

Propuestas de Soluciones TIC emergentes para Personas con Discapacidad

Wilfredo A. TALLEDO

Facultad de Ingeniería, Universidad Peruana de Ciencias Aplicadas
Santiago de Surco, Lima, Lima 33, Perú

Samuel O. VILLEGAS

Facultad de Ingeniería, Universidad Peruana de Ciencias Aplicadas
Santiago de Surco, Lima, Lima 33, Perú

y

Alfredo BARRIENTOS

Facultad de Ingeniería, Universidad Peruana de Ciencias Aplicadas
Santiago de Surco, Lima, Lima 33, Perú

RESUMEN

Actualmente se quiere lograr una total inclusión social dentro del Perú, en la cual también considera a las personas con alguna deficiencia que les impida realizar actividades comunes de la vida cotidiana. La búsqueda de su mayor participación se potenció con la declaración de la Ley N.º 29973, Ley General de la Persona con Discapacidad, que promueve la inclusión de estas personas, casi 6% de la población, en la vida política, económica, social, cultural y tecnológica. Esto crea la necesidad de conocer y saber qué opciones existen para facilitar la adopción de estas personas y también para mejorar su calidad de vida. Esta investigación propone proyectos que mejoren la independencia y estilo de vida de las personas con discapacidad mediante el empleo y/o la adaptación de las características de tecnologías y soluciones existentes. En este trabajo se describe una forma de resolver la limitación comunicativa de los sordomudos empleando Creative, una cámara interactiva gestual, en un sistema de reconocimiento de señales; una aplicación de entrenamiento fonético como tratamiento para la pos-cirugía del labio leporino, una aplicación de una nueva interfaz móvil para una mayor accesibilidad de los ciegos, y una aplicación de autoaprendizaje braille con un teclado orientado para ciegos.

Palabras Claves: Discapacidad, Tecnología de apoyo, lenguaje de señas, Ceguera, Accesibilidad.

1. INTRODUCCIÓN

En el Perú, según la primera encuesta nacional especializada sobre discapacidad realizada en diciembre de 2012 por el Instituto Nacional de Estadística e Informática, existen 1'575,402 de personas con algún tipo de discapacidad, el 5.2% de la población en dicho año [1]. Esto muestra y pone en conocimiento público un aspecto social de la realidad presente en el país. En lo legal, se tiene la Ley N.º 29973 (Ley General de la Persona con Discapacidad), con el propósito de promover un ambiente de inclusión en diferentes aspectos de la vida como ciudadano. A pesar de estas iniciativas, en otras estadísticas de Lima Metropolitana, se tiene cómo referencia que los servicios de accesibilidad tienen un bajo índice. Está determinado que el acceso a salud para personas con discapacidad es del 34%,

adicionalmente, un 23% de la población afirma tener dificultades para acceder a servicios tan básicos como transporte y sólo un 10% recibe un tipo de tratamiento adecuado [15]. Tomando como referencia estas estadísticas, se puede asumir un alarmante bajo conocimiento de tecnologías adaptativas y, en consecuencia, mucho menos indicios de dar soluciones a problemas básicos de comunicación.

Adicionalmente, se realizó una revisión de soluciones modernas para llegar a la conclusión contextual del problema planteado. Se revisaron soluciones para diferentes limitaciones, por ejemplo, para las del habla y audición, se descubrieron muchas soluciones de interpretación de señas donde gran parte se basan en cámaras 2D cuyo principal problema es la complejidad de procesamiento y su considerable porcentaje de fallas en interpretación. Para las personas invidentes se encontraron varias soluciones que simulan un teclado *perkins* para el ingreso de información orientado a sus limitaciones. A pesar del aporte, creemos que se puede mejorar el concepto de la solución con ciertas funcionalidades adicionales. Asimismo, se encontraron muchas aplicaciones móviles para este público, pero sólo una integraba varias de estas en una sola interfaz y no incluía ciertas aplicaciones que serían de mayor utilidad para estas personas. Además, sólo está disponible en el sistema operativo Android. Por otro lado, en la propuesta para el tratamiento post-cirugía del labio leporino no hubo ninguna solución tecnológica. Sólo se recopiló que se usan ciertos instrumentos en forma de manguera para la interpretación terapéutica de los resoplidos del paciente.

