

Modelo de Software para un Entorno Ubicuo de Enseñanza

N. León - nleonl@usmp.pe

Resumen— El avance tecnológico permite potenciar los entornos de enseñanza y ensayar nuevas formas de interacción entre el docente y los contenidos mostrados con medios audiovisuales en clases presenciales. La tecnología de comunicación novedosa basada en la interacción natural de usuario, permite apoyar la difusión de conocimientos y contenidos sin la preocupación para la interacción con los dispositivos de presentación.

El presente trabajo describe el modelo realizado para implementar un entorno de interacción natural mediante gestos entre el docente y los medios de presentación de contenido educativo, con el objetivo de iniciar la transformación del aula de enseñanza presencial a un entorno ubicuo.

Keywords— Entorno ubicuo, Gestos, Interacción natural de usuario, Dispositivos de presentación, enseñanza presencial.

Abstract— Technological progress allows enhancing learning environments and allows testing new forms of interaction between the teacher and the contents displayed with audiovisual media classes. Innovative communication technology based natural user interaction helps to support the dissemination of knowledge and content without concern for interaction with display devices.

This paper describes the model made to implement a natural interaction environment by gestures between the teacher and the means of presenting educational content, in order to initiate the transformation of classroom teaching in ubiquitous environment.

Keywords— Ubiquitous environment, gestures, natural user interaction, presentation devices, teaching.

I. INTRODUCCIÓN

Usar tecnología es un factor esencial para la educación universitaria, la universidad genera y acompaña los cambios en este nuevo contexto social y tecnológico, en el que la penetración, las nuevas tecnologías de información y comunicaciones (NTIC) en las universidades peruanas viene desde un proceso lento [3] a un actual proceso acelerado, caracterizado por el crecimiento de la adquisición de nuevas tecnologías en los hogares del país, cuyo indicador por trimestre supera el 80% [4].

Los contenidos educativos universitarios han sido los primeros en adoptar los cambios, la virtualización de contenidos es una de las adopciones aceptadas abrumadoramente rápido [3].

El dictado de clase también se está virtualizando con la grabación de videos y clases en línea; Coursera [4], es uno de los ejemplos más recientes y exitosos en el que el concepto MOOC (Massive Online Open Course) alcanza un gran impacto mediático cuando las universidades más prestigiosas proponen una formación a distancia planificada para alcanzar un elevado volumen de usuarios gracias a su carácter abierto,

participativo y con una metodología de inscripción gratuita [5]. En el Perú, los esfuerzos en esta línea aún son incipientes.

En cuanto a la interacción con las computadoras, mediada por teclado y mouse, se encuentra en un proceso de cambio paradigmático hacia la adopción de otras formas de interacción; ahora las computadoras son capaces de reconocer a la persona y a sus gestos, esta capacidad apoyada por el extraordinario desarrollo de los dispositivos digitales en los últimos tiempos [1] permite el acceso a los contenidos y a las clases desde cualquier lugar [2] acercándonos de manera acelerada a las características de la interacción propuesta en la computación ubicua [8].

En el dictado de clases de manera presencial, los cambios han sido escasos, se inició con la implementación de herramientas multimedia en la época de los 90 y se ha mantenido en la actualidad, el docente interactúa con los dispositivos de manera manual y apoyada por el control remoto.

Este trabajo describe las acciones realizadas para iniciar la transformación del aula de dictado de clases presencial en un entorno ubicuo.

El trabajo propone apoyar la incorporación de la tecnología de interfaz natural de usuario, en el marco de la interacción del docente con los dispositivos multimedia existentes en el aula; específicamente las presentaciones mostradas por el proyector. Se propone desarrollar los componentes necesarios para interactuar con el presentador de contenidos mediante gestos del cuerpo y la voz, dejando de lado el uso del teclado, mouse o control remoto.

