

  Christian Quintero¹;
 Diego Roa²

¹ Universidad Militar Nueva Granada, Bogotá - Colombia,
christian.quintero@unimilitar.edu.co

² Universidad Militar Nueva Granada, Bogotá - Colombia,
u1201198@unimilitar.edu.co

ISSN-p: 0123-7799
ISSN-e: 2256-5337

Vol. 25, nro. 53, e2131, 2022

Recibido: 17 agosto 2021
Aceptado: 21 febrero 2022
Disponible: 24 marzo 2022

©Instituto Tecnológico Metropolitano
Este trabajo está licenciado bajo
una Licencia Internacional
Creative Commons Atribución
(CC BY-NC-SA)



Cómo citar / How to cite

C. Quintero; D. Roa, “Aproximación musical a través de la interpretación gestual”, *Tecnológicas*, vol. 25, nro. 53, e2131, 2022.
<https://doi.org/10.22430/22565337.2131>

Resumen

Este trabajo tuvo como propósito explorar entornos virtuales desde un sistema interactivo que permitiera interpretar los gestos manuales hacia una aproximación musical. La metodología implementada se realizó desde un enfoque cuantitativo, donde se establecieron tres fases de recolección de datos: secuencialidad de gestos, signos de Curwen y aproximación a la realidad. Este estudio fue dividido en tres etapas longitudinales conceptualización, diseño e implementación y evaluación para ampliar el espectro en la interpretación de los gestos. El método Kodály implementado en este trabajo, el cual ha sido adaptado en Colombia, utiliza los gestos manuales para asimilar con mayor rapidez conceptos musicales. La evaluación fue desarrollada con una muestra poblacional de 15 usuarios en cada fase para la recolección de datos. Los resultados se identificó una rápida adaptación en la interpretación de los gestos en relación con las notas musicales. Los tiempos y errores obtenidos permitieron medir aspectos en la ejecución de los gestos. Se adaptó un conjunto de gestos de la mano a un entorno virtual en el que los usuarios pueden tocar la melódica mediante la interpretación gestual, lo que demostró que el reconocimiento de gestos puede ayudar hacia una aproximación musical. Finalmente, los usuarios lograron un alto nivel de comprensión gracias al uso de movimientos naturales de la mano, lo cual fue un proceso significativo para los usuarios en una primera experiencia mediante el reconocimiento de gestos manuales.

Palabras clave

Reconocimiento de gestos, sistemas interactivos, entornos virtuales, lenguaje natural.

Abstract

This article describes a virtual environment of an interactive computer system that can interpret hand gestures as pitches in a tonal scale. The methodology implemented here is based on a quantitative approach with three data collection stages: sequence of gestures, Curwen's hand signs, and approximation to reality. This study was divided into three general stages conceptualization, design and implementation, and evaluation, to expand the scope of hand gesture recognition. The Kodály method implemented in this paper, which was adapted in Colombia, uses hand gesture to help students better understand musical concepts. The evaluation was conducted with a population sample of 15 users in each data collection stage. The results show that users adapted quickly to hand gesture recognition and its relationship with musical notes. Times and errors obtained in the tests were used to measure aspects of execution. A set of hand gestures was adapted to a virtual environment where users can play melodia by gestural interpretation, which showed that gesture recognition can help music instruction. Finally, users achieved a high level of understanding thanks to the use of natural hand movements, and it was a meaningful process for them because it was their first experience with a hand gesture recognition.

Keywords

Gesture recognition, interactive systems, virtual environments, natural languages.

1. INTRODUCCIÓN

La teoría musical de Zoltán Kodály como metodología de la enseñanza permite aprender la estructura de la música por medio de cantar y gestos manuales [1], [2]. Este pensamiento pedagógico permite que la música pueda ser direccionada hacia todas las personas.

Asimismo, las personas han tomado la costumbre de realizar gestos que simulan los movimientos que los artistas hacen en las presentaciones mientras tocan sus instrumentos, llegando a involucrar varias partes del cuerpo, como los dedos, manos, brazos o pies [3]. Los movimientos de la mano y los dedos que se realizan en la ejecución del piano tienen un nivel de destreza bajo, pero cuando se relacionan de una forma interpretativa de música el desafío y habilidad aumentan de forma significativa. Por este motivo hacer un movimiento mímico de ellos es sencillo para las personas, permitiendo así representar un piano de aire, este término se describirá a través de las secciones de este texto.

Los diferentes tipos de gestos interpretan y transmiten un mensaje, los artistas expertos, quienes tocan un instrumento real, inducen este mensaje. De igual forma, pueden percibir con mayor facilidad el resultado de estos movimientos, hasta lograr una denominación como instrumentos de aire [3]. Además, esto se puede considerar como una expresión desde las personas en como perciben e imaginan la música, y estudiar las relaciones entre estos gestos y sonidos puede contribuir a nuestro conocimiento de cómo los gestos ayudan a estructurar nuestra experiencia musical.

Por otra parte, Kodály adaptó los gestos manuales de Curwen para aplicarlos en su método de enseñanza [4]. Los signos manuales de Curwen representan visualmente las notas de la escala y sus relaciones entre sí. Los novatos de la música asocian los signos manuales a números, esto permite la facilidad de recordar las notas por medio de un conjunto de signos, más que intentar recordar un pentagrama desde el principio, lo que se convierte en un método de aprendizaje más cómodo para el estudiante [1].

El cuerpo en la cibercultura ha establecido un contexto general de argumentos sobre la manifestación en la teoría social y cultural. En el ciberespacio, en una desmaterialización del cuerpo en espacios virtuales permite ocupar una ubicación ambivalente, este puede expresar una idea de dejar atrás la carne, deshacerse del cuerpo y cargar la conciencia como datos en áreas de la cibercultura contemporánea; otra forma se estableció junto con la realidad de las experiencias encarnadas del uso de la computadora [5].

La tecnología es una forma de llegar a las futuras generaciones, para que estas se interesen en la música y así ampliar la cultura musical que hay dentro de país. Esto se puede denominar *usos subculturales del ciberespacio*. Esto refiere a formas de usar tecnologías informáticas que subvierten de alguna manera las normas sociales o las formulaciones dominantes de para qué sirve la tecnología, las que podrían ser opuestas a las estructuras políticas y económicas corporativas que dominan cada vez más la forma del ciberespacio [5], de la misma manera que podrían usar otras formas de comunicación y aquellos que señalan una relación expresiva con la tecnología a través de actividades subculturales.

La contribución de este artículo es la adaptación de un conjunto de gestos manuales en un entorno virtual hacia una aproximación musical orientado a gestos en un piano de aire.

En tal sentido, el artículo presenta diferentes secciones que permite establecer los conceptos y procesos realizados. Para esto se presenta una sección de estado de arte, la cual relata diferentes trabajos relacionados al campo de investigación que permita resaltar criterios de otros autores desde el aprendizaje y la tecnología en interfases gestuales orientadas a la música. Las siguientes tres secciones presentan conceptos que permiten aclarar términos y definiciones. En la metodología permite enunciar la estructura diseñada para el desarrollo de la investigación, presentando las fases en la implementación de pruebas

y análisis. Cabe aclarar que cada fase se planteó desde una conceptualización, diseño e implementación y evaluación. Esto con el fin de poder contrastar los resultados obtenidos en las diferentes fases. La sección de método indica el enfoque, la muestra poblacional, los indicadores y describe los procedimientos en la recolección de datos. La sección de resultados presenta de forma gráfica los datos recopilados por los usuarios que implementaron las pruebas diseñadas. Finalmente, las dos últimas secciones, discusión y conclusiones, describen la interpretación de los datos obtenidos en las diferentes pruebas, esto permite inferir diferentes hallazgos sobre los análisis de datos.

2. ESTADO DEL ARTE

La interacción humano-computador (HCI, por sus siglas en inglés) viene aproximadamente desde 1960, comenzando, en la mayoría de las áreas, desde la investigación universitaria. Sin embargo, las interfaces gestuales fueron las últimas en ser exploradas [6], siendo que estas tienen en cuenta el tiempo de aprendizaje, ergonomía e intuición por parte del usuario [7].

Los gestos manuales se han implementado para controlar la reproducción de música (pausar, detener, reproducir) por medio de cámaras web [8]. Además de este acercamiento al uso de los gestos para la manipulación de la música, también se han estudiado los gestos realizados al tocar un clarinete real, con el fin de identificarlos y aplicarlos a una interacción humano-computador [9]; incluso se han estudiado los gestos que no son obvios y la influencia que estos tienen dentro del sonido final en los instrumentos de aire [10].