En evidencia de estas falencias, planteamos soluciones prácticas con base a tecnologías y productos revisados. Para la solución de reconocimiento de señas, se ha identificado una plataforma que evita la necesidad de desarrollar algoritmos complejos para el reconocimiento de los dedos de la mano. Para el teclado braille, se revisó la tecnología NFC para agregar funcionalidades de autoaprendizaje con un sintetizador de voz. Para la aplicación interfaz en el sistema Android, se plantea agregar un par de soluciones de reconocimiento de billetes y lector de libros en formato DAYSI en la plataforma Windows Phone 8.1. Y, por último, para el tratamiento post-cirugía del labio leporino se planteó una solución de autoentrenamiento con el uso de la API de reconocimiento de voz de la plataforma Windows Phone 8.1. La organización de este trabajo está compuesta por cuatro secciones. En la sección dos, Revisión de Literatura, describimos los trabajos e información relevante en TICs y temas

relacionados que permiten abordar las soluciones con mayor conocimiento. En la sección tres, Propuestas de Soluciones, detallamos los resultados del esfuerzo plasmado en propuestas. Finalmente, en la sección cuatro, Conclusiones, se resume cómo se ha llegado a la solución para el problema planteado.

2. REVISIÓN DE LITERATURA

A continuación, se hará una revisión de la literatura relacionada a las propuestas planteadas antes de abordarlas posteriormente. Con respecto al reconocimiento de señas, se revisó el campo de aplicación de la tecnología de reconocimiento de gestos, los cuales fueron descritos como cuatro por Murthy y Jadon [2]: Realidad Virtual, Interacción Robótica, Aplicaciones de Computadoras y Tabletas y Videojuegos.

En el campo de los sistemas de reconocimiento de señas, se encontró que estos se basan en tres tipos de aproximaciones para poder obtener la información necesaria para operar:

- Aproximación por Modelo Kinemático. Consiste en obtener los parámetros de una imagen 2D con base a un modelo 3D. Tiene como limitaciones que la mano de por sí carece de una textura estándar y para identificar los bordes se tiene que hacer uso de técnicas de contrastes fuertes [2].
- Aproximación por vista. Comparación intensa de diferentes imágenes de un modelo 2D de la mano para obtener su significado [2].
- La aproximación por características de bajo nivel. Está basada en la argumentación que no es necesario obtener imágenes de alta calidad, sino en el uso de algoritmos para identificar puntos críticos de la mano, como la palma [2].

Cuando la primera cámara Kinect sale al mercado, surgen muchas propuestas de reconocimiento de señas (*Human Gesture Recognition*) por la nueva capacidad de identificar las articulaciones del cuerpo al procesar las imágenes en 3D. Esta funcionalidad resuelve muchos problemas de sistemas similares, cuyas limitaciones exige el uso de varios algoritmos complejos para poder identificar la mano del usuario del resto de la imagen [19]. A pesar de estas mejoras, esta primera versión de Kinect también tenía limitaciones confirmadas en [20] por el tamaño de las manos a detectar y lo difícil que era identificar los bordes entre los dedos. Para resolver estas nuevas limitaciones, fue necesario, también, aplicar ciertos algoritmos para la identificación del movimiento de los dedos como el Finger-Earth Mover's Distance (FEMD).

Por su parte, Intel lanza al mercado una nueva cámara de reconocimiento de gestos que posee una serie de funcionalidades embebidas para la detección de dedos y pulgares [27]; esta resuelve por sí sola muchas limitaciones de otros sistemas, incluso el mencionado anteriormente. Asimismo, la compañía pone a disposición de los desarrolladores muchos proyectos donde se muestran ejemplos de aplicaciones simples usando las nuevas funcionalidades de rastreo de dedos y pulgares en [22,23], y su infinidad de usos en soluciones de tecnología.

Con relación a la propuesta de la aplicación interfaz para personas invidentes, se pudo confirmar un mercado actual de aplicaciones móviles que ha ido creciendo a pasos agigantados, incluyendo muchas aplicaciones móviles con características específicas para personas invidentes. El fabricante con más iniciativa para la accesibilidad ha sido de desde los inicios del mercado, Apple [5], seguido de esta, Android, el sistema

operativo más usado en la actualidad. Este último ha liberado en el mercado una considerable cantidad de aplicaciones para problemas específicos de personas invidentes. Ambas marcas presentan soluciones que se enfocan en problemas del día a día que las personas invidentes afrontan, como las siguientes:

- Elección de llamadas por comando de voz (Voice Command – Android) [7]
- Reconocimiento de billetes (LookTel Money Reader – iOS) [3]
- Identificación de objetos (ScanLife Barcode and QR Reader – Android) [4]
- Navegación (Blind Navigator – Android) [8]
- Lectores de Contenido (VoiceOver – iOS/TalkBack - Android) [6,7]

Estas aplicaciones nos permiten dar un vistazo de la gran población de aplicaciones para personas invidentes.