Este trabajo se justifica desde el punto de vista práctico, debido a la propuesta de creación de un entorno que permitirá experimentar las influencias de las nuevas tecnologías en una clase presencial universitaria que facilite la identificación de variables asociadas al uso de estas tecnologías y experimentar el beneficio de las mismas. Desde el punto de vista teórico, generará reflexión y discusión, tanto sobre el conocimiento existente respecto a los ambientes de interacción natural aplicados al dictado de cátedra universitaria y su influencia en el proceso de enseñanza aprendizaje. Desde el punto de vista metodológico, esta investigación está generando un modelo, que conllevará a la implementación de entornos de interacción natural en la enseñanza universitaria. Esto producirá conocimiento válido y confiable para implantar nuevos componentes.

La organización del trabajo es la siguiente: en la Sección II se explica la evolución de la interacción de usuario con la computadora. La Sección III muestra detalles de la implementación de la solución. La Sección IV las pruebas realizadas con la solución desarrollada. Finalmente, se pueden encontrar las conclusiones en la Sección V.

II. LA EVOLUCIÓN DE LAS INTERFACES

Desde el dominio de las computadoras de escritorio los paradigmas interactivos han evolucionado constantemente, a partir de la interacción mediante la línea de comando con uso del teclado, luego la aparición de la computadora personal provista de una interfaz gráfica y con un dispositivo adicional llamado mouse que aseguraba una interacción fácil del usuario, a los esfuerzos enormes para lograr la inmersión del usuario en los sistemas residentes en la computadora que se bautizó como realidad virtual. En este esfuerzo se involucra la creación de una serie de dispositivos adicionales como lapiceros, cascos, lentes entre otros [5], con el fin asegurar la inmersión del usuario; encontrándose en este camino con la impetuosa necesidad de aumentar la información virtual con información real o viceversa y aparece con el concepto de realidad aumentada, la cual contribuye con una serie de códigos “Marcadores” adheridos a objetos reales sumándole ventajas a la interacción de usuario (Azuma, Weon, Lee, Jiang, & You, 1999), toda esta evolución acompañada de dispositivos de captura y procesamiento de imágenes de vanguardia finalmente evoluciona a un nuevo concepto llamado interfaz natural de usuario, en que el usuario puede interactuar con la computadora sin usar los medios físicos inmediatos, por lo que la interacción se forma natural al mundo real.

A inicios de los años noventa se acuñó el término de computación ubicua, teoría que visualiza a las computadoras en todas las cosas y con capacidad de entender al usuario de manera total [5], englobando y capturando los esfuerzos anteriormente mencionados.

La creación de dispositivo Kinect hace simple la interacción del usuario con las computadoras; este compacto, cuyo precio asequible y de fácil uso puede ser adquirido por cualquier persona como si fuera otro dispositivo o una cámara de video [6], provisto de micrófono multiarreglo, sensores con lentes de color RGB y de profundidad permite recoger datos que proporcionan capacidades de captura de movimientos en 3D de todo el cuerpo, reconocimiento facial y de voz. El sensor de profundidad compuesto de un proyector láser infrarrojo combinado con un metalóxido complementario monocromo semiconductor (CMOS), que captura vídeo datos en 3D, cuyo rango de detección de profundidad puede ser ajustado, primero ilumina la escena con el proyector infrarrojo y captura el patrón de profundidad, que es invisible para el ojo humano. El sensor de imagen CMOS lee el patrón de profundidad que está ahora distorsionado debido a los distintos objetos en la escena. Esta distorsión se procesa por un System on Chip (SoC) conectado al sensor de imagen CMOS, produciendo una imagen de profundidad de la escena. Los datos son entonces transmitidos al computador a través de “Universal Serial Bus” (USB). Este método es exacto en el rango de 80 cm a 4 metros y puede trabajar en condiciones de poca luz, como se puede encontrar en el interior de una habitación haciendo la interacción natural e intuitiva [8]. El sensor tiene un ajuste en su motor de inclinación; característica muy útil para ajustar las alturas de los objetos en el campo de visión, con un campo de visión horizontal de 57 grados, un campo de visión vertical de 43 grados, rango de inclinación física: ± 27 grados, el rango de profundidad del sensor para la Windows tiene dos opciones: Default Mode (0 – 0,8 metros, fuera de rango, 0,8 – 4 metros,