Otras tecnologías desarrolladas presentan el seguimiento del movimiento de los dedos de las manos mientras se toca una guitarra, esto obtiene un puntaje orientado a encontrar el camino más corto entre todas las posiciones posibles de los dedos [11]. Por otro lado, el proyecto Cyber Composer es un sistema que ayuda a crear música con los gestos de la mano, tanto a músicos como a aquellos que no lo son. No obstante, es realizado por un método no invasivo, ya que requiere de unos guantes diseñado para el reconocimiento de los gestos [12].

El uso de sensores biométricos también ha sido implementado para la recolección de señales de electromiografía del antebrazo como el dispositivo *Myo gesture*, puesto que contienen información relevante sobre la actividad eléctrica producida por el proceso neuromuscular. Por lo tanto, es posible analizar esas señales para un reconocimiento de gestos a partir de los movimientos de contracción o relajación de dedos, mano, brazo y antebrazo. Por ejemplo, el trabajo realizado por [13] presentan una comparación de tres modelos computacionales para el reconocimiento de gestos, para lo cual interpretaron un conjunto de ocho gestos, concluyendo que existe un mejor rendimiento computacional a las métricas de sensibilidad y precisión mediante el modelo “bolsa de características”.

Ahora bien, retomando la interpretación de gestos en el campo musical, se presenta el desarrollo de una aplicación en realidad aumentada denominada “FunPianoAR”, que pretende activar los intereses de los usuarios para mejorar la experiencia del aprendizaje del piano. Asimismo, la aplicación en realidad aumentada presenta una forma novedosa de mostrar elementos tridimensionales en el teclado del piano real como guía para al usuario en el uso de la aplicación [14]. Otro trabajo orientado al instrumento del piano, abordado desde el concepto de la gamificación, propone una técnica para superar la falta de motivación intrínseca por medio de una realidad aumentada para reducir la carga cognitiva, y un sistema de retroalimentación continua, para hacer el proceso de aprendizaje menos extenuante y más motivador [15].

En la misma línea de tecnologías para la interpretación de gestos y orientado en el piano como instrumento musical, se presenta el estudio de [16] con el dispositivo de alta precisión en tiempo real, Leap Motion, el cual desarrolla un piano virtual permitiendo así a los usuarios tocarlo en tiempo real moviendo sus manos por encima de este dispositivo. No obstante, el anterior estudio no presenta un análisis de la captura de los datos y solo realiza un grupo de gestos. En contraste, los resultados que se presentan en este artículo extienden el conjunto de gestos implementados por [16].

Finalmente, la propuesta de un sistema de realidad virtual (RV) de piano utiliza dos sensores Leap Motion para el seguimiento de la mano de amplio alcance [17]. Sin embargo, esta propuesta no se centra en la RV, sino que la extiende con elementos tangibles y otros sensores, obteniendo así resultados experimentales que demuestran que el sistema propuesto de piano RV proporciona a los usuarios una gran experiencia al tocar este, siendo además útil para el aprendizaje de este instrumento.

3. REPRESENTACIÓN DE GESTOS MANUALES

Los tipos de gestos que hacen los intérpretes mientras se toca una pieza musical pueden variar de acuerdo con el tipo de instrumento que se utiliza. En el caso de los artistas, quienes tocan un instrumento real, se puede percibir con mayor facilidad el resultado de estos movimientos; sin embargo, los oyentes denominados como espectadores o intérpretes tácticos, quienes realizan movimientos tocando instrumentos que se pueden denominar como instrumentos de aire [3], no tienen un modo de percibir realmente lo que están realizando, solo representan una imitación del movimiento de forma empírica.

Es necesario aclarar que no todos los movimientos que hace un artista en medio de la presentación repercuten en un sonido, sino que pueden ser los llamados gestos auxiliares de la presentación, así que se deben tener en cuenta los movimientos que realmente significan un sonido dentro de los demás. Así como los movimientos de la mano cuando se toca un piano, que se hacen para posicionar los dedos sobre las teclas correctas del piano, estos se pueden considerar gestos productores de sonido, ya que se hacen específicamente para realizar un sonido [3], [18].

Algunos pianistas experimentados dicen que con solo escuchar una pieza de música evoca el movimiento involuntario de sus dedos, y realizar los movimientos respectivos que se ajustan a la pieza de música [19]. Así podemos ver que los instrumentos de aire no se acomodan solo a los oyentes, sino también para los músicos expertos. Sin embargo, para llegar a ser expertos en música, sin importar cuál sea el instrumento en el que se especialicen, deben pasar por etapas de estudio musical y luego refinar sus conocimientos desde la práctica.

3.1 Método Kodály y su adaptación en Colombia

Zoltán Kodály era un compositor húngaro que decidió dejar de componer para dedicarse a la enseñanza de esta, recopilando la música cultural de su país al darse cuenta de que había analfabetismo musical, incluso, en las personas cultas de la música [2]. Basó su método de enseñanza en canciones y cómo un niño podía aprender el lenguaje con tan solo escucharlo y reproducirlo constantemente, entonces las personas podían aprender la música interpretando de forma oral las canciones [20].

Los principios básicos del método Kodály están centrados en la premisa que “la música pertenece a todos, no es un juguete para unos pocos seleccionados. La música es un alimento

espiritual para todo el mundo” [20]. Por esto, Kodály sugiere que la enseñanza musical empiece con música folclórica. Sin embargo, este método debe ser adaptado y no adoptado, ya que se basa en la música folclórica de cada región, y la diferencia entre la música húngara y latina hace que si no se adapta, este terminará fracasando como método de enseñanza musical [1].

Aunque Kodály sugiere la utilización de la música folclórica, no es necesario que se comiencen por estas canciones “pesadas” y pasadas de moda. Se puede comenzar por música más allegada a los estudiantes, como rock, pop o heavy metal, entre otras, y llevar la instrucción hacia un tipo de música que le sea más fácil de comprender [1], [2].

Ahora bien, el Método Kodály adaptado en Colombia presenta una vinculación para todas las personas, esto infiere que no hay una separación de personas con dones musicales o personas sin apreciación musical. El instrumento principal, que es la voz, utiliza la música de cada cultura, y a partir de esta se entiende la música de otras culturas y de todos los elementos que la comprenden. El método no tiene un orden preestablecido, va de acuerdo con el desarrollo del niño y la música de su entorno. Esta adaptación fue iniciada por el Grupo Kodály Colombia de la Pontificia Universidad Javeriana, a partir de la recopilación de canciones populares tradicionales, rondas, rimas y juegos infantiles en un contexto Colombiano [1], [21].

Este método de enseñanza de música utiliza, como se dijo antes, el canto, en conjunto con gestos manuales que ayudan a que los estudiantes relacionen con mayor facilidad las notas a los movimientos que hacen con su mano, aunque se hace necesario que el repertorio musical de los estudiantes contenga muchas canciones con patrones melódicos similares para que pueda conectar los patrones de una canción con los de otra. De otro modo, con patrones aleatorios de música, esto no serviría debido a la dificultad de la comprensión de los patrones de notas. Además, incluye los gestos desarrollados por Curwen [4] para dar una orientación a los estudiantes.

Los signos manuales de Curwen representan visualmente las notas de la escala y sus relaciones entre sí. Los novatos de la música asocian los signos manuales a números, lo que facilita recordar las notas por medio de un conjunto de signos, más que intentar recordar un pentagrama desde el principio, lo que se constituye en un método de aprendizaje más cómodo para el estudiante [1].

4. INTERACCIÓN HUMANO-COMPUTADOR

Hoy en día, las formas de comunicación se extienden en nuevas técnicas que permite relacionar la interacción hombre-máquina, la relación que debe existir entre el usuario y la máquina, entendiendo por máquina el entorno, interfaz, dispositivo o herramienta que pueda ser utilizada por humanos. Evidencia los conceptos, principios y técnicas por medio de los cuales se busca minimizar las cargas de percepción para el usuario. Este proceso de diseño debe lograr un resultado que demuestre un manejo fácil e intuitivo. Habla del proceso centrado en el usuario por medio de ciertos conceptos: eficacia, eficiencia, usabilidad [22].

Es importante enfocarnos en la interacción no verbal para generar el sentido de realismo en las comunicaciones con personajes virtuales. [23] hablan sobre el rol del comportamiento no verbal en el diálogo social, donde el objetivo es crear agentes conversacionales computarizados, estos deben disponer de habilidades y competencias humanas.

De igual forma, en la comunicación humana existe un amplio rango de canales comunicativos, pues hay una mezcla en proporciones variables de lo verbal y lo no verbal, las

cuales son representadas por el cuerpo mediante gestos, miradas e interacciones afectivas como el reconocimiento de expresiones faciales y emociones.