Para abordar la solución de tratamiento de labio leporino, se debe entender un poco más sobre esta misma. Una disfunción del velofaríngeo, o labio leporino, se refiere a la condición en la que este no se cierra por completo al momento de producir el sonido oral de comunicación [17]. Este tiene definido tres tipos de disfunción:

- Insuficiencia Velofaríngeal [17]: este tipo se caracteriza por ser un defecto físico en el paladar que evita cerrar por completo la boca con la nariz a nivel del paladar blando [16].
- Incompetencia Velofaríngeal [17]: va relacionado con un desorden neuromotor o psicológico que provoca movimientos estructurales pobres.
- Aprendizaje erróneo del Velo faríngeo [17]: va relacionado con el movimiento equivocado del Velofaríngeo al emitir sonidos.

En algunos escenarios, después de haber procedido con una cirugía correctiva de la hendidura, el paciente puede desarrollar articulaciones incorrectas al momento de emitir sonidos del habla. Estas anomalías al hablar son conocidas también como producciones compensatorias [17] relacionadas a su previa condición. En consecuencia, el tratamiento principal y básico después de una cirugía correctiva es la retroalimentación auditiva del habla. Esta terapia consiste en el contraste del sonido oral que el paciente emite y el sonido correcto que debería emitir [16]. La terapia de dictado no es efectiva si no se ha realizado una cirugía previa por ser un impedimento físico. El tratamiento de retroalimentación oral se enfoca en una disfunción velofaríngeal relacionada al incorrecto aprendizaje para hablar [17] debido a una condición física previa.

3. PROPUESTAS DE SOLUCIONES

En esta sección presentamos las propuestas de soluciones para las siguientes limitaciones: comunicación de personas invidentes, tratamiento del labio leporino y comunicación de personas sordomudas.

Teclado Braille con Funcionalidades de Autoaprendizaje

Problema. El principal problema que presentan las personas con ceguera es el ingreso de texto en los dispositivos móviles de pantalla táctil debido a que el teclado está orientado más para las personas sin esta deficiencia.

Trabajos similares. Frey *et al.* [10] propusieron y desarrollaron una solución que transforma todo el dispositivo

móvil en un teclado braille: *BrailleTouch*. Este funciona mediante la representación de los caracteres en la pantalla al dividirla en seis bloques. El ingreso de los caracteres se realiza a través de la presión de los distintos puntos simulando su escritura braille. Similarmente, Mascetti *et al.* [11] dan otra solución, *TypeInBraille*, que consiste en dividir la pantalla por la mitad, en dos sectores. El ingreso de cada letra se realiza mediante tres toques en la pantalla, simulando cada par de puntos de las filas del carácter en braille. La ventaja de *TypeInBraille* sobre el primero es el hecho que sólo necesita que el dispositivo móvil reconozca mínimamente 3 toques simultáneos, mientras que el otro requeriría hasta 5.

Southern *et al.* [12] ofrecen una comparación de estas dos soluciones y otras más. Los resultados de este aporte son:

- *TypeInBraille*, presentada por Mascetti *et al.* [12] Ingreso de una letra en tres toques basándose en las filas del carácter en braille. Alcanzó una velocidad de 6.3 palabras por minuto con un error de ingreso de 3%.
- *BrailleType*, presentada por Oliveira *et al.* [13] Ingreso punto por punto. Alcanzó una velocidad de 1.45 palabras por minuto con un 9.7% de error.
- *Perkinput*, presentada por Azenkot *et al.* [14] Ingreso similar a *TypeInBraille* pero en dos toques y orientado a las columnas de la letra en braille. Alcanzó una velocidad de 6.1 palabras por minuto con un error de 3.5%.
- *BrailleTouch*, presentada por el equipo del autor. Ingreso de una letra mediante la presión coordinada en la pantalla para la representación del carácter. Alcanzó una velocidad de 23.2 palabras por minuto.

Todas las soluciones dan una salida para la limitación de ingreso de información. Sin embargo, falta solucionar la brecha de aprendizaje que, muchas veces, es lo más complicado. Ninguna de las soluciones plantea una propuesta de autoaprendizaje ni tutorial para una persona que no sabe usar teclado braille y mucho menos el sistema braille. Se puede causar un gran impacto si se supera esta limitación y se mejora el proceso de aprendizaje con un método más didáctico.