parámetros normales, 4 – 8 metros, se recoge información pero no es óptima, > 8 metros, fuera de rango). Near Mode (0 – 0,4 metros, fuera de rango, 0,4 – 3 metros, parámetros normales en donde la mejor calidad se encuentra a los 2 metros, 3 – 8 metros, se recoge información pero no es óptima, > 8 metros, fuera de rango) [12]. El flujo de datos es de 80 x 60, 320 x 240, 640 x 480 @ 30fps de resolución de imagen de profundidad y 640 x 480 32-bit de color @30fps, 1280 x 960 RGB @ 12fps, Raw YUV 640 x 480 @ 15fps de imagen Color. [7].

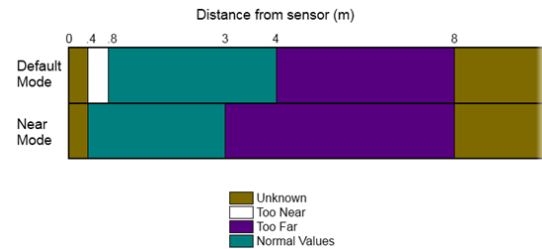


Ilustración 1 Campo de visión KINECT fuente [4]

Además el sistema de seguimiento es capaz de rastrear hasta 6 personas, incluyendo 2 personas activas, rastrear 20 articulaciones por persona.

El sistema de audio es muy sólido provisto de 32-bit @ 64 kHz y un sistema de cancelación de eco que aumenta la entrada de voz y reconocimiento de voz múltiple permite el reconocimiento de hasta 26 lenguas distintas. Esto agiliza el camino de la inclusión de la interacción natural en el desarrollo de software de realidad aumentada el “Gestvelopment” (Gestual Development o tecnología GDM) [4], por tanto Kinect, permite desarrollar componentes capaces de interpretar movimientos y voz y trasladarlos a acciones en el entorno de desarrollo gracias a la arquitectura que trae consigo, permitiendo desarrollar interfaces inteligentes capaces de entender los movimientos, voz, posiciones y gestos del cuerpo.

2.1. Reconocimiento de gestos

El reconocimiento de gestos es un gran esfuerzo de las Ciencias de la Computación [8] y la Tecnología del lenguaje, ambas especializaciones se fusionan en el esfuerzo de interpretar la amplia gama de gestos en comunicación, los cuales varían mucho entre los contextos y culturas. Las emociones universales tienen influencias culturales en los niveles absolutos de precisión en el reconocimiento y en las sentencias de la intensidad exterior y la experiencia subjetiva interna [9]. El reconocimiento de gestos puede llevarse a cabo con técnicas división por computadora y procesamiento de imágenes. Los gestos son detectados en los movimientos corporales significativos a las personas, los esfuerzos de ambas disciplinas se han centrado en el reconocimiento de emociones en la cara y el reconocimiento de gesto con la mano. El reconocimiento de gestos permite a los seres humanos interactuar con la máquina de forma natural sin ningún tipo de dispositivos intermediarios.

En este trabajo vamos a enfocarnos en el reconocimiento de gestos con Kinect en el que el algoritmo Skeletal tracking logra identificar partes del cuerpo de las personas que están en el campo de visión del sensor, pudiendo obtener puntos o

articulaciones que hacen referencia a las partes del cuerpo de la persona y hacen un seguimiento identificando gestos y/o posturas del esqueleto y su posición en el espacio de una forma sencilla y luego almacena los datos en una matriz para su posterior utilización en el algoritmo que se esté desarrollando.

