INTERACCION
Revista Digital de AIPO
Asociación Interacción Persona-Ordenador
Vol. 2, No 2 (2021)
ISSN 2695-6578

Audio Binaural y Ganancia de Rotación en Entornos Virtuales

Binaural Audio and Rotation Gain in Virtual Environments

Arcadio Reyes-Lecuona
Dpt. Tecnología Electrónica
I.U.I. en Telecomunicación,
E.T.S.I. Telecomunicación,
Universidad de Málaga,
Campus de Teatinos, 29071,
Málaga, SPAIN
areyes@uma.es

Ana Márquez-Moncada
E.T.S.I. Telecomunicación,
Universidad de Málaga,
Campus de Teatinos, 29071,
Málaga, SPAIN

Hauke Luis Bottcher
E.T.S.I. Telecomunicación,
Universidad de Málaga,
Campus de Teatinos, 29071,
Málaga, SPAIN

Daniel González-Toledo
Dpt. Tecnología Electrónica
E.T.S.I. Telecomunicación,
Universidad de Málaga,
Campus de Teatinos, 29071,
Málaga, SPAIN

María Cuevas-Rodríguez
Dpt. Tecnología Electrónica
E.T.S.I. Telecomunicación,
Universidad de Málaga,
Campus de Teatinos, 29071,
Málaga, SPAIN

Luis Molina-Tanco
Dpt. Tecnología Electrónica
I.U.I. en Telecomunicación,
E.T.S.I. Telecomunicación,
Universidad de Málaga,
Campus de Teatinos, 29071,
Málaga, SPAIN

Recibido: 28.10.2021 | Aceptado: 16.12.2021

Palabras Clave

Realidad Virtual
Audio Binaural
Ganancia de Rotación

Resumen

Este artículo presenta un experimento exploratorio para determinar la influencia del audio 3D binaural y sus características en la detección de una manipulación en la rotación de la cabeza de un usuario dentro de un entorno virtual. Para ello, se manipula la rotación aplicando una ganancia, amplificando o atenuando los giros de la cabeza del usuario. Los resultados apuntan a que la modalidad visual proporciona más información para poder detectar dicha ganancia. Sin embargo, la manipulación de las diferencias interaurales del audio binaural puede afectar significativamente dicha detección, incluso en presencia de la modalidad visual. Asimismo, se corrobora que la exposición a entornos virtuales con la rotación de la cabeza manipulada provoca cibermareo en una mayoría de los usuarios.

Keywords

Virtual Reality
Binaural audio
Rotation gain

Abstract

This paper presents an exploratory experiment to determine the influence of 3D binaural audio and its characteristics in the detection of a manipulation in the rotation of the user's head in a Virtual Environment. To do so, we manipulate the head rotation, applying certain gain, amplifying or attenuating the user's head rotation. Our results point out that the visual modality provides more cues to detect such a gain. However, the manipulation of interaural differences in binaural audio may significantly affect such detection, even with the

visual modality present. Moreover, we corroborate that exposing the users to a virtual environment manipulating the head rotation produces cybersickness in most of the users.

1. Introducción

En un sistema de Realidad Virtual inmersiva, la rotación de la cabeza del usuario es capturada por dispositivos de seguimiento de la cabeza, normalmente integrados en los cascos de Realidad Virtual o HMD (Head Mounted Displays) y usada para la correspondiente actualización del punto de vista del mundo virtual que se muestra a dicho usuario. Normalmente, el ángulo de rotación en el mundo real de la cabeza del usuario se corresponde de forma directa con el ángulo de giro del punto de vista en el entorno virtual. Sin embargo, cuando la rotación aplicada al punto de vista no corresponde a la rotación realmente medida en la cabeza del usuario, decimos que hay una ganancia de rotación (G_r), que se define como el factor que relaciona ambas rotaciones. Cuando $G_r < 1$, el punto de vista gira más despacio que la cabeza del usuario, y más rápidamente cuando $G_r > 1$.

Utilizar una ganancia de rotación en el mundo virtual con respecto al real es útil para, por ejemplo, conseguir un ángulo visual mayor en el mundo virtual sin necesidad de girar mucho la cabeza (Sargunam, Moghadam, Suhail, & Ragan, 2017). También se usa en las técnicas de la redirección de la marcha (*Redirected Walking*), ideadas por (Razzaque, Kohn, & Whitton, 2001). Se trata de técnicas usadas para superar las limitaciones del espacio físico real donde se desarrolla la experiencia, en aquellos casos en los que el mundo virtual sea más grande. En ellas se redirige el camino del usuario de forma sutil, sin que éste lo note, de tal forma que al caminar tienda a volver al centro del espacio físico en el que se está desarrollando la experiencia. Por lo tanto, usando estas técnicas se puede navegar en espacios virtuales grandes con un espacio real limitado.