Descripción de la propuesta. Esta propuesta plantea solucionar el problema del aprendizaje del braille mediante una retroalimentación asistida con el uso de un sintetizador de voz, el reconocimiento de voz y tarjetas NFC con relieves braille. En el aprendizaje del braille hay una etapa en la cual se empieza a enseñar el sistema de seis puntos y no puede soportarse en la propuesta por usar maquetas muy grandes para una etiqueta NFC. Sin embargo, la etapa dura como máximo 2 días y es para entender el braille. En la siguiente etapa se empieza el adiestramiento en el reconocimiento táctil de las letras, el cual sí es soportado por esta solución. Tarjeta por tarjeta, el usuario recibirá una interpretación de cada relieve que pase por el dispositivo sin necesidad de tener un profesor. Una vez que el usuario se sienta seguro de sus conocimientos, podrá empezar a interpretar los relieves y decir qué letra cree que le corresponde a la tarjeta. El sistema procesará la voz a texto y comparará con la información de la tarjeta NFC para dar el resultado con el sintetizador de voz.

Descripción de la arquitectura. Dentro de los componentes de la propuesta, se plantea utilizar un dispositivo móvil que soporte una cantidad mayor de 5 “touch-points” y que pueda reconocer NFC. Otro componente es el sintetizador de voz. Esta tecnología permite emitir mensajes de voz que serán usados para informar al usuario mientras ingresa un mensaje.

Internamente se estará realizando una concatenación de las letras y/o palabras que formarán el mensaje final. Near Field Communication (NFC) es una tecnología que permite reconocer tarjetas NFC mediante el uso de radio frecuencia de corto alcance. Estas pueden almacenar información para luego ser consultados mediante el uso de un Smartphone. El sistema operativo usa un tipo controlador de comunicación por proximidad (NFP DDI) para intercambiar información con estas tarjetas que serán parte fundamental del proceso de autoaprendizaje. Para comprobar lo aprendido en el sistema braille por parte del usuario, el proceso de autoevaluación debería soportar reconocimiento de voz integrado del sistema operativo.

Descripción del proceso. La forma de procesamiento de la propuesta está separada en tres procesos:

El primero consiste en almacenar la información de letras o palabras en tarjetas NFC con el relieve braille respectivo. Estas tarjetas se usarán como herramienta para entrenar al usuario en el reconocimiento táctil de los relieves con una retroalimentación del sintetizador de voz.



El segundo se basa en la comparación de lo que el usuario ha interpretado del relieve de una tarjeta con la interpretación que la tarjeta posee. El sistema resolverá la comparación y comunicará al usuario por medio del sintetizador de voz.



El tercer proceso se basa en el reconocimiento de toques en la pantalla del móvil, organizadas por regiones transversales. El reconocimiento de “touch-points” está relacionado a estas regiones, las cuales tienen asignada uno de los seis puntos braille. Una vez se detecte que un grupo de regiones ha sido presionado, se procesa la combinación y se obtiene la letra o palabra correspondiente a la combinación realizada.



Sistema de Reconocimiento de Señas con Intel Gesture Camera

Problema. Las personas sordomudas al no poder emitir sonidos y recibir mensajes por medio auditivo, tienen una gran dificultad para comunicarse. Por ello se hace uso del lenguaje de señas, el cual consiste en una serie de combinaciones de gestos

con las manos para enviar y recibir mensajes con otros que sepan este lenguaje. Sin embargo, no todas las personas lo conocen.