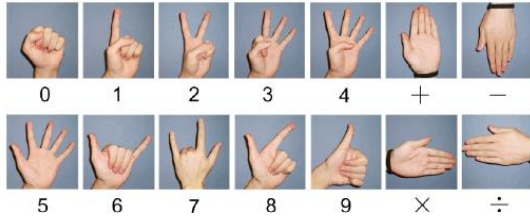


Ilustración 2 Gestos con la mano Fuente: [12]

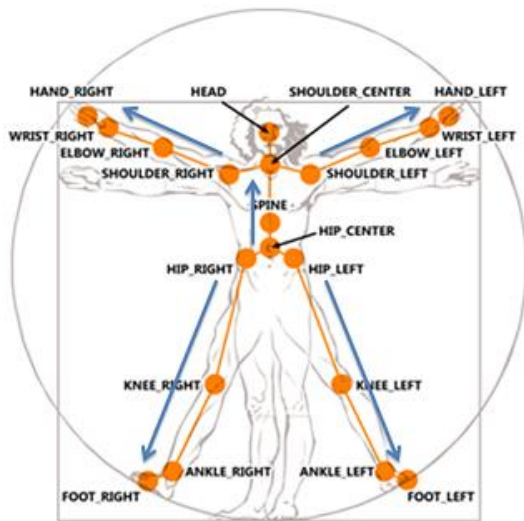


Ilustración 3 Skeleton Data Fuente [10]

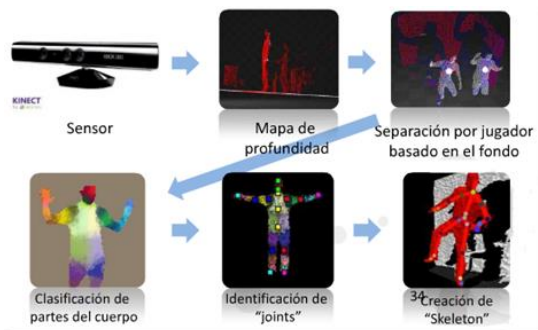


Ilustración 4 Identificación de articulaciones para crear el esqueleto formado por estos puntos Fuente [10]

La nueva versión del SDK 1.7 de Kinect trae consigo una serie de algoritmos para el reconocimiento de gestos con la mano lo cual facilita los algoritmos de este tipo [14], por ejemplo, agarrar, empujar, arrastrar, definir un área de trabajo y señalar. Los algoritmos permiten la identificación de los píxeles de la imagen, la extracción de características de los píxeles identificados con el fin de clasificar la mano en una de un conjunto predefinido de poses, y reconoce la ocurrencia de secuencias específicas que se plantean como gestos.



Ilustración 5 reconocimiento de gestos por Kinect, Fuente propia.

Fuente: [7]

III. IMPLEMENTACIÓN

Para el desarrollo de un entorno de interacción natural entre el docente y software de presentación en una clase presencial se ha seguido el siguiente proceso.

2.2. Identificación espacio y componentes

Se identificó el espacio del aula de clase y los componentes que contiene.

Espacios: compuesto por el espacio del dominio del profesor delante de la pizarra y lugar de proyección; espacio del dominio del estudiante, lugar de recepción de clases; espacio de proyección, lugar en donde el proyector mostrará los contenidos.

Componentes tecnológicos: Módulo de dispositivos multimedia, integrado por un proyector multimedia, computadora, conexión a internet, puntos de acceso a la alimentación eléctrica disponible y acceso al computador mediante puerto USB; Módulo de interfaz natural, integrado por un sensor kinect for Windows y el software de interfaz entre el sensor y la computadora.

Aula de Clase

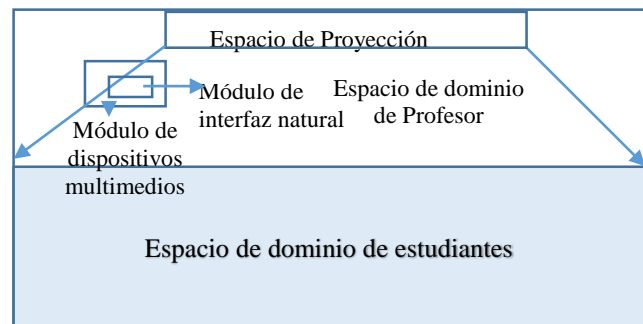


Ilustración 6 Componentes de aula de clase. Fuente propia.

2.3. Identificación de actores

Identificación de características de docentes: Grupo de personas que interactúan con las presentaciones de manera permanente.