Conseguir ganancias de rotación efectivas, pero no perceptibles, es un tema activo de investigación en Realidad Virtual. La ganancia de rotación no debe perturbar la experiencia de los usuarios ni su comodidad. Ganancias muy altas pueden ser muy molestas y perceptibles, produciendo cibermareo (Oculus, 2017).

Se pueden encontrar diversos estudios sobre esto en la literatura, basados sobre todo en el umbral de detección de dicha manipulación. Esto es, la máxima ganancia para la que los usuarios no pueden diferenciar entre la rotación virtual y la física. La mayoría de estos estudios han considerado únicamente una Realidad Virtual puramente visual, sin la contribución de otras modalidades sensoriales. (Brument, Marchal, Olivier, & Argelaguet, 2020), (Steinicke, Bruder, Jerald, Frenz, & Lappe, 2009), (Congdon & Steed, 2019a).

Sin embargo, el audio también puede proporcionar información espacial relevante. De forma natural, al oír un sonido podemos saber de qué dirección viene. En particular, las diferencias interaurales son clave para la percepción de la lateralidad del sonido, es decir, para que el oyente pueda estimar el ángulo de azimut del que proviene el sonido dentro del plano horizontal (Woodworth, 1938). Cuando una fuente sonora está situada en un lateral, el sonido llega antes al oído ipsilateral y un tiempo más tarde, dependiendo de la ubicación relativa de la fuente respecto al oyente, al oído contralateral. De igual forma, la sombra acústica que produce la cabeza provoca una atenuación extra en el sonido que llega al oído contralateral respecto al ipsilateral. Estas diferencias interaurales son conocidas como ITD (Interaural Time Difference) e ILD (Interaural Level Difference). En este trabajo nos centraremos en las ITDs, que son causadas por la diferencia de distancia entre los dos oídos Estas diferencias interaurales, además de otros indicios auditivos, están incluidas dentro de lo que se conoce como la HRTF (Head Related Transfer Function). Las HRTFs modelan el filtrado direccional de la señal en la entrada del canal auditivo debido a las modificaciones provocadas por la oreja, la cabeza y el cuerpo del oyente. De este modo la generación de señales de audio espacializadas binaurales se basa generalmente en la convolución de una señal monoaural con estas HRTFs (Algazi & Duda, 2011).

En este trabajo, se plantea un estudio sobre la capacidad de los usuarios de detectar si existe o no una ganancia de rotación en un sistema de Realidad Virtual, considerando para ello la influencia que puede tener el audio espacial y, más concretamente, las diferencias interaurales de tiempo. Por este motivo, se plantea un experimento en el que se puedan comparar tres situaciones: un entorno virtual mudo o de "sólo vídeo", un entorno virtual sin estimulación visual relevante para el giro de la cabeza o de "sólo audio", y un escenario en el que se combinen ambos tipos de indicios ("vídeo+audio").

2. Trabajos relacionados

En un trabajo muy influyente, Wallach (Wallach, 1987) estudió los mecanismos que subyacen a la aparente estabilidad del

mundo visual humano a pesar de los movimientos de la cabeza, los ojos o todo el cuerpo. Cuando las personas giran la cabeza de izquierda a derecha, la imagen proyectada en las retinas se mueve en dirección opuesta al movimiento de la cabeza. Sin la cabeza girando, tal desplazamiento de imagen parecería algo en movimiento; pero cuando se correlaciona con el giro de la cabeza, no se ve ningún movimiento del entorno. Wallach estudió qué pasaría si la imagen se moviera en coordinación con el movimiento de la cabeza, pero la extensión de ese movimiento fuera menor (o mayor) de lo habitual. Su objetivo era descubrir si dicha manipulación se detectaría. Para ello ideó un dispositivo que permitía que los movimientos de la cabeza causaran desplazamientos de una imagen en cualquier porcentaje deseado. En sus experimentos, descubrió que los sujetos sanos perciben su entorno como estable cuando éste no se mueve más del 3% del giro real de la cabeza en cualquier dirección. Esto demostró que hay un proceso de compensación que corrige el desplazamiento de la imagen que normalmente acompaña al movimiento de la cabeza, dando así una apariencia de estabilidad.