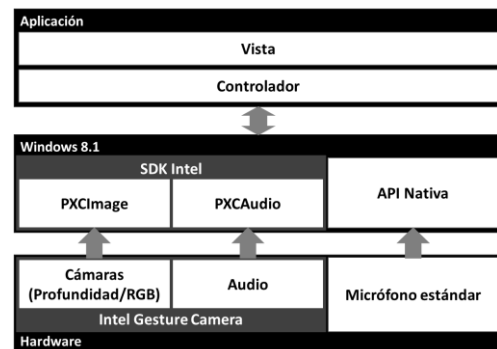
Trabajos similares. Existen varias soluciones para el reconocimiento de señas desde varios años atrás. Todas tratan, en diferentes formas, de solucionar la complejidad del reconocimiento de las señas, ya sea por medio del reconocimiento visual o por el uso de sensores. G. R. S. Murthy *et al.* [23] resumen las mejores y más completas soluciones relacionadas al reconocimiento de señas hasta el 2009 por visión o sensores. T. Starner y A. Pentland desarrollaron un sistema basado en el reconocimiento visual con una cámara 2D. El sistema soporta el sistema americano de lenguaje de señas y es capaz de reconocer 14 gestos en tiempo real para manipular pantallas de un ordenador y objetos de este [23]. Md. Hasanuzzaman, V. Ampornaramveth, M. A. Bhuiyan, Y. Shirai y H. Ueno presentaron un sistema de reconocimiento de gestos basados en la captura de imágenes. El algoritmo de reconocimiento de las manos está basado en una segmentación de las tonalidades de piel y la relación por patrones. Parte de este procesamiento combinaba otros algoritmos, el coeficiente de correlación y el mínimo calificado por distancia, para la interpretación de las señas. Este prototipo fue capaz de reconocer diez gestos con las manos y dos gestos faciales, los cuales eran retransmitidos a robots a través de una plataforma de software. Este sistema realiza el reconocimiento de patrones de imágenes con una cámara 2D [23]. Elena Sánchez - Nielsen, Luis Anton - Canalís y Mario Hernández desarrollaron un sistema de reconocimiento de señas con base a la captura rápida de segmentación para la obtención de la imagen de una mano en movimiento. Para el reconocimiento de gestos, hace uso de la métrica “Hausdorff” para la comparación de formas. El sistema obtuvo como resultado el reconocimiento de 26 gestos con las manos, con un 90% de precisión [23]. Zhou Ren *et al.* [24] crearon un sistema robusto de reconocimiento de señas con base al reconocimiento de imágenes en 3D. Este sistema usa la cámara de profundidad del Kinect para poder reconocer las manos del resto del cuerpo en adición al reconocimiento RGB. Para el reconocimiento de señas, se hace uso de la métrica “Finger-Earth Mover’s Distance” (FEMD), que compara la distancia entre las distribuciones de dos imágenes 2D, extrapoladas de los modelos 3D capturados por la cámara Kinect. La precisión del reconocimiento de señas de este sistema fue de 90.6% para cálculos aritméticos y lógicos de comparación de valores, de las señas interpretadas. Después de haber revisado estos trabajos, se concluye que, en general, estas soluciones tienen un desempeño aceptable para el reconocimiento de señas, pero sus limitaciones van relacionadas a su complejidad y a la limitada cantidad de gestos capaces de reconocer. Adicionalmente, en muchos casos, la correcta interpretación depende de la velocidad del movimiento de las manos. En la vida real, una persona que se comunica por medio de señas puede interpretar una cantidad de gestos con las manos a una velocidad relativamente rápida.

Descripción de la propuesta. Esta propuesta consiste en crear un sistema robusto para el reconocimiento de señas, habilitando la comunicación de personas que no practiquen este lenguaje. Para lograr dicha comunicación, se hará uso de un ordenador para la automatización de la lógica de interpretación de señas y una cámara Intel Gesture para conocimiento de dedos, pulgar y articulaciones. La ventaja de esta solución está en el uso de una plataforma que facilita el procesamiento de reconocimiento de los dedos con un nivel de precisión muy superior a otras propuestas. Asimismo, se requerirá del uso de un

synetizador de voz para completar el ciclo de comunicación una vez se tenga procesado el mensaje.

Descripción de la arquitectura. La arquitectura de esta propuesta tiene perspectivas a nivel de software y componentes de hardware de otros proveedores.

La perspectiva física está compuesta por dos componentes básicos: un ordenador con el sistema operativo Windows 8 y una cámara Intel Gesture. La primera capa de software es la de drivers y controladores de buses de información para la cámara y audio. La siguiente capa también tiene componentes de software, estos están disponibles para acceder a la información de las imágenes y sonidos capturados por los dispositivos electrónicos de la cámara. En esta se encuentra el módulo PXCImage, el componente de software que tiene las funcionalidades para acceder a la información capturada por los sensores de la cámara. Adicionalmente, se hará uso del módulo PXCAudio para la información que recibirá el usuario en casos de limitaciones auditivas. El software intérprete podrá acceder a las funcionalidades de detección de dedos por medio del módulo PXCImage. Este aplicará un algoritmo relacionado a los pulgares, los dedos y cercanía con el centro de la palma para las diferentes señas. Para interpretaciones más complejas como el uso de los brazos, se usarán las funcionalidades de reconocimiento de extremidades y articulaciones. La ventaja de esta nueva plataforma es la reducción del margen de error en el reconocimiento de las articulaciones de la mano y extremidades, sin necesidad de aplicar demasiadas capas lógicas. Esto permite mejorar y ampliar el vocabulario que el sistema puede reconocer, una desventaja que otros sistemas presentaban. Por otro lado, otro componente de software de esta arquitectura es el API nativo de Windows para poder enviar datos de salida a un sintetizador de voz. El software intérprete, el último componente de esta propuesta, debería estar diseñado bajo un concepto de centralización de funcionalidades.