Identificación de características de estudiantes: Este grupo recibe la información e interactúan con las presentaciones de manera esporádica.

2.4. Identificación de Funcionalidades

Funcionalidades de interacción del docente: El docente solicita activar las diapositivas, avanzar, retroceder y terminar la presentación.

Tabla 1 Patrones de gestos propuestos

Gestos	Acciones
Cerrar el brazo derecho	Avanza la diapositiva
Cerrar el brazo izquierdo	Retrocede la diapositiva
Cerrar la mano	Termina la presentación

Estructura de clases:

Estos algoritmos se estructuraron en tres grupos que permiten identificar los gestos para las manos, los gestos para el cuerpo y responder a la voz.

Knicect API: Componente que analiza la escena de la imagen, reconoce el cuerpo y genera el skeleton, con información de los joint y la distancia hacia el dispositivo [10], en el proyecto se ha trabajado en Default Mode.

Las funcionalidades para avanzar y retroceder las diapositivas involucran los algoritmos de los movimientos de ambos brazos. Los algoritmos del movimiento con los brazos incluyen un tiempo promedio entre 200 y 500 milisegundos, este tiempo fue definido por acciones de prueba en donde se escogió el mejor promedio de movimientos en personas con estaturas y volumen corporal similar.

Swipe left: Identifica el gesto de cerrar el brazo izquierdo y acciona la función de retroceder diapositiva.

Swipe Right: Identifica el gesto de cerrar el brazo derecho y acciona la función de avanzar diapositiva.

Las funcionalidad cerrar la presentación de diapositivas involucran dos algoritmos, el primero le pertenece a la acción de mostrar la mano y bajarla hasta posición horizontal.

El segundo permite detectar la voz, específicamente la palabra "Terminar".

Hand trackingP: Componente del software que retorna la localización específica de las manos.

Close: componente que identifica la acción de la mano y acciona la función de terminar la aplicación.

BodyP: Recibe la información del sensor y genera información relacionada con el cuerpo, como las articulaciones, la distancia del cuerpo al dispositivo.

VoiceP: Componente que identifica la palabra "Terminar", y acciona la función, terminar presentación.

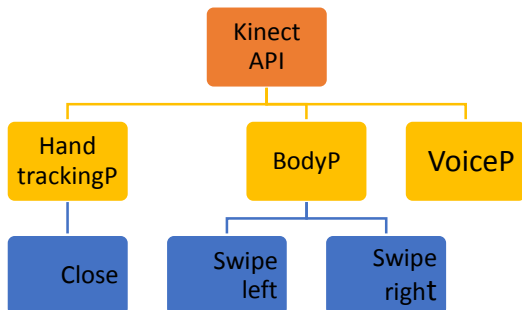


Ilustración 7 Modelo de Clases, Fuente propia.

Funcionalidades de interacción del estudiante: El estudiante necesita apreciar el contenido de la presentación con el menor ruido posible, se define ruido, como cualquier acción que perturbe la atención del estudiante con respecto a la presentación en la diapositiva.

1.1. Identificación de restricciones

Identificaciones de restricciones: Las distancias de acción del profesor.

Altura del sensor de movimientos: La altura en que debe ser colocado el sensor Kinect para el dictado de clases está en el rango de 60 cm a 1.5 metros.

La distancia de influencia: Entre el docente y el sensor Kinect debe haber como mínimo un metro de distancia y como máximo tres metros.

El grado de iluminación sobre el campo de influencia: La iluminación sobre la pizarra es longitudinal, tiene interruptor independiente, un arreglo de cuatro grupos de tres lámparas fluorescentes cuyo apagado y encendido es en paralelo, lo que permite la visibilidad de la proyección [8].

Restricciones de estatura del docente: Las estaturas del docente varían desde 1 metro a 1.90 centímetros.