Desde entonces se han realizado numerosos estudios para estimar los umbrales de detección de diferentes tipos de manipulación en la rotación de la cabeza. Jaelk et al. (Jaekl, Jenkin, & Harris, 2005) midieron la cantidad de movimiento visual que se puede añadir sin que los participantes dejen de percibir un escenario como estable. Los movimientos de la cabeza se monitorizaron por medio de un tracker mecánico de baja latencia y la información se utilizó para actualizar una pantalla visual montada en el casco. 10 participantes sentados, tenían que ajustar la ganancia de rotación de su HMD hasta que se sintiera estable, mientras movían la cabeza de distintas formas. Los umbrales de detección obtenidos para la ganancia de rotación fueron 0,84 y 1,41, es decir, se percibió como estable entre esas ganancias. Curiosamente, la ganancia con mayor probabilidad de ser percibida como estable fue 1.2, en lugar de 1, que equivaldría a no aplicar ganancia de rotación.

En otro estudio, Engel et al. (Engel, Curio, Tcheang, Mohler, & Bülthoff, 2008) propusieron a 10 participantes que se movieran dentro de una habitación con esquinas, y fueran ajustando la ganancia de rotación de forma dinámica durante los ensayos, en lugar de configurarla previamente. Las ganancias a utilizar estaban entre 0.5 y 2, y estos límites se escogieron porque siempre eran detectables. De forma consistente con los resultados de Jaelk et al., los participantes no solían notar ganancias entre 0.8 y 1.4, y todos los participantes encontraron que las ganancias menores que 1 eran menos cómodas. Las pruebas de los usuarios mostraron que las ganancias se vuelven más evidentes cuando hay fuentes de luz o sonido del mundo real, como personas que hablan en un rincón del laboratorio o que las luces no están completamente apagadas. Este estudio afirma que los usuarios emplean inconscientemente tales puntos de referencia del mundo real para orientarse en el espacio de seguimiento, y que una auralización capaz de renderizar sonidos 3D debería aumentar la capacidad de detección de la ganancia de rotación.

Por otro lado, Bruder et al. (Bruder, Interrante, Phillips, & Steinicke, 2012) estudiaron la detección de la ganancia de rotación en participantes en silla de ruedas y de pie, con la idea de investigar las diferencias entre estas dos opciones. Los participantes tenían que responder a la pregunta de si el movimiento virtual fue mayor o menor que el físico, sólo pudiendo dar 2 opciones de respuesta (sí, no), y a partir de dichas las respuestas se determinaban los umbrales de detección. El ángulo de rotación virtual se escaló con ganancias de rotación entre 0,4 y 1,6 en pasos de 0,2. El escenario virtual usaba sobre todo la modalidad visual, aunque también se utilizaban a veces señales auditivas. Los resultados fueron umbrales de 0,68 y 1,26 para estar de pie y 0,77 y 1,26 para participantes en silla de ruedas. El estudio propone que estas pequeñas diferencias se pueden deber a diferencias en la propiocepción de giro.

Los trabajos mencionados no prestan mucha atención a la modalidad auditiva, pero también hay estudios que han trabajado sobre esto. En un experimento sobre técnicas de redirección (Nilsson, Suma, Nordahl, Bolas, & Serafin, 2016) se exploró la posible influencia del sonido en un paradigma de redirección de la marcha, y si domina o no la modalidad auditiva sobre la visual, utilizando tres condiciones: sin audio (solo se mostraron imágenes), audio estático (el sonido se espacializó, pero no se vio afectado por la ganancia aplicada) y audio en móvil (el sonido fue espacializado y alineado con la fuente virtual). La condición de audio estático se incluyó con el propósito de determinar si los efectos observados fueron causados por la mera presencia de sonido. Los participantes estuvieron expuestos a ganancias de rotación que variaban de 0,5 a 1,5 en incrementos de 0,1. Se concluyó que no hay gran diferencia entre los tres casos estudiados: para no audio los umbrales de detección fueron (0.77 y 1.10), para audio estático (0.80 y 1.11) y para audio móvil (0.79 y 1.08). La mayoría de los sujetos opinaron que el audio móvil no fue muy útil a la hora de juzgar la ganancia de rotación; aun así, los participantes aseguraban que al añadir el audio la experiencia fue más inmersiva y realista.

Serafin et al. (Serafin, Nilsson, Sikstrom, De Goetzen, & Nordahl, 2013) estudiaron una situación de sólo audio mediante dos experimentos en los que se trata de redirigir a los participantes con estímulos auditivos, sin que ellos lo sepan. El primero consistía en realizar rotaciones con distintas ganancias e informar de si la rotación percibida era menor o mayor que la física, y en el segundo los usuarios caminaban por una línea curva en el mundo real, que se percibía como recta en el virtual. De estos experimentos se concluyó que los usuarios pueden ser girados físicamente alrededor de un 20% más o un 12% menos que la rotación virtual percibida, y que los usuarios pueden ser redirigidos en un arco circular mientras creen que están caminando en línea recta. Se descubrió también que es menos probable que los participantes perciban el redireccionamiento cuando se usa realimentación visual que cuando esta es puramente auditiva, lo que resulta consistente con la idea de que la visión sea superior a la audición cuando se refiere a la estimación de la localización espacial de objetos (Goldstein, 2010).