Estos analizan cómo los usuarios deben ser capaces de interactuar naturalmente con los computadores. En la comunicación cotidiana entre humanos es indispensable utilizar el lenguaje vocal y el corporal en la gestualidad para expresar emociones, estado de ánimo, actitudes y atención, y para así lograr que otros entiendan el sentido de la comunicación.

Según [24], es conveniente que el computador reconozca los gestos del usuario y otras expresiones.

4.1 El gesto

El gesto puede describirse como un aspecto fisiológico que puede ser consciente e inconsciente, y se exterioriza a través de algún movimiento del cuerpo y que se interpreta por un observador, como lo manifiesta [25]. De este modo, estos gestos se convierten en unas normas fijas e invariables que se transforman según su uso y la propósitos e intenciones entre sujetos que lo utilizan.

La interpretación de gestos en el aire pueden ser un acercamiento inicial con la música, este nos puede permitir creer que existen vínculos importantes entre el sonido musical y los gestos, pero esto requiere de un estudio serio. Por otro lado, el proceso de lectura del gesto y la interpretación implica inicialmente la percepción de un cambio de estado de la persona observada que datan de sus emociones, pensamientos y sentimientos. Lo anterior se da gracias a una serie de convenciones culturales que permite interpretarlos como un todo.

Quiere decir, entonces, que la interacción gestual entre sujetos, o entre los sujetos y objetos, en este caso la HCI, se desarrolla, por lo general, en lugares y momentos específicos convencionalizados por la cultura [26].

El reconocimiento de gestos es un tema de las ciencias de la computación y la tecnología del lenguaje que tiene como fin interpretar gestos humanos a través de algoritmos matemáticos. Estos pueden ser comúnmente originados por la cara y las manos. Los enfoques actuales incluyen reconocimiento facial y el otro estudiado es el reconocimiento de gestos de la mano como enfoque en esta investigación. Asimismo, la recolección de datos puede ser obtenida por técnicas de visión por cámaras, dispositivos móviles, procesamiento de imágenes y sensores de luz.

Finalmente, el gesto, más que un medio de comunicación auxiliar al lenguaje desde una forma corporal es un pensamiento corporeizado que es dinámica, y se puede interpretar de múltiples maneras como apoyo a funciones propias del diseño. En la misma línea, las expresiones gestuales no son naturales ni innatas ni fáciles de aprender, comprender o recordar. Sin embargo, estos pueden ser adaptados a condiciones particulares de un contexto [15], [27]. En [28] se plantean dos escenarios i) adaptación para el reconocimiento de gestos culturalmente aceptados-apropiados, y ii) proponer gestos que puedan ser aprendidos y apropiados por las personas, siendo una de sus conclusiones que las formas de interacción de los humanos con herramientas, máquinas e instrumentos musicales a través de gestos es innata.

4.2 Dispositivo de interacción gestual

La orientación objetual hacia la interacción gestual requiere de un dispositivo electrónico que permita la detección de las manos, además de que este pueda obtener una precisión de los datos capturados para poder analizar y representar el reconocimiento de gestos. Por anteriores consideraciones, se escogió la herramienta Leap Motion Controller (LMC), la cual

tiene una precisión no mayor a 0.2 mm [29]. La distancia que regresa el dispositivo es en milímetros y tiene un tiempo de reacción en microsegundos capaz de reconocer que mano es la que está evaluando (derecha o izquierda), cada dedo por separado, y las articulaciones que estos tienen. Esto lo hace por medio de sensores ópticos y luz infrarroja, de modo que puede hacer el cálculo de profundidad, movimiento y rotación que realizan las manos [30].

Mediante las funciones del LMC, los gestos a realizar serán posibles de identificar con mayor certeza, logrando un sistema más confiable, siempre y cuando se encuentre en un ambiente con buena iluminación que permita que los sensores ópticos del LMC puedan hacer el contraste de las siluetas de las manos.

5. MÉTODOS

Esta sección presenta los métodos implementados para dar alcance al propósito de explorar entornos virtuales desde un sistema interactivo que permita interpretar los gestos manuales hacia una aproximación musical. Para esto, los signos manuales de Curwen son adaptados en un entorno virtual para ser interpretados en tiempo real por los usuarios.

Asimismo, se extendió esta fase de gestos del método Kodály [20] por medio de tres fases, enfocándose en la secuencialidad de gestos, los gestos manuales diseñados por Curwen y una aproximación a tocar el instrumento real a través de los gestos del piano de aire.

Todas estas fases se evaluaron por medio de la misma secuencia de notas para que los resultados se pudiesen agrupar y valorar de tal forma que se pudiera ver la diferencia entre los gestos. A parte, la retroalimentación, dentro de la aplicación prototipo que recibían los usuarios, era auditiva cada que realizaban uno de los gestos establecidos y visual, que le indicaba qué gesto tenía que realizar y cuál era el gesto subsiguiente.

Teniendo lo anterior en cuenta, cada fase se dividió en tres etapas. La conceptualización del problema, diseño e implementación de una solución, y la evaluación. En la conceptualización se estudió el problema que presentaba cada fase y se consultó la teoría ya existente para comenzar a utilizarla para proponer una solución.

En diseño e implementación se utilizó la teoría consultada y refinada en la primera parte y se procedió a buscar los mejores gestos posibles para la interpretación de las notas musicales y se pasó a codificarlas para que los usuarios pudiesen interactuar, esto con el fin de verificar que los gestos fueran fáciles de entender para los usuarios.

La muestra escogida por oportunidad con quince personas fue en edades entre los 20 y los 25 años, sin distinción de género. Los indicadores en la recolección de datos se establecieron en la captura del tiempo y la cantidad de errores que realizaban por cada nota antes de pasar a la siguiente nota.

El enfoque de la recolección de datos es de origen cuantitativo [31], para esto se estableció la aplicación de pruebas divididas en las tres fases. La Fase I, dividida en dos partes; en la primera, el usuario tenía que seguir la secuencia de la escala musical (do, re, mi, fa, sol, la, si) de forma repetida; en tanto, en la segunda, el usuario debía seguir la secuencia de notas para el intervalo musical “Cumpleaños feliz”, realizando los gestos del piano de aire. La Fase II se implementó de forma similar que, en la Fase I, una primera parte de notas en secuencia y una secuencia de notas para la misma canción, sin embargo, se diferenciaba por los gestos, ya que esta fase se centraba en los gestos de Curwen modificados. En la Fase III se aplicó la prueba únicamente con la secuencia de la canción seleccionada “Cumpleaños feliz” (do, do, re, do, fa, mi, do, do, re, do, sol, fa, do, do, do, la, fa, mi, re, si, si, la, fa, sol, fa), realizando los gestos del piano de aire para la octava central y, para las cuatro octavas restantes, se modificaron los gestos del piano de aire para adecuarse a los espacios definidos

para las octavas y las notas. El intervalo musical seleccionado fue considerado como símbolo cultural, puesto que la gran parte de la población en Colombia reconoce el ritmo y melodía, por lo tanto, se facilitaba seguir su ritmo mientras realizan los gestos manuales de cada nota.

En el protocolo de la recolección de datos los usuarios debían realizar las pruebas en un orden establecido, con el fin de inferir cómo los usuarios evolucionaban en el desarrollo de los gestos conforme transcurrían las pruebas. Para esto la captura de los datos se realizó de forma automática mediante archivos de registro generados por cada fase. Finalmente, son etiquetados y almacenados para su posterior análisis.

En la Fase III no se recolectaron datos de forma automática en la interpretación de los gestos, ya que el manejo espacial para la definición de las octavas no era el suficiente y los usuarios no podían manejar bien este espacio. Es decir, era difícil para los usuarios el mantener sus manos dentro del espacio definido para poder ejecutar las notas necesarias para cumplir con la ejecución de la nota que pedía la aplicación. A pesar de este inconveniente los usuarios aplicaban la prueba. Sin embargo, se recolectaron otros tipos de datos cualitativos mediante el relato de la experiencia generada en el desarrollo de la prueba de forma general.

5.1 Fase I: secuencialidad de gestos

Conceptualización: inicialmente se buscaba que el usuario identificará los gestos dentro de la aplicación, por ello se pasó a diseñar los gestos para un piano de aire con solo una octava.

Para cumplir con este fin, se revisaron videos de la interpretación de piezas musicales hechas por músicos profesionales.








Los gestos simples, como bajar un dedo para tocar la tecla, de este modo, reproducir la nota musical fueron copiados dentro de los gestos utilizados para la aplicación, con la variación que la nota a reproducir depende directamente del dedo que realice el gesto, mas no de la posición en que se encuentre ubicada la mano. Dicho de otra manera, las notas de la octava se asocian a un dedo en particular.

