	Rodilla Izquierda	Openni-Nite	Precisión	5,443353407
			Exactitud	0,114944724
			Coef. Corr.	0.930126
		SDK	Precisión	2,993643798
			Exactitud	0,24247458
			Coef. Corr.	0.832224

En la tabla 4.1 se muestran la precisión, exactitud y coeficientes de correlación de las mediciones hechas con Kinect al voluntario 1, mientras en que las tablas 4.2 y 4.3 se muestran las de los voluntarios 2 y 3 respectivamente (las gráficas de estos voluntarios pueden ser vistas en los anexos).

En todos los casos, los coeficientes de correlación entre los promedios de las curvas obtenidas mediante Kinect y el sistema Vicon fueron superiores a 0.8, por lo que se puede decir que existe una similitud significativa entre las gráficas que se obtiene a través de ambos sistemas. Las precisiones fueron mayores en todos los casos al usar Kinect SDK en comparación con OpenNI; lo mismo sucede con la exactitud, excepto en el voluntario 3.

4.2. Discusión de los resultados obtenidos

De los resultados que se han obtenido, se puede apreciar lo siguiente:

- Al observar las figuras 3.5, 3.7, 3.9 y 3.11, se puede apreciar que la fase de balanceo de la marcha, donde hay una variación del ángulo de la rodilla mayor, tiene en general una forma más cercana a la del patrón obtenido con Vicon, siendo la figura 3.11 la excepción.

- Lo opuesto se aprecia en la fase de apoyo de la marcha, donde hay menor variación del ángulo de la rodilla, ya que la forma de la variación del ángulo de la rodilla en esta fase dista más de la del patrón obtenido con Vicon.

La razón de lo mencionado anteriormente podría ser debido al ruido presente en la captura de datos del Kinect. Como ya se mencionó en el capítulo 2, el rastreo de las piernas es más inestable y ruidoso en comparación con el rastreo de la parte superior del cuerpo. Una variación de ángulo menor se ve afectada de mayor manera por el ruido que una variación más pronunciada, puesto que mientras más pequeña la variación, más difícil será distinguirla del ruido.

Otro detalle que se puede observar al ver las tablas 4.1, 4.2 y 4.3 es que las medias de las curvas de movimiento obtenidas con las librerías SDK muestran en general mayor precisión y exactitud que las obtenidas con las librerías del OpenNI, siendo el voluntario 3 la excepción (solo en el voluntario 3 se aprecia una menor exactitud al usar las librerías SDK en comparación con la exactitud obtenida con las librerías del OpenNI); además las formas de las curvas de movimiento obtenidas con el SDK son más similares al patrón obtenido con Vicon que las obtenidas con el OpenNi, particularmente en la fase de balanceo. Esto puede deberse a dos cosas:

- El SDK fue desarrollado por Microsoft, los mismos creadores del Kinect; mientras que el OpenNI fue desarrollado por investigadores y programadores que querían aprovechar las distintas posibilidades que ofrece el Kinect. Que el SDK haya sido desarrollado por los mismos creadores del Kinect podría explicar porque el SDK parece ser más inmune a los errores de rastreo de usuario que el OpenNI.
- El SDK hace una distinción entre el tobillo y el resto del pie. Que todo el pie sea detectado como un solo punto crea un problema en el rastreo del esqueleto, especialmente para esta aplicación, ya que para determinar la cinemática de la rodilla en el plano sagital se están usando las coordenadas Y y Z de las articulaciones del esqueleto virtual. La detección de la distancia del Kinect al pie (coordenada Z) se verá más afectada ya que el punto que representa al pie va a estar variando sobre la parte superior de este, generando mayor error en su lectura y, por ende, en la medición de la variación del ángulo de la rodilla.

Además de las limitaciones propias de cada una de las librerías usadas (OpenNI y SDK) y del ruido presente en la captura de datos del Kinect, es importantes mencionar otra limitación importante en este estudio que puede haber generado mayor error al capturar el movimiento de los voluntarios y por ende haber aumentado la diferencia entre la forma de las curvas de movimiento obtenidas mediante el Kinect y el sistema Vicon: el espacio de trabajo del Kinect. El Kinect solo tiene un rango de visión de 2.3 metros, lo cuál puede ser suficiente para movimientos donde solo se requiera movimientos sin desplazamiento, como en los videojuegos, pero es insuficiente para aplicaciones donde es importante el desplazamiento, como lo es en el análisis de marcha. Para poder realizar un buen análisis de la marcha es necesario que al menos tres pasos completos sean registrados, lo cual rara vez era posible con el reducido rango de visión del Kinect, pudiéndose capturar solo 2 pasos completos en la mayoría de las pruebas.