Otro estudio sobre cómo pueden ser redirigidos los usuarios (Feigl, Kõre, Mutschler, & Philippsen, 2017) afirma que las personas que caminan en línea recta redirigen su movimiento lejos de fuentes de audio fuertes. En este estudio sí que se vio diferencia entre audio estático o dinámico, y entre hombres y mujeres. Se concluyó que una fuente de audio dinámica puede cambiar a los hombres de dirección hasta en un 30% y a las mujeres hasta en un 25%, mientras que audio estático tiene aproximadamente dos tercios de este impacto. También se descubrió que la distancia de la fuente influye, a más distancia menos efectividad.

Finalmente, Congdon et al. (Congdon & Steed, 2019b) investigaron si diferentes tasas de cambio de ganancia (cambio de ganancia repentino, cambio de ganancia lento y ganancia constante) afectan al nivel de ganancia en que los usuarios perciben inconsistencias entre su movimiento virtual y físico. Es decir, si se detectan con más facilidad los cambios repentinos de ganancia o los lentos. El resultado fue que los cambios lentos se detectaban mejor. Además, los usuarios que estaban familiarizados con la realidad virtual detectaban más fácilmente estos niveles de ganancia.

En todos estos estudios previos se ha estudiado el umbral de detección de la ganancia de rotación usando vídeo sin audio, sólo audio o bien condiciones con audio+vídeo. Algunos de estos estudios comparan entre dos de estos tres casos (Nilsson et al., 2016). Pero, hasta donde sabemos, no existen estudios previos que comparen los tres casos. Tampoco se ha explorado la importancia que puede tener la correcta parametrización de las diferencias interaurales en la detección de la ganancia de rotación.

Basándonos en todos estos trabajos previos, nuestras hipótesis son que la estimulación visual es más influyente que la auditiva para detectar la ganancia de rotación, pero que la máxima sensibilidad se producirá en una situación bimodal en la que se proporcione vídeo y audio 3D. Por otro lado, postulamos que la manipulación de las diferencias interaurales pueden también influir en la capacidad del usuario de detectar una ganancia de rotación.Sin embargo, dado que este tipo de experiencias puede producir molestias en los usuarios, nos planteamos primero la realización de un experimento piloto reducido, en el que los usuarios interactúen sólo durante unos pocos minutos y del que podamos extraer experiencia para afinar mejor un diseño experimental que aborde las cuestiones mencionadas.

3. Material y Método

3.1 Participantes

En el estudio han participado 12 sujetos (7 mujeres y 5 hombres) con edades comprendidas entre los 20 y los 29 años reclutados entre los estudiantes de la Universidad de Málaga. Los participantes no recibieron ningún tipo de remuneración por su participación. Todos ellos han reportado no tener problemas auditivos ni de visión. El estudio ha sido revisado y aprobado por el Comité de Ética de la Universidad de Málaga.

3.2 Escenario Virtual y Aparataje

Durante el estudio, los participantes se encuentran inmersos en un entorno de Realidad Virtual implementado haciendo uso de la plataforma de desarrollo de videojuegos Unity (Unity, 2021). Los participantes visualizan el entorno a través de un casco de Realidad Virtual Oculus Rift (Oculus, n.d.) e interactúan con él haciendo uso de sus dos controladores. Además, los participantes llevan puestos unos auriculares SONY MDR-7506 para recibir el audio de la escena.

Este entorno de Realidad Virtual inmersivo está formado por dos componentes: componente visual y componente auditiva. La componente visual muestra un entorno 3D desarrollado con Blender (Blender Foundation, 2015) e importado a Unity, el cual simula una habitación donde el participante se encuentra situado en el centro, tal y como se muestra en la Figura 16, sobre la cruz roja y mirando inicialmente en dirección al escritorio. La componente auditiva consiste en un sonido binaural espacializado gracias

al uso de la librería 3DTI-Toolkit (Cuevas-Rodríguez et al., 2019), también integrada en Unity. El estímulo de audio utilizado es una voz masculina anecoica hablando en inglés extraída de *Music from Archimedes, English male speech, B&K 4003, anechoic* (Bang & Olufsen, n.d.). La HRTF utilizada para la espacialización del audio es la ID_1032 de la base de datos LISTEN (Warusfel, 2003) de 256 muestras con una frecuencia de muestreo de 44.1 kHz.