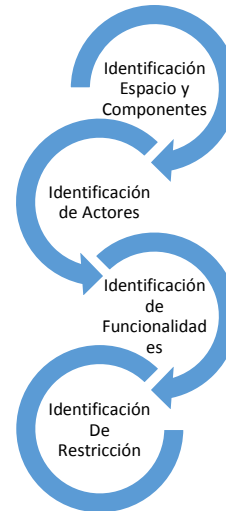


Ilustración 8 Modelo de software, Fuente propia.

IV. PRUEBAS

El equipo experimental, fue instalado en un aula específica, y en el turno noche, por las acciones de instalación del módulo de interfaz natural.

Los estudiantes pertenecían a la carrera de Ingeniería de Computación y Sistemas y se encontraban en el sexto ciclo, su selección fue dictaminada por el aula de clases, seleccionada.

Tabla 2 Cursos y ciclos de prueba

Cursos	Ciclos
Ingeniería de software I	6to ciclo.
Ingeniería de costos	6to ciclo.
Gestión de procesos	6to ciclo

XX Congreso Internacional de Ingeniería Eléctrica, Electrónica, de Computación y Ramas Afines, 5-9 Agosto 2013, UPAO, Perú ISSN 2309-2111

El perfil del docente escogido definió que sea variable para mediar la variable de funcionalidad de la aplicación y el accionar del docente, por tanto se escogió un docente acostumbrado a usar las herramientas de tecnología y un docente con poco uso de herramientas de tecnología.

Tabla 3 Perfil de usuarios

Cursos	Tecnología
Docente de curso de tecnologías	Usa tecnología
Docente de cursos de gestión y administración	No usa tecnología

Tabla 4 Estatura y sexo de docentes

Cursos	Estatura	Sexo
Ingeniería de software I	1.68	Femenino
Ingeniería de Costos	1.65	Masculino
Gestión de procesos	1.70	Masculino

Con respecto a la adaptabilidad del docente, se trabajó la medición de las distancias, los ángulos de seguimiento, y la velocidad de movimiento del brazo.

Se desarrollaron dos capacitaciones antes del dictado de clases con el uso de la herramienta. El docente identificó la posición y las distancias de influencia del hardware y software.

Estas distancias deben ser identificadas antes de ser iniciada la clase, esto permite el uso libre de todos los gestos a los que el docente está acostumbrado fuera del alcance de identificación del software.

En cuanto al modo de impartir la clase presencial, una vez que el docente se acostumbró al uso de la solución, le fue fácil controlar los espacios y las acciones.

En cuanto a la reacción del estudiante en los primeros cinco minutos de la primera clase fue de sorpresa e interrogación hasta que descubrieron la cámara Kinect y la clase volvió a la normalidad y aceptación total.

V. CONCLUSIONES.

El resultado del presente trabajo, es un esfuerzo para experimentar las influencias tecnológicas en las clases presenciales y descubrir las variables que se deben de manejar para adaptarse al cambio.

En el trabajo se ha descubierto algunas variables como el manejo de tiempos, espacios y dispositivos multimediales alternativos.

Esto supone un replanteamiento respecto a los métodos e instrumentos de trabajo del docente en clase presencial y un compromiso por parte de las instituciones educativas para la investigación e implementación de un medio de interacción natural del docente y el software usado en el dictado de clases presenciales.

El experimento desarrollado en este trabajo supone el inicio de una serie de estudios con objetivos de lograr complementar la solución a fin de que permitan la ubicuidad de las clases presenciales.

El estudiante ha aceptado la solución de manera total y en la primera clase, pasando del asombro inicial a la aceptación total.