Descripción del proceso. El proceso principal de esta solución consiste en el reconocimiento e interpretación de señas y emisión del mensaje procesado. Seguido de esto, puede recibirse información en forma de audio que será mostrada en pantalla, en caso la persona presente limitaciones auditivas. El proceso inicia cuando el usuario empieza a realizar gestos para la interpretación de estos. La información es capturada por las cámaras y dispuestas al sistema operativo para poder ser accedida por el controlador por medio del módulo de software PXCImage. Una vez que la información ha sido recibida por el controlador, se emplea la lógica para interpretar los puntos obtenidos. Los algoritmos deben estar orientados a interpretar la ubicación de cada uno de los dedos. La conjugación de posiciones de estos elementos son señas que deberán ser procesadas por un algoritmo

desarrollado en la instanciación. Posteriormente, el controlador enviará las cadenas de texto obtenidas de la interpretación para poder ser emitidas por el sintetizador de voz.

Suite para Personas Invidentes

Problema. Actualmente existen una considerable cantidad de aplicaciones móviles y muchas otras aplicaciones que están orientadas a brindar accesibilidad a las personas invidentes. Sin embargo, el problema de la navegabilidad para las personas invidentes se ha vuelto una dificultad que vale el esfuerzo evaluar soluciones.

Trabajos similares. La concepción de una solución integrada de varias aplicaciones sólo ha sido construida en Android. El nombre de esta aplicación es Mobile Accessibility y consiste en un set de botones donde se accede a funcionalidades básicas y necesarias para una persona con limitaciones visuales [9]: llamadas, contactos, SMS, alarmas, navegación web, calendario, email, geolocalización, configuración y el acceso a otras apps.

Descripción de la propuesta. La solución que se plantea desarrollar es una única interfaz para el usuario invidente con el objetivo de volver accesible el teléfono móvil. Esta interfaz tendría un fácil acceso para las siguientes aplicaciones: llamadas, contactos, SMS, alarmas, web, calendario, email, geolocalización, configuración, detector de billetes, lector de libros en formato DAYSI y el acceso a otras apps. Adicionalmente, se plantea agregar nuevas funcionalidades de ingreso de información por medio del reconocimiento de voz y el uso de un teclado braille, idéntico a la primera propuesta. La factibilidad de desarrollo está sustentada por la existencia de otras soluciones similares.

Descripción de la arquitectura. El primer componente es el sistema operativo que ofrecerá todos los componentes software necesarios para soportar las otras aplicaciones. El siguiente componente es la aplicación interfaz, que contendrá todas las aplicaciones mencionadas en la descripción y servirá como nuevo medio para interactuar con el dispositivo. Los siguientes componentes son las aplicaciones que se mencionaron en la descripción. Cada una de estas debe ser desarrollada e integrada con la interfaz.

Descripción del proceso. La forma como se realizarían las mensajerías será mediante la aplicación teclado braille integrado. Por otro lado, el “voice recognition” y “screen reader” serían los encargados de soportar el proceso de realizar llamadas, ya sea por medio de reconocimiento de voz o navegación por el listado de los contactos. Cabe mencionar que el “screen reader” ofrece una navegación asistida por medio de un sintetizador de voz al dictar las opciones por las que se navega. Para navegar por una ciudad o realizar una búsqueda rápida de un establecimiento comercial, se empleará la aplicación de Geolocalización con el uso de voice recognition y/o screen reader. Para la lectura de libros, se plantea desarrollar e integrar un software que permita la lectura de libros en formatos especiales como DAISY. Sin embargo, no deja la posibilidad de usar un screen reader para acceder a archivos pdf u otras extensiones para información digital. El resto de aplicaciones no se mencionarán por lo trivial de sus funcionalidades. En resumen, esta propuesta propone abarcar plataformas en donde otras aplicaciones similares no se han implementado. Adicionalmente, se agregan funcionalidades de gran aporte, como el lector de libros DAYSI, el

reconocimiento de billetes y la funcionalidad de ingreso de información con el sistema braille.

Tratamiento post-cirugía de Labio Leporino en Smartphone

Problema. El tratamiento post-cirugía de labio leporino es una intervención física que resuelve casi el 80% del problema del habla. Generalmente, estas sesiones son acompañadas por terapeutas que dan consejos con base a lo que van escuchando de los pacientes. Sin embargo, no permiten que el paciente pueda practicar autónomamente.

Trabajos similares. Actualmente no existen soluciones orientadas a un tratamiento post-cirugía de labio leporino. Sin embargo, hay aplicaciones que están orientadas al reconocimiento de fonemas para entrenamiento del habla de personas en sus primeros años académicos. Existen aplicaciones que soportan el reconocimiento de pronunciación para el entrenamiento de lectura de niños que están aprendiendo a hablar. Asimismo, existen varias decenas de aplicaciones que se basan en las funcionalidades de pronunciación como la aplicación “Pronunciation Guide” en Windows Phone [25]. Así como esta, existen otras aplicaciones similares que ofrecen aprendizaje de varios otros lenguajes que están con base a una comparación de textos tomando como fundamento un léxico [26].