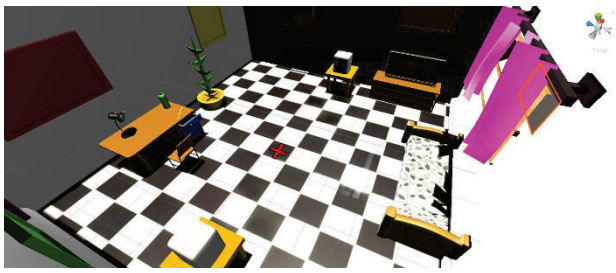

Figura 16: Escenario virtual.

3.3 Protocolo higiénico

Debido a la crisis sanitaria provocada por la COVID-19, se han adoptado una serie de medidas higiénicas que se describen a continuación. El experimento se ha desarrollado en un laboratorio de unos 100m^2, con puertas y ventanas abiertas en todo momento para asegurar la ventilación cruzada, en el que solo están presentes el participante y el experimentador. Se ha monitorizado el nivel de CO_2 para que no supere los 600ppm. Se han seguido las recomendaciones al efecto para experimentos de Realidad Virtual recogidas en (Posselt & Gosselin, 2020): El casco de realidad virtual tiene una funda de silicona para las partes que están en contacto con la piel, que se desinfecta con cada participante, y se usa una sola vez por día. En total se usan 3 cascos de realidad virtual, por lo tanto, cada día hay como máximo 3 participantes. Los auriculares se cubren con una funda desechable. Se proporcionó una mascarilla quirúrgica nueva a los participantes salvo que ya trajeran una de mayor seguridad. Antes de empezar con cada participante, se cambian los dispositivos y se desinfectan, La experimentadora usa guantes de vinilo y se desinfecta las manos con gel hidroalcohólico al 70 % antes de recibir a cada participante.

3.4 Diseño del experimento

Como se desea estudiar cada componente, visual y auditivo, por separado y la influencia de las dos componentes a la vez, se han diseñado tres condiciones:

- Solo Vídeo. Escenario donde el sujeto se encuentra dentro de una habitación virtual y donde aparece dos altavoces sobre sendas mesas situadas a la izquierda o derecha del sujeto (ver Figura 16). En este escenario no aparece ningún estímulo auditivo.
- Solo Audio. Escenario donde el sujeto visualiza únicamente un entorno virtual con un suelo y un cielo, sin ningún detalle visual que permita conocer el giro del punto de vista, y escucha un sonido espacializado a su izquierda o su derecha (ver Figura 17).

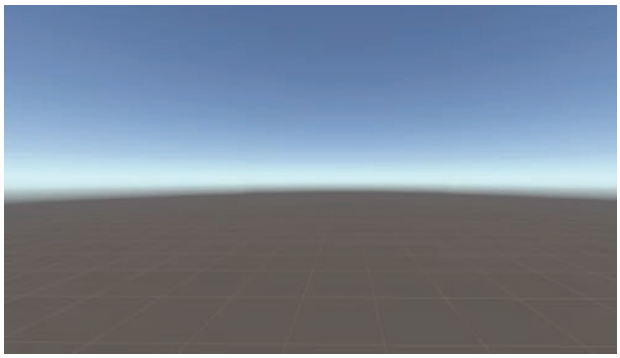

Figura 17: Estímulos visuales en el escenario de sólo audio.

- Audio + Vídeo. Este escenario incluye los dos escenarios mencionados anteriormente, el sujeto se encuentra dentro de la habitación virtual donde se encuentran los altavoces a izquierda y derecha. Uno de ellos emite el sonido espacializado.

En los tres escenarios el sujeto lleva puesto en todo momento el HMD y los auriculares. Cuando nos referimos a la derecha e izquierda del sujeto queremos decir en una posición en coordenadas polares de $\pm 90°$ de azimut y $0°$ de elevación, considerando que el origen $(0°, 0°)$ está justo delante del sujeto.

El estudio presentado explora también la influencia del ITD en el umbral de detección de la ganancia de rotación. Para ello se ha personalizado el ITD del HRTF utilizado para la espacialización de audio. La librería 3DTI Toolkit ofrece esta posibilidad, calculando el ITD individual del sujeto a partir del radio de su cabeza (Cuevas-Rodríguez et al., 2019) mediante la fórmula de woodworth (Woodworth, 1938).

$$ITD = \frac{a}{c}(\theta + sen\,\theta)$$

Donde θ es el azimut interaural de la fuente sonora respecto al oyente, c la velocidad del sonido y a el radio de la cabeza. Este radio se obtiene a partir del perímetro de la cabeza de cada sujeto, que se midió utilizando una cinta métrica, colocada rodeando la cabeza, aproximadamente 2cm por encima de las cejas, pasando sobre la parte superior de las

orejas (Figura 18). Para el estudio se consideran tres posibles condiciones:

- ITD individual. Uso del ITD individual del sujeto, calculado a partir del radio de la cabeza del sujeto.
- ITD reducido. Uso de un ITD menor que el individual, calculado con un radio de cabeza un 20% menor.
- ITD aumentado. Uso de un ITD mayor que el individual, calculado con un radio de cabeza un 20% mayor.