Aunque el dedo pulgar se utiliza dentro de los movimientos propios para tocar el piano, se descartó en la primera fase del proyecto, con el fin de obtener una mayor precisión del gesto por medio del LMC. De este modo, los dedos a utilizar son: meñique, anular, corazón e índice. Por otra parte, ya que las notas musicales son siete (do, re, mi, fa, sol, la, si), la distribución de las notas no sería equitativa. Se agregaron cuatro notas musicales a la mano derecha (do, re, mi, fa) y tres a la izquierda (sol, la, si).

Diseño e implementación: una vez establecidas las notas que irían asociadas a cada mano, se pasó a anexarlas al gesto correspondiente. Como se manifestó anteriormente, las notas dependen del gesto, así que lo más cómodo para el usuario fue utilizar los dedos se forma secuencial. Las notas van desde la mano derecha la nota *do* asociada con el dedo índice, y el dedo meñique con la nota *fa*, mientras la mano izquierda tiene la nota *sol* en el dedo índice y la nota *si* en el dedo anular, teniendo en cuenta las notas y dedos intermedios (véase Tabla 1).

Evaluación: esta etapa se formuló en dos partes, la primera parte consistía en una secuencia ordenada de notas donde el usuario tendría que hacer los gestos pertinentes. Esto se propuso con el fin de que el usuario pudiese aprender los gestos del piano de aire de forma intuitiva. En segunda parte se les pidió a los usuarios que siguieran una secuencia de notas, correspondiente a una canción. Aquí se buscó encontrar la cantidad de veces que el usuario hacía el gesto de forma incorrecta y el tiempo que se demoraba en hacer correctamente el gesto.

Tabla 1. Correspondencia entre gestos secuenciales y notas musicales. Fuente: elaboración propia.

Dedo	Nota	Gesto
Índice derecho	Do	
Corazón derecho	Re	
Anular derecho	Mi	
Meñique derecho	Fa	
Índice izquierdo	Sol	
Corazón izquierdo	La	
Anular izquierdo	Si	

5.2 Fase II: Signos de Curwen















Conceptualización: en la aplicación del método Kodály los gestos son el recurso que permite aprender las bases musicales para cualquier instrumento, pero, estableciendo un medio instrumental como el piano, donde se tomaron los gestos manuales de Curwen y estos pueden ser realizados con una mano, esto permite facilitar el aprendizaje en la ubicación gestual de la interpretación.

Diseño e implementación: los gestos manuales de Curwen fueron tomados por Zoltán Kodály por su fácil aprendizaje y reproducción. Fue necesario adaptar los gestos, puesto que el sistema presentaba una confusión al momento de reconocer aquellos debido a la similitud

entre gestos o a la posición de los dedos, lo que permitía mejorar la captura de datos con los gestos (véase Tabla 2).

Evaluación: se contrasta con los resultados obtenidos en la Fase I, esto permitió definir el uso de la misma secuencia de notas.

Tabla 2. Comparación de gestos de Curwen adaptados y gestos originales. Fuente: elaboración propia.

Nota	Adaptado/Original	
Do		
Re		
Mi		
Fa		
Sol		
La		
Si		

5.3 Fase III: Aproximación a la realidad

Conceptualización: en busca de llevar la experiencia del piano de aire a un punto más próximo a la realidad, se implementaron más octavas que en la Fase I y II.

Los pianos de cola reales tienen siete octavas y media, sin embargo, no se disponía del espacio suficiente para implementar las octavas de un piano real, así que se procedió a aplicar únicamente cinco octavas.

Diseño e implementación: se definieron las octavas mediante una codificación en el aplicativo; asimismo se diseñaron y asociaron los gestos que se utilizarían dentro de la Fase III.

Lo primero fue definir las áreas que utilizarían las octavas, dependiendo del ángulo de visión y la distancia máxima del LMC. Se definieron cuatro áreas con un espacio de

interacción de dimensiones iguales y un área central capaz de contener las dos manos (véase Figura 1).

Los gestos se adaptaron de acuerdo con las que se tuvieron que adaptar a la octava que las manos estuvieran utilizando. A la octava cuarta, ubicada en el centro, se le implementaron los mismos gestos que fueron utilizados en la Fase I del proyecto, con lo cual esta octava tiene las notas por dedo (véase Tabla 1). Por otra parte, las octavas restantes, donde las siete notas de la octava dependen de una mano, tuvieron que utilizarse por medio de otros gestos.

La posición de la palma de la mano permitió ubicar la región dentro de cada octava, esta cambiaría las notas que se pudiesen tocar con la mano (véase Figura 2), es decir, por cada rango espacial, y la nota estaría dedicada al gesto de uno de los dedos, permitiendo que la persona tuviese que desplazar su mano dentro de los rangos establecidos para cada octava extra a la Fase I. Para lograr esto, se tuvo que dividir las octavas, excluyendo la cuatro en tres espacios donde cambiarían las notas de acuerdo con el gesto. Esto se realizó con el fin de aproximar la experiencia a como se toca una pieza de música en un piano en la realidad.

Evaluación: Los usuarios interactuaron con la aplicación de forma exploratoria, donde la precondition se estableció en la ejecución de la Fase I y II por parte del mismo usuario; esto permitía lograr una habilidad difusa desde la parte gestual hacia los movimientos de las manos para tocar un piano. Igualmente, se presentó una canción como referencia sonora y visual. La referencia visual se simbolizó en las notas textuales que el usuario debía mimetizar desde los gestos diseñados.

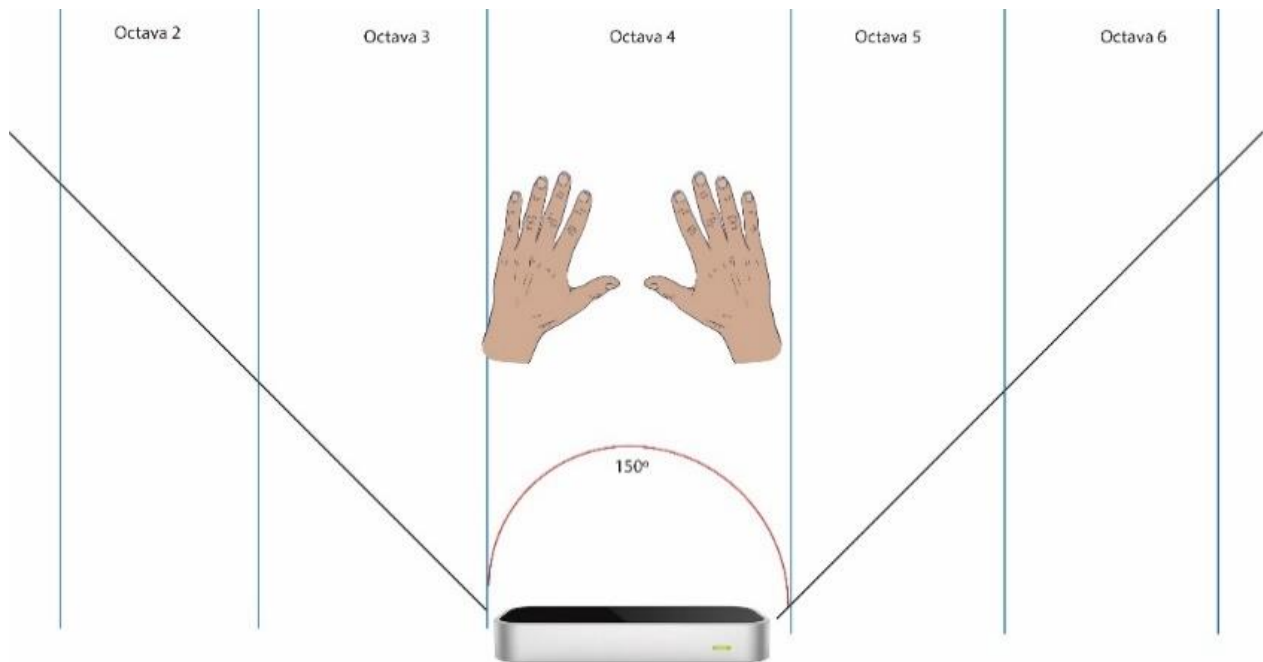


Figura 1. Regiones de movimiento definidas para las octavas dependiendo el ángulo de visión y la distancia máxima del LMC. Fuente: elaboración propia.