Descripción de la propuesta. Esta propuesta plantea mejorar un proceso terapéutico para niños que se han sometido a una operación correctiva de labio leporino. La aplicación consistirá en la evaluación de su pronunciación mediante el reconocimiento de fonemas. El sistema identificará los defectos en la pronunciación y ofrecerá recomendaciones en base a dichos defectos. Las recomendaciones serían de tipo terapéutico y ofrecerán la retroalimentación necesaria para que el usuario pueda mejorar mientras es evaluado con nuevas palabras. El aporte esencial de la propuesta se centra en que la aplicación permita brindar las recomendaciones adecuadas en relación a los errores identificados.

Descripción de la arquitectura. La arquitectura de esta propuesta se enfoca exclusivamente en componentes de software. El primer componente es la API de Speech Recognition. Este brinda las funcionalidades necesarias para poder acceder a la información captada por el dispositivo móvil. La API posee funcionalidades que permite reconocer una palabra dependiendo de la configuración del léxico que tenga el sistema operativo. Esta última funcionalidad es muy importante para identificar los errores en pronunciación y dependerá del léxico del idioma configurado. El siguiente componente de software es la aplicación móvil desplegada en el sistema operativo del teléfono inteligente. La aplicación realizará el procesamiento de las palabras y retornará el resultado con las recomendaciones relacionadas a los errores identificados en la pronunciación. A pesar de ser una propuesta nueva en el enfoque del problema a solucionar, la arquitectura es muy similar a cualquier otra aplicación que usa las funcionalidades de reconocimiento de voz.

Descripción del proceso. El proceso principal de la aplicación es el reconocimiento de voz. Este inicia cuando se le indica al usuario, de forma textual, pronunciar una palabra. Las vibraciones en el aire son capturadas por componentes del hardware del dispositivo y dispuestas al sistema operativo. Posteriormente, el controlador recibirá la información por medio de la API de reconocimiento de voz para proceder a comparar la palabra captada con la palabra mostrada en un inicio. La API podrá darle una pronunciación a la palabra que se capturó si este

no aparece en el léxico configurado. Con esta “nueva” palabra se podrá identificar qué sílabas o letras el usuario no está pronunciando correctamente. Para estos errores, se brindarán las recomendaciones respectivas que usualmente se daría en una terapia de pronunciación. Una vez que se tenga el resultado, este se mostrará al usuario de forma textual en la interfaz.

4. CONCLUSIONES

Se ha planteado las propuestas innovadoras con base a la necesidad general de soluciones adaptativas. Estas propuestas son consideradas como aporte, ya sea por la aplicación de una nueva versión de la tecnología o la concepción de una nueva solución. La propuesta de reconocimiento de señas usa una cámara Intel con funcionalidades de reconocimiento de dedos y pulgares, solucionando la limitación de comunicación de personas sordomudas. La solución para la terapia de recuperación post-cirugía de labio leporino es una propuesta completamente nueva e innovadora. Finalmente, la propuesta de teclado braille con funcionalidades de autoaprendizaje y la suite de aplicaciones de apoyo para personas invidentes son mejoras de sistemas similares para problemas de comunicación.