Para compensar esta limitación, se tomaron 5 pruebas por cada voluntario; tomando al final los 3 mejores ciclos de marcha de estas pruebas. Sin embargo, que solo se pueda obtener un ciclo de marcha completo por cada prueba ya es una fuerte limitación para este estudio.

CONCLUSIONES

- Este trabajo de tesis ha permitido marcar una línea base para futuros estudios y proyectos con el Kinect para el análisis de marcha, al obtener valores cuantitativos y cualitativos del Kinect con tres voluntarios diferentes de la cinemática de la rodilla, permitiendo así conocer la precisión y exactitud del Kinect para el análisis de marcha y sus limitaciones.
- En el proceso de obtención de los ángulos de la rodilla mediante el Kinect, la corta distancia de trabajo del Kinect demostró ser un problema para realizar esta tarea. El rango de trabajo del Kinect no es suficiente para poder tomar buenas medidas de la rodilla, ya que difícilmente se puede obtener tres pasos completos de los voluntarios, algo que es necesario puesto que con tres pasos se tiene un ciclo completo en cada pierna.
- La presencia de ruido en la captura de movimiento del Kinect también fue otra fuente de error al realizar el análisis de la marcha en los voluntarios,

especialmente en la primera fase del ciclo de marcha, es decir, la fase de apoyo, la cual se ve más afectada por el ruido, disminuyendo la precisión y exactitud de las medidas en esa parte del ciclo de marcha, puesto que la variación del ángulo de la rodilla en esta parte es menor en comparación con la fase de oscilación, haciéndola más difícil de diferenciarla del ruido.

- Al hacer la comparación entre los datos obtenidos con el Kinect y con el sistema Vicon, se puede concluir que el SDK es más óptimo para el análisis de marcha con el Kinect que el Openni, puesto que, aunque la exactitud del SDK no es tan superior a la del Openni, si lo es su precisión, algo que es importante si se quiere tomar medidas de las articulaciones.
- Los resultados obtenidos en este estudio confirman que se puede realizar una evaluación de la marcha humana con el Kinect, pero este no es adecuado para uso clínico como lo concluido en el estudio de Alexandra Pfister [19], reafirmando que la poca precisión del Kinect no lo hace apropiado para ser usado clínicamente en la medición de la cinemática de la rodilla. Aunque, no se espera que el Kinect tenga la exactitud ni precisión del sistema de captura de movimiento 3D Vicon debido a la aplicación de este tema de tesis, si es posible obtener variables espacio-temporales, como se menciona en el estudio de Ross A. Clark [18], que permitirán el monitoreo de la evolución de los pacientes que estén siguiendo un tratamiento de rehabilitación física de forma rápida, sencilla y barata.

RECOMENDACIONES

- Es necesario poder tomar más muestras que las que permite el corto rango de trabajo del Kinect. Para esto se sugieren dos alternativas:
 - Hacer uso de más de un Kinect, para aumentar su área de trabajo y así poder tomar más medidas, lo que permitiría tomar mejores datos. Para hacer esto, sería necesario utilizar un algoritmo de calibración que permita detectar al usuario con los dos Kinect y procesar la información de ambos al mismo tiempo. Además, sería necesario controlar la cámara infrarroja del Kinect, para así decidir en qué momentos un Kinect está registrando

información y cuando no, para evitar que un Kinect cause errores en la detección del otro.