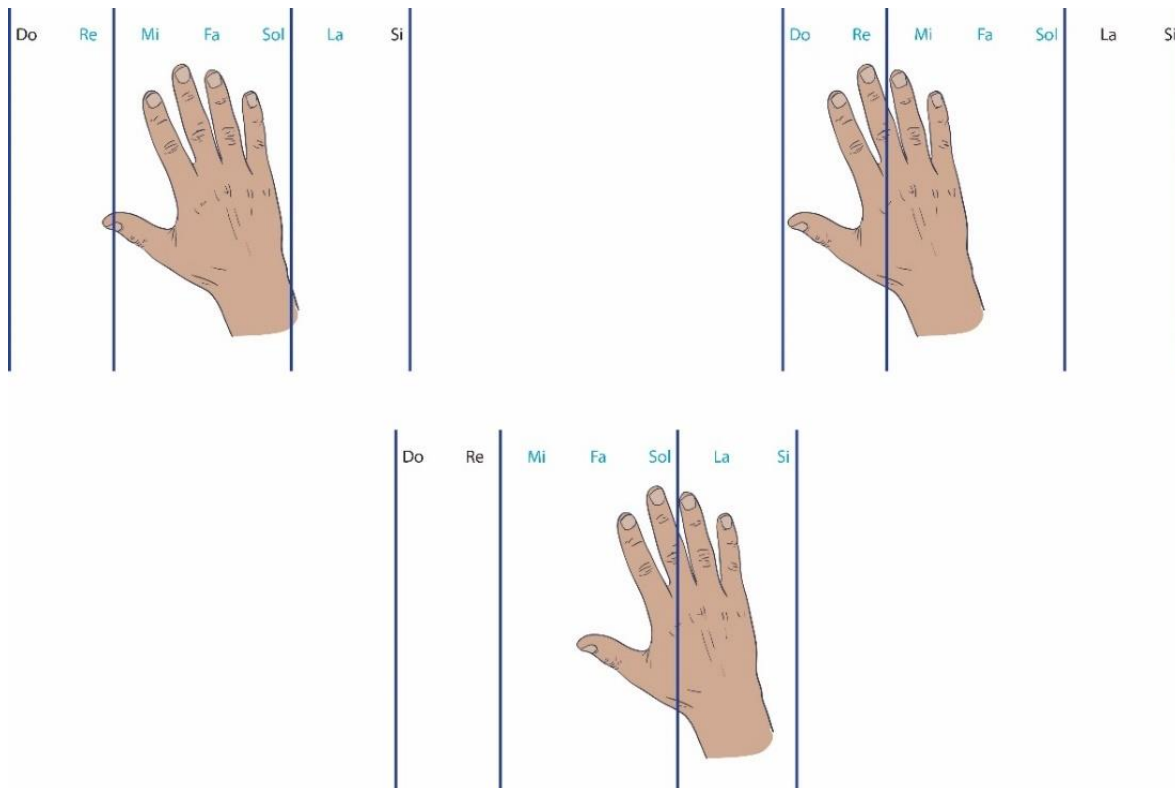


Figura 2. Posicionamiento de la palma de la mano que permite la distribución de las notas para las regiones de octavas. Fuente: elaboración propia.

6. RESULTADOS

El diseño e implementación de un entorno virtual permitió integrar objetos tridimensionales para representar un piano y manos. El propósito principal es recrear los movimientos de las manos y dedos en relación con los gestos musicales, además, cada vez que se reproducía una nota musical se representaba en el piano. Adicionalmente, se visualizaron los diferentes gestos en la interfaz para que el usuario pudiese referenciar el gesto a interpretar. La Figura 3 y la Figura 4 presentan el entorno virtual mediante la interfaz diseñada, esta fue desarrollada en el *software* Unity – Motor de Juegos, permitiendo así la integración de elementos tridimensionales y el dispositivo LMC para representar el movimiento en tiempo real.

Se recolectaron los datos con relación a los tiempos de ejecución de cada fase y el número de errores de los usuarios al interpretar un gesto que no correspondía con la nota musical.

Para la toma de los tiempos en las pruebas se tuvo en cuenta la diferencia de tiempo entre la indicación visual del primer gesto a realizar hasta la interpretación correcta del último gesto, reconocido así por el sistema, y al final de cada prueba, los datos registrados de forma automática mediante archivos de registro generados por cada fase, a su vez, etiquetado y almacenados por cada usuario.



Figura 3. La imagen representa la aplicación de identificación de gestos secuencial mediante la nota musical *do*. Fuente: elaboración propia.



Figura 4. Representación de la aplicación de identificación de gestos de Curwen mediante la nota musical *si*. Fuente: elaboración propia.

A continuación, se presentan las gráficas que muestran el tiempo promedio de todos los usuarios, teniendo como unidad de medida los segundos. Igualmente, el número de errores promedio mostrados en las gráficas por cada fase.

En la Figura 5 se presenta una mayor complejidad en la interpretación de los gestos de las notas *do*, *sol* y *la*. En tal sentido, el gesto asociado a la nota *do* es la primera interpretación que hace el usuario en la aplicación. Además, esto se representan como el número de errores al realizar el gesto como se muestra en la Figura 6. Por otra parte, las notas *sol* y *la* son interpretadas por la mano izquierda, es decir, que el usuario debe cambiar de mano para continuar con la secuencia de gestos, lo que cambia su atención de una mano a otra; sin embargo, el usuario mejora la interpretación de los gestos con la mano izquierda, puesto que en la última nota *si* se evidencia un menor tiempo y error en la realización del gesto.

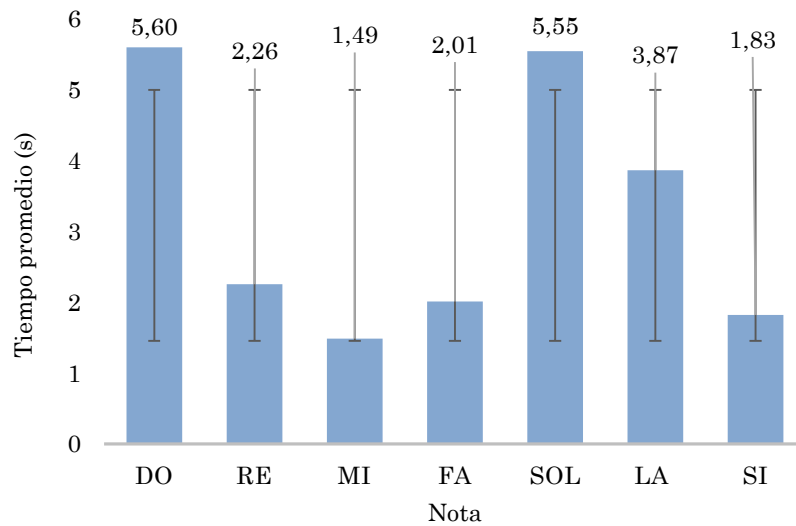


Figura 5. Tiempo promedio en piano de aire en secuencia de notas con la desviación estándar en líneas negras. Fuente: elaboración propia.

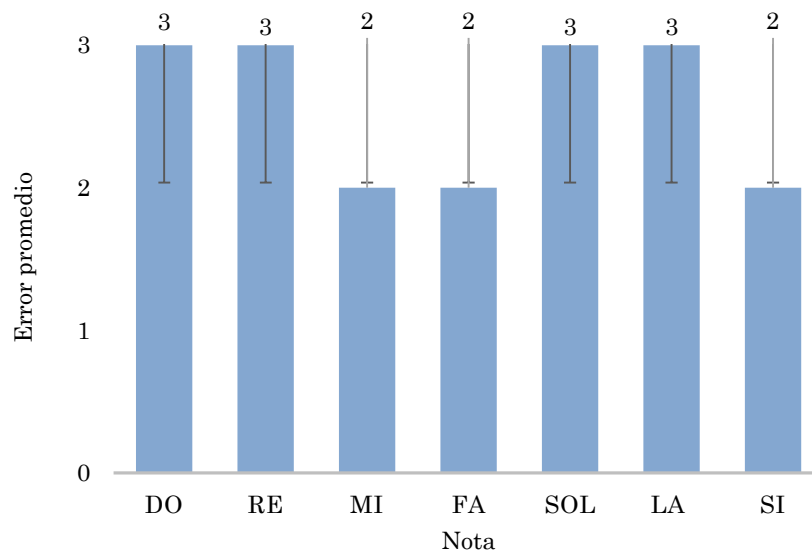


Figura 6. Errores promedio en piano de aire con secuencia de notas con la desviación estándar en líneas negras. Fuente: elaboración propia.

Luego, cada usuario interpretó la secuencia “Feliz cumpleaños”, empleando el gesto asociado a cada nota musical como se presenta en la Tabla 1. Los resultados permiten observar que el tiempo en general que el usuario realizó el gesto de forma correcta es menor a la fase anterior. Aunque, se sigue presentando dificultades en las notas Do, Sol y La (véase Figura 7). Adicionalmente, la nota Si aumentó el tiempo en el cual se realizaba el gesto. Es de anotar que las notas Sol, La y Si se realizan con los dedos de la mano izquierda puesto que se debe realizar los gestos con ambas manos. Esto puede aumentar los tiempos a causa de que deben realizarse gestos de forma no secuencial entre ambas manos. Sin embargo, el número de errores se mantuvo en promedio entre las actividades de interpretación secuencial (véase Figura 6) y no secuencial (véase Figura 10).