5. REFERENCIAS

- [1] Instituto Nacional de Estadística e Informática, 2013. Instituto que norma, planea, dirige, coordina, evalúa, supervisa y proporciona estadísticas oficiales del Perú. (<http://www.inei.gov.pe>) Consultado el 20 de agosto de 2013.
- [2] GRS Murthy and RS Jadon. A review of vision based hand gestures recognition. *International Journal of Information Technology and Knowledge Management*, 2(2):405–410, 2009. W.-K. Chen, *Linear Networks and Systems*. Belmont, Calif.: Wadsworth, pp. 123-135, 1993 (Book style).
- [3] Apple, Apps for Visually Impaired. (http://appadvice.com/applist_ipad_client_view/apps-for-the-visually-impaired) Consultado el 08 de junio de 2014.
- [4] Apps4Android, Best 10 Android Apps for the Visually Impaired. (<http://www.apps4android.org/?p=4107>) Consultado el 08 de junio de 2014.
- [5] Google Play, Google TalkBack. (https://play.google.com/store/apps/details?id=com.google.android.marvin.talkback&hl=es_419) Consultado el 08 de junio de 2014.
- [6] Apple, iOS A Wide range of features for a wide range of needs. (<https://www.apple.com/accessibility/ios/>) Consultado el 08 de junio de 2014.
- [7] Google Play, Voice Command (<https://play.google.com/store/apps/details?id=com.ToxicBakery.apps.voicecommand>) Consultado el 08 de junio de 2014.
- [8] Google Play, Blind Navigator (<https://play.google.com/store/apps/details?id=com.blindnavigator>) Consultado el 08 de junio de 2014.
- [9] CodeFactory, Mobile Accessibility (<http://www.codefactory.es/en/products.asp?id=415>) Consultado el 08 de junio de 2014.
- [10] Frey, B., Southern, C., Romero, M. 2011 BrailleTouch: Mobile Texting for the Visually Impaired. En: HCCI 2011, LNCS 6767, págs. 19-25. Springer-Verlag, Berlin Heidelberg
- [11] Mascetti, S., Bernareggi, C., Belotti, M. 2011 TypeInBraille: Quick Typing on Smartphones by Blind Users. En: Università Degli Studi Di Milano: Milán, Italia.
- [12] Olviera, J., Guerreiro, T., Niclau, H., Jorge, J., Gonçalves, D., 2011. BrailleType: Unleashing Braille over Touch Screen Mobile Phones. En: HCCI 2011, págs. 100-107, Springer Berlin.
- [13] Azenkot, S., Wobbrock, J. O., Prasain, S., Ladner, R. E. 2012 Input Finger Detection for Nonvisual Touch Screen Text Entry in Perkinput. En: Proc. Graphics Interface. Canadian Information Processing Society, Toronto.
- [14] Wccf Tech (<http://wccftech.com/kinect-sports-rivals/>) Consultado el 08 de junio de 2014.
- [15] PROPOLI, Sistematización de la intervención de PROPOLI en discapacidad: Lecciones aprendidas y recomendaciones para futuras intervenciones (<http://www.bvcooperacion.pe/biblioteca/bitstream/123456789/2019/1/BVCI0001650.pdf>) Consultado el 08 de junio de 2014.
- [16] SpeechBuddy, Top 4 Cleft Palate Speech Therapy techniques (<http://www.speechbuddy.com/blog/speech-therapy-techniques/top-4-cleft-palate-speech-therapy-techniques/>) Consultado el 08 de junio de 2014.
- [17] Cincinnati Children’s, Speech Therapy for Cleft Palate or Velopharyngeal Dysfunction (VPD) (<http://www.cincinnatichildrens.org/assets/0/78/759/781/65e90133-9243-4926-a065-8a97951944fb.pdf>) Consultado el 08 de junio de 2014.
- [18] Microsoft 2014 ([http://msdn.microsoft.com/en-us/library/hh378451\(v=office.14\).aspx](http://msdn.microsoft.com/en-us/library/hh378451(v=office.14).aspx)) Consultado el 08 de junio de 2014.
- [19] Muhammad Asad, Charith Abhayaratne, 2013. Kinect Depth Stream Pre-Processing for Hand Gesture Recognition.
- [20] 23Kinect 2014 (<http://123kinect.com/everything-kinect-2-one-place/43136/>) Consultado el 09 de junio de 2014.
- [21] Intel® Perceptual Computing SDK Tutorials (<https://software.intel.com/en-us/articles/intel-perceptual-computing-sdk-tutorials>) Consultado el 09 de junio de 2014.
- [22] Intel Development Resources ([https://software.intel.com/en-us/search/site/?f\[0\]=bundle%3Aarticle](https://software.intel.com/en-us/search/site/?f[0]=bundle%3Aarticle)) Consultado el 09 de junio de 2014.
- [23] G. R. S. Murthy, R. S. Jadon. A Review of Vision Based Hand Gestures Recognition, 2009.
- [24] Zhou Ren, Jingjing Meng, Junsong Yuan y Zhengyou Zhang. Robust Hand Gesture Recognition with Kinect Sensor, 2013.
- [25] Pronuntiation Guide (<http://www.windowsphone.com/en-us/store/app/pronunciation-guide/05f729b1-5279-e011-986b-78e7d1fa76f8>) Consultado el 28 de junio de 2014.
- [26] Learn Pronuntiation English (<http://www.windowsphone.com/en-us/store/app/learn-pronunciation-english/0de5e6a9-3d84-4e0e-bf68-1dd3d9aa26ca>) Consultado el 28 de junio de 2014.
- [27] Developing applications using Intel® Perceptual Computing Technology (<https://software.intel.com/en-us/articles/developing-applications-using-intel-perceptual-computing-technology>) Consultado el 28 de junio de 2014.