VI. REFERENCIAS

- [1] A. Ascarza, «E-Learning en el Perú: Un estado del arte,» San Marcos, Lima, 2007.
- [2] INEI, «Las Tecnologías de Información y Comunicación en los Hogares,» INEI, Lima, 2011.
- [3] Universia, «Educación a distancia: Oferta virtual crece en universidades del país,» 19 Julio 2010. [En línea]. Available: <http://noticias.universia.edu.pe/en-portada/noticia/2010/07/19/691220/educacion-distancia-oferta-virtual-crece-universidades-pais.html>.
- [4] Coursera, «Coursera,» 2012. [En línea]. Available: <https://www.coursera.org/>.
- [5] M. Gea, R. Montes, RojasBelen, A. Marin, A. Cañas, I. Blanco, V. Gámiz, A. Del Río, D. Bravo-Lupiañez, M. Cádiz y C. Gutierrez, «Formación abierta sobre modelos de enseñanza masivos: nuevas tendencias hacia el aprendizaje social,» Granada, 2012.
- [6] B. Cope y M. Kalantzis, «Aprendizaje ubicuo,» *Ubiquitous Learning. Exploring the anywhere/anytime possibilities for learning in the age of digital media*, p. 244, 2009.
- [7] G. Meneses, «Universidad: NTIC, interacción y aprendizaje,» *EduTec. Revista Electrónica de Tecnología Educativa*, p. 15, 2006.
- [8] M. Weiser, «The Computer for the 21st Century,» 1991.
- [9] «Theory and Research in HCI: Morton Heilig, Pioneer in Virtual Reality Research,» September 2008. [En línea]. Available: http://www.telepresenceoptions.com/2008/09/theory_and_research_in_hci_mor/.
- [10] J. Tong, J. Zhou, L. Liu y Z. Pan, «Scanning 3D Full Human Bodies using Kinects,» *IEEE Virtual Reality*, 2012.
- [11] L. D'Souza, I. Pathirana, D. McMeel y R. Amor, «Kinect to Architecture,» Nueva Zelanda, Auckland, 2011.
- [12] L. Cruz, L. Djalma y L. Velho, «Kinect and RGBD Images: Challenges and Applications,» Rio de Janeiro, 2012.
- [13] Microsoft, «Kinect for developers,» 2013. [En línea]. Available: <http://msdn.microsoft.com/en-us/library/jj131033.aspx>.
- [14] ACM, «Conferencias y eventos de la ACM y otras organizaciones,» [En línea]. Available: <http://www.acm.org/calendar-of-events>. [Último acceso: 22 June 2012].
- [15] D. Matsumoto, «Culture and nonverbal behavior,» San Francisco, 2006.
- [16] Z. Ren, M. Junsong y Z. Zhang, «Robust Gesture recognition With Kinect Sensor,» Redmon, 2010.

- [17] Microsoft, «Tracking User With Kinect Skeletal Tracking,» 2011. [En línea]. Available: <http://msdn.microsoft.com/en-us/library/jj131025.aspx>. [Último acceso: 24 September 2012].
- [18] MSDN, «Reto SDK kinect: Reconocer gestos con Skeletal tracking,» 22 August 2011. [En línea]. Available: <http://blogs.msdn.com/b/esmsdn/archive/2011/08/22/reto-sdk-kinect-reconocer-gestos-con-skeletal-tracking.aspx>. [Último acceso: 24 September 2012].
- [19] Microsoft, «Communicate with computers naturally,» 03 2013. [En línea]. Available: <http://www.microsoft.com/en-us/kinectforwindows/>. [Último acceso: 10 05 2013].
- [20] J. Webb y A. James, «Beginning Kinect Programming With the Microsoft Kinect sdk.,» 2012.
- [21] MINEDU, «Criterios normativos para el diseño de locales de educación básica regular, niveles de inicial, primaria, secundaria y básica especial,» Lima, 2006.
- [22] R. Azuma, J. Weon, B. Lee, P. Jiang y S. You, «Tracking in unprepared environments for augmented reality systems,» Computers & Graphics, 1999.
- [23] L. D'Souza, I. Pathirana, D. McMeel y R. Amor, «Kinect to Architecture,» 2011.
- [24] L. Cruz, D. Lucio y V. L. , «Kinect and RGBD Images: Challenges and Applications,» 2012.



Norma León Lescano, es ingeniera de Computación y Sistemas de la Universidad de San Martín de Porres (USMP), tiene estudios de Maestría en Administración de Empresas y E Business, está realizando el último año del doctorado en Ingeniería de Sistemas de Información. Es jefa del Laboratorio de Investigación Aplicada de la escuela de Ingeniería de Computación y Sistemas de la USMP, es

catedrática universitaria.