Figura 18: Medida del perímetro de la cabeza de los participantes.

Para estimar la capacidad de detección de ganancia de rotación, cada una de las condiciones experimentales anteriores (tipo de escenario e ITD) se ejecutan con los siguientes valores de ganancia de rotación: G_r = 0.6, 0.8, 1, 1.2, 1.4 y 1.6. Por lo tanto, tenemos un total de 3 niveles audio/vídeo, 3 niveles de ITD y 6 niveles de ganancia de rotación, lo que hace un total de 54 condiciones experimentales diferentes.

3.5 Procedimiento

Tras ser informados del procedimiento del estudio y firmar un consentimiento de participación, los participantes completan un cuestionario demográfico. Seguidamente, se les mide el perímetro de la cabeza para la individualización del ITD. Para poder comenzar el experimento, el sujeto debe ponerse el HMD, los auriculares, sostener los controladores y colocarse de pie en el centro de la sala.

El experimento está formado por tres bloques, uno para cada condición sensorial: solo Vídeo, solo Audio, Audio + Vídeo. Cada bloque está formado por 18 condiciones (3 configuraciones de ITD x 6 valores de G_r). Tanto los bloques como los ensayos dentro de cada bloque están aleatorizados siguiendo un contrabalanceado utilizando cuadrados latinos.

Los participantes deben realizar un descanso entre bloques, de la duración que deseen, quitándose los auriculares y el casco de Realidad Virtual y sentándose si lo considera necesario. Al inicio de cada bloque aparecen dos ensayos de entrenamiento (con ITD individual y G_r = 1) donde el usuario podrá familiarizarse con el entorno y realizar las preguntas que desee al experimentador.

Para comenzar cada ensayo el sujeto debe colocarse mirando hacia la posición inicial, donde aparece un botón que debe pulsar. En este momento aparece una flecha que indica hacia qué lado debe dirigirse para posicionarse frente al altavoz (escenario Solo Vídeo), o frente a la fuente de audio (escenario Solo Audio) o ambos (escenario Audio + Video). Una vez el sujeto se encuentra delante de la imagen o fuente sonora, debe pulsar un botón del controlador para confirmar que ha completado el giro, apareciendo entonces la siguiente pregunta: *¿ha detectado manipulación del giro?*, y dos botones con las opciones *Sí* y *No* (Figura 19). El sujeto debe elegir una de las dos opciones y volver a la posición inicial para comenzar con el siguiente ensayo dentro del bloque. Un mensaje en la pantalla indica la finalización de cada bloque.

En total, la duración del experimento por cada participante, incluyendo los descansos fue de aproximadamente media hora.

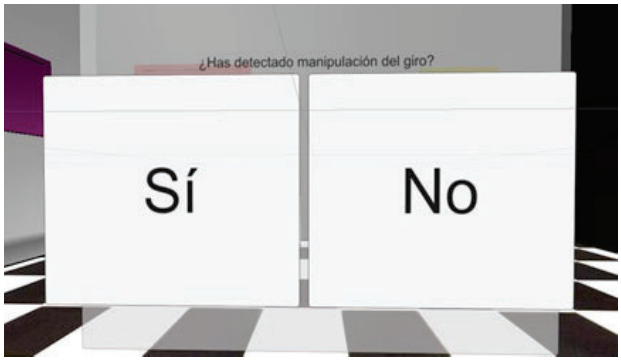

Figura 19: Interfaz para que el participante informe de si ha detectado manipulación en la ganancia de rotación o no después de cada ensayo.

4. Resultados

La Figura 5 muestra las respuestas de los participantes a la pregunta de si detectaban manipulación o no. Más concretamente, la figura representa la proporción de respuestas en las que el participante no detectó ninguna manipulación en el giro, o lo que es lo mismo, se percibió una ganancia G_r=1.