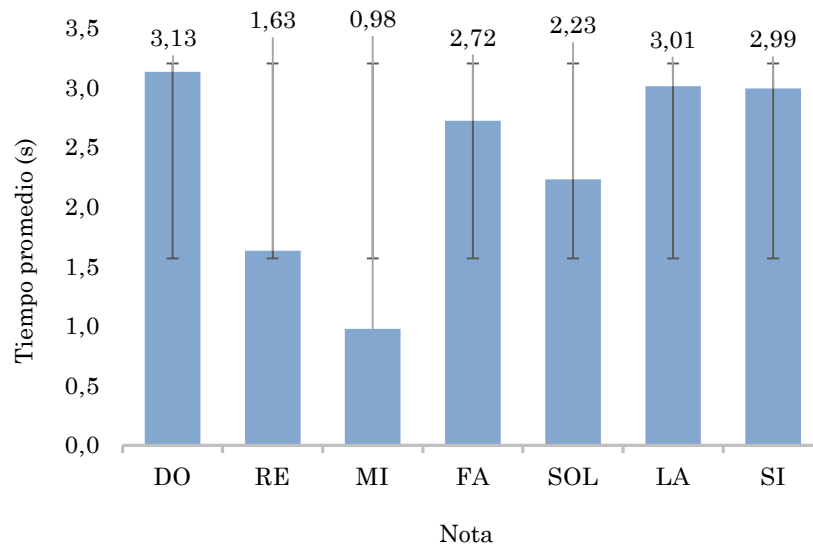


Figura 7. Tiempo promedio en piano de aire para la canción “Feliz cumpleaños” con la desviación estándar en líneas negras. Fuente: elaboración propia.

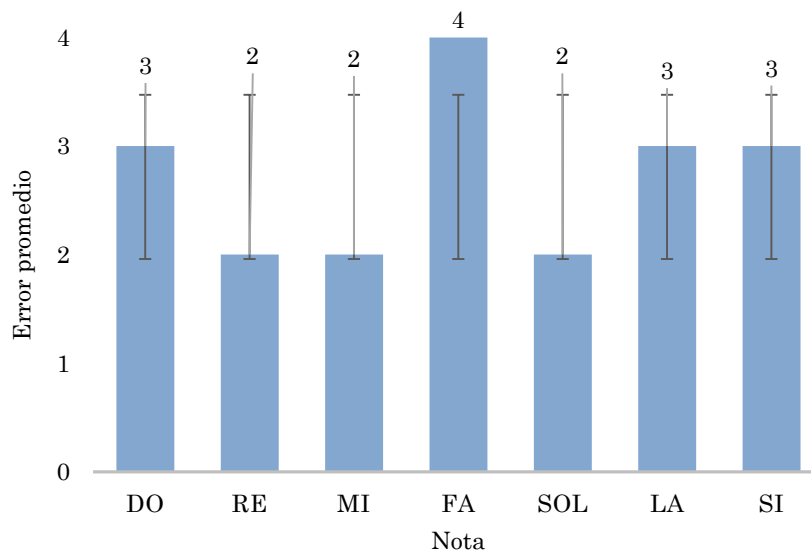


Figura 8. Errores promedio en piano de aire con secuencia de la canción “Feliz cumpleaños” con la desviación estándar en líneas negras. Fuente: elaboración propia.

Ahora bien, se presentan los datos de la fase II con la secuencia de notas interpretando los gestos de Curwen adaptados (véase Tabla 2). Obteniendo así, unos tiempos significativamente altos en las notas Fa, Sol y Si (véase Figura 9). Puesto que duplican el tiempo promedio de los otros gestos. Así mismo, se presentan un alto número de errores en la interpretación del gesto. Es de resaltar, que en la nota Fa, tiene un número de errores demasiado alto que posiblemente se debe a un reconocimiento de otros gestos similares como Do y Si que el gesto se realiza con los dedos en contracción hacia la palma (véase Figura 10).

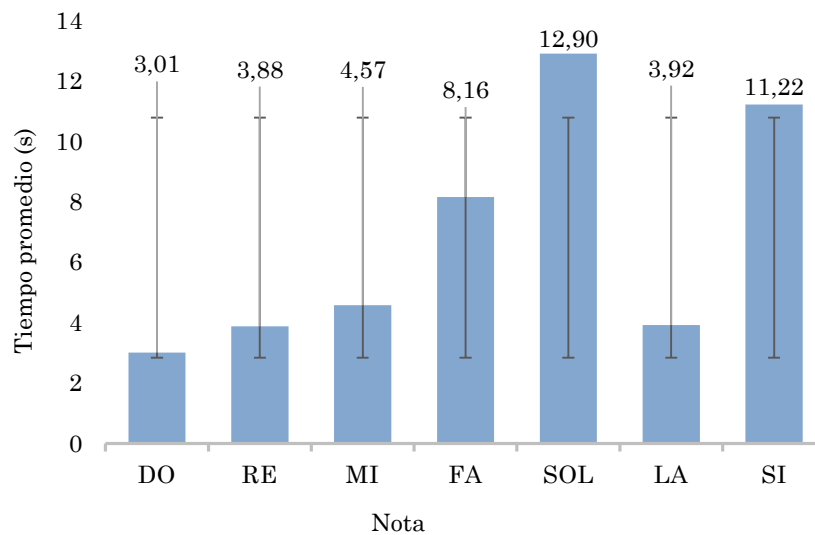


Figura 9. Tiempo promedio de gestos de Curwen con secuencia de notas con la desviación estándar en líneas negras. Fuente: elaboración propia

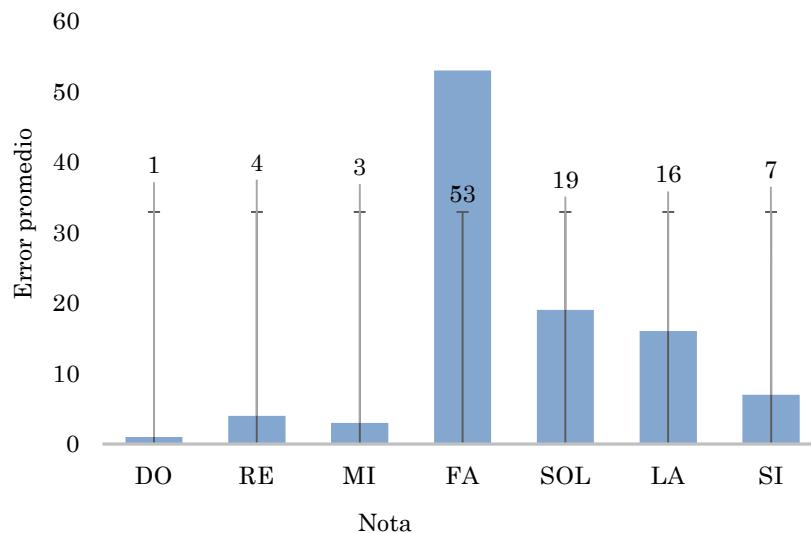


Figura 10. Errores promedio de gestos de Curwen en secuencia de notas. Fuente: elaboración propia.

Posteriormente, la secuencia de notas que representaban la canción “Feliz cumpleaños” fue interpretada por cada gesto de Curwen adaptado a cada nota musical. En tal sentido, los

resultados presentan dificultades en la realización de los gestos asociados con las notas Sol, La y Si (véase Figura 11). No obstante, el posicionamiento de la mano para interpretar el gesto de la nota Sol tienen una posición vertical, y que posiblemente el sensor no pueda reconocer el gesto debido a falta de información cuando se captura la postura de la mano y de sus dedos.

Finalmente, el número de errores de las notas Re y La tienen los valores más altos en comparación de las otras notas. Aunque, estas notas tienen una similitud en el gesto, los cual pueden confundir al usuario (véase Figura 12).