- Otra alternativa sería que los voluntarios se coloquen sobre una faja caminadora cuando estén en frente del Kinect, puesto que así sería posible tomar muchas muestras dentro del rango de visión del Kinect. Para esto sería necesario utilizar un algoritmo de calibración diferente, que permita detectar al usuario sin que la faja le cause errores, puesto que, como se mencionó en la sección del Nite, objetos cercanos generan errores en la detección del usuario; de aquí la necesidad de un algoritmo de calibración diferente.
- Es necesario mejorar la detección del Kinect en la fase de apoyo del ciclo de marcha. Una alternativa sería hacer uso de filtros para disminuir los efectos del ruido en la detección de los puntos que representan las articulaciones, para así poder mejorar la precisión y exactitud del Kinect en la fase de apoyo del ciclo de marcha.
- Para mejorar la exactitud de los datos, es posible el uso de un sistema de sensores Kinect. Se han reportado trabajos similares en otras aplicaciones utilizando dos Kinects con resultados alentadores. De acuerdo a la literatura revisada, el uso de sensores inerciales también podría ser una alternativa de solución.

FUENTES

- [1] Whittle, M. W. (2014). *Gait analysis: an introduction*. Butterworth-Heinemann.
- [2] Reátegui Zevallos, N. V. (2012). Guía de análisis de movimiento patológico para evaluación post tratamiento quirúrgico de fractura de fémur.
- [3] Medscape. <http://www.medscape.com>
- [4] Wren, T. A., Gorton, G. E., Ounpuu, S., & Tucker, C. A. (2011). Efficacy of clinical gait analysis: A systematic review. *Gait & posture*, 34(2), 149-153.
- [5] McGinley, J. L., Baker, R., Wolfe, R., & Morris, M. E. (2009). The reliability of three-dimensional kinematic gait measurements: a systematic review. *Gait & posture*, 29(3), 360-369.
- [6] Kawamura, C. M., de Moraes Filho, M. C., Barreto, M. M., de Paula Asa, S. K., Juliano, Y., & Novo, N. F. (2007). Comparison between visual and three-dimensional gait analysis in patients with spastic diplegic cerebral palsy. *Gait & posture*, 25(1), 18-24.
- [7] El-laithy, R. A., Huang, J., & Yeh, M. (2012, April). Study on the use of Microsoft Kinect for robotics applications. In *Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION* (pp. 1280-1288). IEEE.
- [8] Jana, A. (2012). *Kinect for windows SDK programming guide*. Packt Publishing Ltd.
- [9] Schwarz, L. A., Mkhitarian, A., Mateus, D., & Navab, N. (2011, March). Estimating human 3d pose from time-of-flight images based on geodesic distances and optical flow. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on* (pp. 700-706). IEEE.
- [10] Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., ... & Moore, R. (2013). Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, 56(1), 116-124.
- [11] Zhang, Z. (2012). Microsoft kinect sensor and its effect. *IEEE multimedia*, 19(2), 4-10.
- [12] Draelos, M., Deshpande, N., & Grant, E. (2012, September). The Kinect up close: Adaptations for short-range imaging. In *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2012 IEEE Conference on* (pp. 251-256). IEEE.
- [13] Bo, A. P. L., Hayashibe, M., & Poignet, P. (2011, August). Joint angle estimation in rehabilitation with inertial sensors and its integration with Kinect. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE* (pp. 3479-3483). IEEE.

- [14] OPENNI, O. OpenNI user guide. 2010.
- [15] Sensor, P. NITE 1.3 Algorithms notes, Version 1.0, PrimeSense Inc. 2010.
- [16] Tong, K., & Granat, M. H. (1999). A practical gait analysis system using gyroscopes. *Medical engineering & physics*, 21(2), 87-94.
- [17] Mayagoitia, R. E., Nene, A. V., & Veltink, P. H. (2002). Accelerometer and rate gyroscope measurement of kinematics: an inexpensive alternative to optical motion analysis systems. *Journal of biomechanics*, 35(4), 537-542.
- [18] Clark, R. A., Bower, K. J., Mentiplay, B. F., Paterson, K., & Pua, Y. H. (2013). Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables. *Journal of biomechanics*, 46(15), 2722-2725.
- [19] Pfister, A., West, A. M., Bronner, S., & Noah, J. A. (2014). Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. *Journal of medical engineering & technology*, 38(5), 274-280.
- [20] Informe del Proyecto "Sistema robótico tipo exoesqueleto para las extremidades inferiores que permita reproducir patrones de movimiento en el plano sagital para marcha normal o movimientos asistidos dirigidos conforme a procedimientos médicos" CONTRATO: 203-FINCyT- IA-2013. Coordinador General: Dante Elías. Pontificia Universidad Católica del Perú.