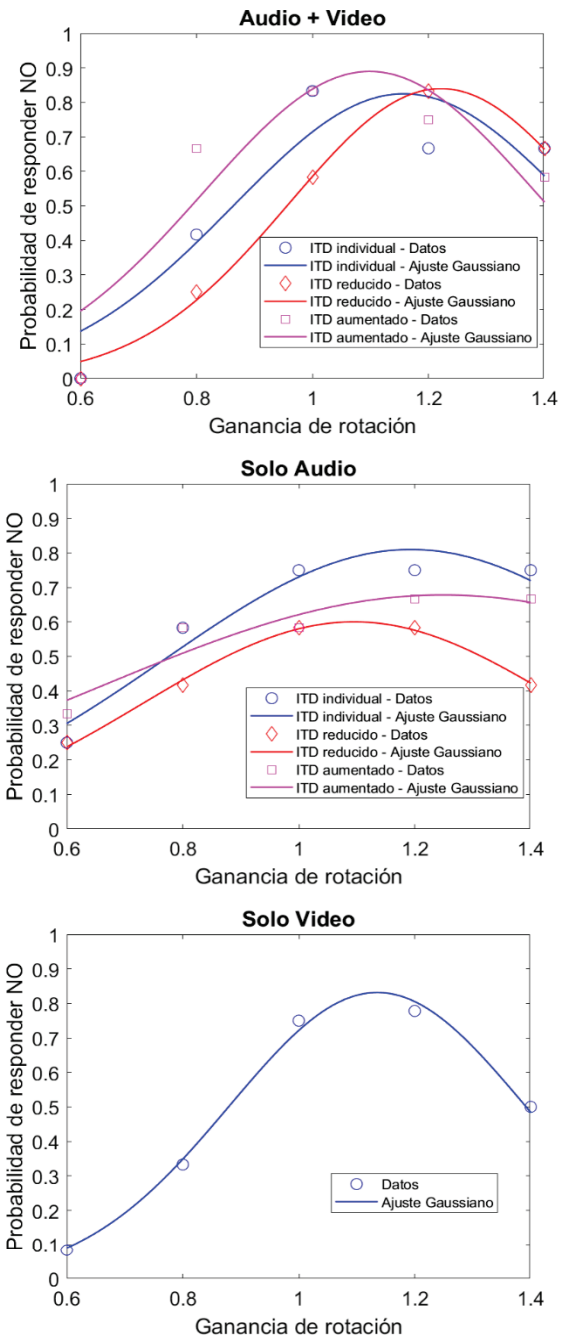

Figura 20: Probabilidades de que los participantes contestaran 'no' en el experimento, lo que significa que no detectaban ninguna manipulación de la rotación. Se muestran las curvas correspondientes a las condiciones Audio+Vídeo (a), Solo Audio (b) y Solo Vídeo(c).

La Figura 5a muestra dichas tasas para la condición Audio+Vídeo, la Figura 5b hace lo propio para la condición Solo Audio y la Figura 5c para la condición Solo Vídeo. En los dos primeros casos, en los que hay estimulación auditiva, se diferencian las situaciones en las que las diferencias de tiempo interaurales son las que corresponden a cada participante (ITD individual) de aquellas en las que se manipuló incrementándolo el radio de la cabeza en un 20% (ITD aumentado) o reduciéndolo en un 20% (ITD reducido).

Además de las tasas, señaladas en la Figura 5 con círculos, rombos y cuadrados, se ha realizado un ajuste psicométrico de estas curvas, asumiendo que la probabilidad de detectar la manipulación de la rotación sigue una ley gaussiana con respecto a la ganancia de rotación. Ese ajuste se muestra con líneas continuas en la citada figura. Tanto los datos obtenidos como el ajuste de ganancia se han llevado a cabo para los valores de G_r simétricamente alrededor de 1, entre 0.6 y 1.4.

Lo primero que llama la atención es que las curvas de ajuste gaussiano no están centradas en $G_r = 1$. Las ganancias de rotación ligeramente superiores a 1 son más difícilmente detectables en las tres condiciones experimentales. Esto es consistente con los resultados obtenidos en trabajos anteriores, como se ha comentado anteriormente. Otro resultado claro es que las condiciones con estimulación visual permiten detectar con mayor sensibilidad la manipulación de la rotación, como puede deducirse de la mayor apertura de la campana para la condición de Sólo Audio.

Puesto que los datos de que disponemos son respuestas sí/no a la pregunta de si se detecta manipulación en la rotación, el número de respuestas afirmativas y negativas para cada una de las diferentes condiciones se sometió a una prueba χ^2, que no detectó ninguna diferencia significativa. No obstante, esto es explicable dado el reducido número de medidas (una respuesta para cada condición por cada participante). A la vista de los resultados que se muestran en la Figura 5, es de esperar que un experimento de mayores dimensiones, sí que muestre diferencias significativas entre las diferentes condiciones, así como una influencia del ITD.

Al finalizar cada bloque, los participantes puntuaban en una escala de 1 a 10 si habían sentido mareo. Siete de ellos declararon haberse sentido mareados en un grado superior a 5. Esto sucedía especialmente con las condiciones con video, como muestra el valor medio de las puntuaciones: 3.75 (Audio+Vídeo), 2.5 (Solo Audio) y 3.92 (Solo Vídeo). Un análisis ANOVA arrojó que estas diferencias son marginalmente significativas ($F_{2,22} = 3.372$; $p=0.053$).