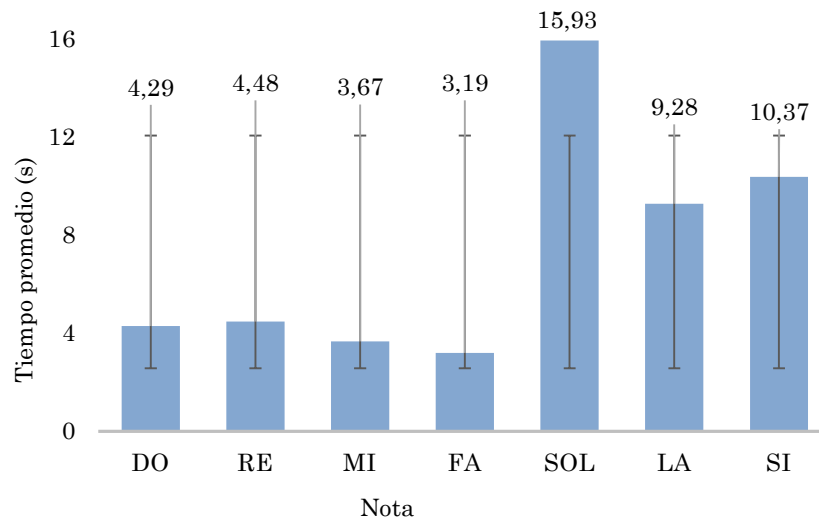


Figura 11. Tiempo promedio en la canción “Feliz cumpleaños” con gestos de Curwen con la desviación estándar en líneas negras. Fuente: elaboración propia.

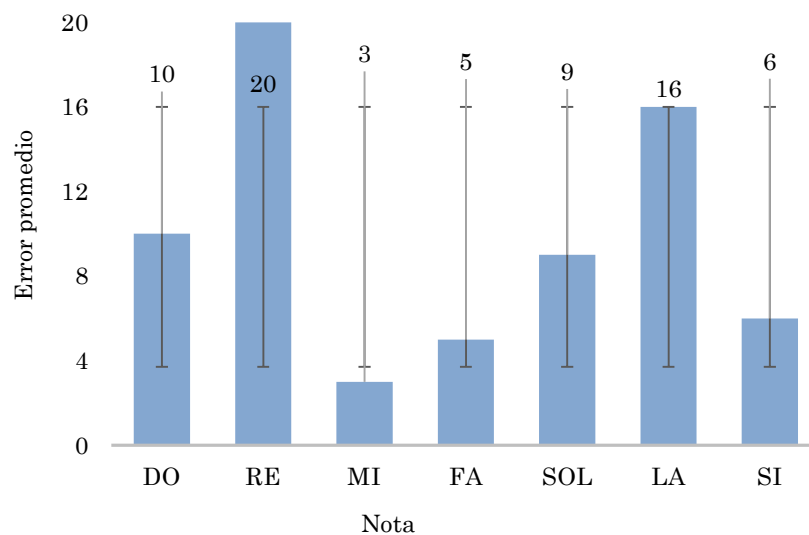


Figura 12. Errores promedio en la canción “Feliz cumpleaños” con gestos de Curwen con la desviación estándar en líneas negras. Fuente: elaboración propia.

7. DISCUSIÓN

Esta sección presenta inferencias con base en los resultados presentados en la sección anterior. Adicionalmente, estos son contrastados por las fases y por otros referentes que presentan trabajos en la misma línea.

El primer aspecto refiere a la implementación del entorno virtual el cual tenía características diferentes en la forma de interacción con el computador. En tal sentido, este tipo de interacción era novedosa para los usuarios puesto que no es un sistema comercial y no se presenta de forma cotidiana. Además, los usuarios muestran un interés inicial en el uso de la aplicación puesto que visualizar sus manos de forma virtual les llama la atención y los incentiva a realizar las actividades propuestas. Por otro lado, se les persuadió a que realizaran las pruebas. Lo anterior se puede afirmar con otros trabajos que dan lugar a una mayor precisión en la interpretación de las notas, ya que algunos usuarios mencionaron que la aplicación no era convencional y llamaba la atención a primera vista, pues tenía componentes de realidad mixta [15]. No obstante, se debe incorporar un mecanismo de retroalimentación a la aplicación desarrollada, dado que es un componente que permite evidenciar progresos en los usuarios, tal como lo indica la literatura [15].

Otro aspecto por resaltar es la representación de guías visuales de los gestos a realizar, los cuales permitieron facilitar el reconocimiento del gesto para su interpretación. De forma similar, la aplicación FunPianoAR presenta pistas para las próximas notas a tocar para facilitar el aprendizaje de la música.

Los tiempos para la Fase I (véase Figura 5 y Figura 7) presentaron una mejoría de una secuencia a otra, es decir, usuarios que realizan de forma más rápida la identificación y el gesto en la representación de la nota musical. Sin embargo, los errores que se cometían se mantuvieron en un nivel promedio (véase Figura 6 y Figura 8). Según lo mencionado anteriormente, esta inferencia se refuerza con el trabajo de un piano virtual con la misma secuencia de gestos [16]. Los autores indican que las pruebas con los usuarios tienen un buen rendimiento. No obstante, se identificaron características de mejora en alta precisión y en el campo de visión (FOV) que hacen posible controlar el piano virtual en tiempo real.

Igualmente, se puede considerar un sistema de retroalimentación de fuerza o un sistema de referenciación espacial. Esto hace que el usuario no pueda localizar sus manos virtuales rápidamente porque tiene que intentar mover sus manos reales para encontrar la ubicación de las manos virtuales, lo que hace que los usuarios no encuentren fácilmente la ubicación correcta en el entorno tridimensional.

Las pruebas realizadas para la Fase II, donde se implementaron los gestos adaptados de Curwen, muestran que los usuarios toman más tiempo para realizar las últimas tres notas musicales *sol*, *la*, *si* como se puede identificar en la Figura 9 y Figura 11. Lo anterior no significa que no haya mejoría en la realización de los gestos, pues también se puede evidenciar que el número de errores que se realizan antes de efectuar el gesto de forma correcta disminuyen en la mayoría de las notas (véase Figura 10 y Figura 12).

Finalmente, en la fase III, se recolectaron las apreciaciones de los usuarios sobre la experiencia en las diferentes fases, estas apreciaciones fueron positivas, ya que sentían que era próximo a la experiencia de tocar un piano real y se podía aprender jugando con la aplicación. Por otro lado, algunos aspectos negativos fue el poco espacio que tenían para mover sus manos al momento de interpretar las notas musicales. Es de resaltar que la implementación de la Fase III, *Aproximación a la realidad*, trata de resolver el campo de visión (FOV) limitado. De forma similar, el trabajo presentado por [14] hace una división del teclado del piano en cuatro zonas y cada zona utiliza un marcador independiente como referencia al espacio real. Asimismo, el trabajo con dos sensores de Leap Motion permiten

ampliar el FOV para el seguimiento de la mano de amplio alcance [17], para lo cual puede ser implementado en fases posteriores para mejorar el seguimiento y captura de datos.

8. CONCLUSIONES

Los gestos adaptados desde los gestos de Curwen fueron implementados en un entorno virtual como interpretación gestual en un piano de aire, obteniendo así, una aproximación musical mediante un nivel de comprensión alto con movimientos de la mano de forma natural, lo que evidencia un proceso significativo para los usuarios mediante una primera experiencia con la aplicación y el reconocimiento de los gestos.

La interpretación de los gestos de forma natural para el usuario fue característica en cada fase del proceso a través de la ejecución de una secuencia de notas para una canción establecida. En tal sentido, los resultados presentan una curva de aprendizaje incremental.

No obstante, se presentaron algunas situaciones de equivocación por parte del usuario en el orden estipulado de las notas, pero el usuario realizaba una corrección del gesto, lo que disminuía el error en la interpretación del gesto.

Por otro lado, algunos de los gestos Curwen se deben rediseñar nuevamente, esto con el fin de lograr un nivel de comprensión y gestualidad más sencilla para los usuarios, con esto se espera el aumento en la eficiencia al ejecutar los gestos.

La posición en la que el usuario realice la actividad en el aplicativo hace que se puedan capturar datos que no correspondan con los demás, esto debido a que los usuarios toman descansos en algunas ocasiones, pues la posición no tiene ningún soporte que permita descansar los brazos, y el tiempo que el usuario realiza la actividad tiene una fatiga corporal en sus extremidades superiores.

Los objetos incorporados de forma virtual como los modelos tridimensionales de las manos utilizadas en el aplicativo deben ser personalizables, es decir por parte de las apreciaciones de los usuarios, donde sugieren que en la representación de las manos no personifican una identidad propia. Estos pueden llegar a incomodar al usuario a la hora de realizar la actividad.

La usabilidad ofrecida a los usuarios se debe mejorar, pues los reconocimientos de los gestos personalizados no es un componente que esté en desarrollado en el LMC, al igual que el campo de visión (FOV), que este afectó al momento de evaluar la Fase III, ya que el espacio para cada octava era muy pequeño para ejecutar los gestos.

9. AGRADECIMIENTOS

Se agradece al programa de Ingeniería en Multimedia de la Universidad Militar Nueva Granada por los espacios de trabajos de grado que permite profundizar temáticas en relación con sistemas interactivos. Este trabajo no contó con ningún apoyo financiero.

CONFLICTOS DE INTERÉS DE LOS AUTORES

Los autores manifiestan que no existe ningún conflicto de interés, ya sea financiero, profesional o personal que pudiera surgir de la publicación del presente artículo.

CONTRIBUCIÓN DE LOS AUTORES

Christian Quintero supervisó el proyecto mediante la orientación de los métodos implementados, realizó el análisis e interpretación de los resultados y la escritura y edición final del artículo.

Diego Roa implementó el sistema para la recolección y análisis de datos y realizó la escritura del borrador inicial del artículo.

10. REFERENCIAS

- [1] A. L. Zuleta Jaramillo, “El método Kodály y su adaptación en Colombia”, *Cuad. Músic. Artes Vis. Artes Escén.*, vol. 1, no. 1, pp. 66-95, Oct. 2004. <https://revistas.javeriana.edu.co/index.php/cma/article/view/6420>
- [2] M. Lucato, “El método Kodály y la formación del profesorado de música”, *Rev. electrónica LEEME*, no. 7, pp. 1-7, May. 2001. <https://ojs.uv.es/index.php/LEEME/article/view/9725/9161>
- [3] R. I. Godøy; E. Haga; A. R. Jensenius, “Playing ‘Air Instruments’: Mimicry of Sound-Producing Gestures by Novices and Experts”, Springer, Berlin, 2006, pp. 256–267. https://doi.org/10.1007/11678816_29
- [4] M. Houlahan; P. Tacka, *Kodály Today: A Cognitive Approach to Elementary Music Education*. Oxford University Press, 2008.
- [5] D. Bell, *An introduction to cybercultures*, (1st ed.). London: Routledge. 2001. <https://doi.org/10.4324/9780203192320>
- [6] B. A. Myers, “A brief history of human-computer interaction technology”, *Interactions*, vol. 5, no. 2, pp. 44-54, Apr. 1998. <http://dx.doi.org/10.1145/274430.274436>
- [7] M. Nielsen; M. Störring; T. B. Moeslund; E. Granum, “A procedure for developing intuitive and ergonomic gesture interfaces for HCI”, en *Gesture-Based Communication in Human-Computer Interaction*, pp. 409-420, 2004. http://dx.doi.org/10.1007/978-3-540-24598-8_38
- [8] N. Henze; A. Löcken; S. Boll; T. Hesselmann; M. Pielot, “Free-hand gestures for music playback: deriving gestures with a user-centred process”, in *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia - MUM '10*, 2010, pp. 1–10. <http://dx.doi.org/10.1145/1899475.1899491>
- [9] C. Cadoz; M. M. Wanderley, “Gesture-music”. 2000. <https://hal.archives-ouvertes.fr/hal-01105543>
- [10] M. Wanderley, “Non-obvious Performer Gestures in Instrumental Music”, en: Braffort A., Gherbi R., Gibet S., Teil D., Richardson J. (eds) *Gesture-Based Communication in Human-Computer Interaction*, Springer, Berlin, 1999, pp. 37–48. http://dx.doi.org/10.1007/3-540-46616-9_3
- [11] A.-M. Burns; M. M. Wanderley, “Visual Methods for the Retrieval of Guitarist Fingering”, en *Proceedings of the 2006 conference on New Interfaces for Musical Expression*, 2006, pp. 196-199. <https://dl.acm.org/doi/10.5555/1142215.1142263>
- [12] H. H. S. Ip; K. C. K. Law; B. Kwong, “Cyber Composer: Hand Gesture-Driven Intelligent Music Composition and Generation”, in *11th International Multimedia Modelling Conference*, Melbourne, 2005. pp. 46–52. <http://dx.doi.org/10.1109/MMMC.2005.32>
- [13] H. S. Cabezas; W. J. Sarmiento, “Evaluación de modelos para el reconocimiento de gestos en señales biométricas, para un usuario con movilidad reducida”, *Tecnológicas*, vol. 22, pp. 33-47, Nov. 2019. <http://dx.doi.org/10.22430/22565337.1513>
- [14] H. Zeng; X. He; H. Pan, “FunPianoAR: A Novel AR Application for Piano Learning Considering Paired Play Based on Multi-Marker Tracking.” *J. Phys. Conf. Ser.*, vol. 1229, no. 1, pp. 1-7, 2019. <https://iopscience.iop.org/article/10.1088/1742-6596/1229/1/012072/pdf>
- [15] W. Molloy; E. Huang; B. C. Wünsche, “Mixed reality piano tutor: A gamified piano practice environment”, *ICEIC 2019 - Int. Conf. Electron. Information, Commun.*, Auckland, 2019. <http://dx.doi.org/10.23919/ELINFOCOM.2019.8706474>
- [16] W. Qiao; R. Wei; S. Zhao; D. Huo; F. Li, “A real-time virtual piano based on gesture capture data”, en *ICCSE 2017 - 12th Int. Conf. Comput. Sci. Educ.*, Houston 2017, pp. 740-743. <http://dx.doi.org/10.1109/ICCSE.2017.8085592>
- [17] I. Hwang; H. Son; J. R. Kim, “AirPiano: Enhancing music playing experience in virtual reality with mid-air haptic feedback”, en *IEEE World Haptics Conf. WHC 2017*, Munich, 2017, pp. 213-218 <http://dx.doi.org/10.1109/WHC.2017.7989903>
- [18] R. I. Godøy, “Motor-mimetic music cognition”, *Leonardo*, vol. 36, no. 4, pp. 317-319, Aug. 2003. <http://dx.doi.org/10.1162/002409403322258781>

- [19] J. Haueisen; T. R. Knösche, “Involuntary motor activity in pianists evoked by music perception”, *J. Cogn. Neurosci.*, vol. 13, no. 6, pp. 786-792, Aug. 2001. <http://dx.doi.org/10.1162/08989290152541449>
- [20] P. M. Howard, “Kodaly Strategies for Instrumental Teachers: When the Kodály method is extended to string and other instrumental instruction, it offers opportunities for enhancing student musicianship”, *Music Educ. J.*, vol. 82, no. 5, pp. 27-33, Mar. 1996. <http://dx.doi.org/10.2307/3398929>
- [21] S. C. Guillén Martínez, “El método musical Kodaly en su nivel de iniciación para favorecer el proceso lecto-escritor en los niños de 5 y 6 años”, *Instituto latinoamericano de Altos Estudios*, 2014. <https://ilae.edu.co/libros/207>
- [22] E. FrØkjaer; M. Hertzum; K. Hornbæk, “Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated?”, en *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, vol 2, no. 1. Springer, Dordrecht, 2000, pp. 345-352, <http://dx.doi.org/10.1145/332040.332455>
- [23] T. Bickmore; J. Cassell, “Social Dialogue with Embodied Conversational Agents”, en *Advances in natural multimodal dialogue systems*, Springer, 2005, pp. 23-54. http://dx.doi.org/10.1007/1-4020-3933-6_2
- [24] A. Jaimes; N. Sebe, “Multimodal human-computer interaction: A survey”, *Comput. Vis. image Underst.*, vol. 108, no. 1-2, pp. 116-134, Oct. 2007. <https://doi.org/10.1016/j.cviu.2006.10.019>
- [25] P. Guiraud, *El Lenguaje Del Cuerpo*. Fondo de Cultura Económica, 2013.
- [26] E. Hall, *Beyond Culture*. Anchor Books, 1977.
- [27] D. A. Norman, “Natural user interfaces are not natural”, *Interactions*, vol. 17, no. 3, pp. 6-10, Jun. 2010. <http://dx.doi.org/10.1145/1744161.1744163>
- [28] L. Cortés-Rico; G. Piedrahita-Solórzano, “Interacciones basadas en gestos: revisión crítica”, *Tecnológicas*, vol. 22, pp. 119-132, Dic. 2019. <http://dx.doi.org/10.22430/22565337.1512>
- [29] F. Weichert; D. Bachmann; B. Rudak; D. Fisseler, “Analysis of the accuracy and robustness of the Leap Motion Controller”, *Sensors*, vol. 13, no. 5, pp. 6380-6393, May. 2013. <http://dx.doi.org/10.3390/s130506380>
- [30] LEAP Motion, “API Overview — Leap Motion C# SDK v3.2 Beta documentation”, s.f. https://developer.archive.leapmotion.com/documentation/cpp/devguide/Leap_Overview.html
- [31] R. Hernández Sampieri; C. Fernández Collado; M. del P. Baptista Lucio, *Metodología de la investigación*, 6.a ed. McGraw-Hill, 2014.