5. Discusión y conclusiones

Esta contribución presenta un estudio experimental piloto que explora la posible influencia del audio tridimensional y sus parámetros en la detección de una manipulación en la rotación de la cabeza en un entorno virtual por medio de la

aplicación de una ganancia de rotación. Lo primero que debemos señalar es que, como sospechábamos que la manipulación de la rotación de la cabeza iba a provocar cibermareo (Oculus, 2017), decidimos desarrollar un experimento muy breve, con sólo una repetición por participante de cada una de las 54 condiciones experimentales, distribuidas en tres bloques de 18 con descansos entre ellos. Aun así, más de la mitad reportaron un nivel de mareo superior a 5 sobre 10, lo que corrobora la necesidad de tomar esta precaución. Pero, al mismo tiempo, al disponer de tan pocos datos, es imposible obtener resultados estadísticamente significativos. No obstante, el análisis cualitativo de los resultados revela una serie de conclusiones interesantes.

- La ganancia de rotación es más fácilmente detectable con estimulación visual que auditiva. En el caso de una estimulación multimodal auditiva y visual, la sensibilidad mostrada es similar a la que se encuentra con solo visual.
- A pesar de lo anterior, cuando se manipulan características del renderizado de audio, como las diferencias interaurales de tiempo, el efecto del audio 3D se percibe claramente, lo que demuestra que éste sí que tiene una influencia en la detección, incluso cuando la modalidad visual está presente.
- El cibermareo parece ser mayor para la estimulación visual que para la auditiva, aunque los datos son muy escasos como para obtener resultados significativos. Además, los descansos entre bloques no fueron tan largos como para evitar que el cibermareo provocado en uno de dichos bloques influyera en el siguiente.

Por todo lo anterior, podemos considerar que este experimento piloto corrobora el interés de este estudio, pero requiere de un importante rediseño que permita obtener más datos para conseguir significación estadística, al tiempo que reduzca la exposición a los estímulos que provocan cibermareo en los participantes. Para ello, proponemos centrar el futuro experimento en la combinación de modalidades visuales y auditivas, con objeto de reducir así todo lo posible la extensión de las pruebas y evitar el cibermareo. En la misma línea, sería conveniente utilizar ganancias de rotación más cercanas a 1, ya que es probable que las situaciones con ganancias muy extremas sean las principales responsables de ese cibermareo. En este sentido, sería interesante estudiar si cuando la estimulación auditiva contribuye a esconder la ganancia de rotación se produce menos cibermareo en la situación contraria.

Por otro lado, sería interesante comprobar si una manipulación del renderizado acústico que incluya no sólo el ITD, sino toda la HRTF podría provocar un efecto aún más potente. Por ello, nuestra propuesta es desarrollar el experimento manipulando la elección de las respuestas al impulso de la cabeza de forma que la ganancia de rotación visual sea diferente a la auditiva. Esto permitiría averiguar de qué manera se resuelve el conflicto generado por diferentes ganancias en las dos modalidades, lo que haría posible construir un modelo que prediga los umbrales de detección de ganancia de rotación a partir de las dos ganancias aplicadas en las dos modalidades.

6. Reconocimientos

Este trabajo ha sido parcialmente financiado con fondos del proyecto PID2019-107854GB-I00, del Ministerio de Ciencia e Innovación, correspondiente a la convocatoria 2019 de Proyectos I+D+i 2019.

Referencias

Algazi, V., & Duda, R. (2011). Headphone-based spatial sound. IEEE Signal Processing Magazine, 28(1), 33–42. https://doi.org/10.1109/MSP.2010.938756

Bang, & Olufsen. (n.d.). Music for Archimedes CD Track Listing at cyList. Retrieved April 28, 2021, from https://www.cylist.com/List/405112822/bang-olufsen-music-for-archimedes-cd-track-listing

Blender Foundation. (2015). blender.org - Home of the Blender project - Free and Open 3D Creation Software. Retrieved May 12, 2020, from Blender.Org website: https://www.blender.org/

Bruder, G., Interrante, V., Phillips, L., & Steinicke, F. (2012). Redirecting walking and driving for natural navigation in immersive virtual environments. IEEE Transactions on Visualization and Computer Graphics, 18(4), 538–545. https://doi.org/10.1109/TVCG.2012.55

Brument, H., Marchal, M., Olivier, A.-H., & Argelaguet, F. (2020). Influence of Dynamic Field of View Restrictions on Rotation Gain Perception in Virtual Environments. https://doi.org/10.1007/978-3-030-62655-6_2

Congdon, B. J., & Steed, A. (2019a). Sensitivity to rate of change in gains applied by redirected walking. 25th ACM Symposium on Virtual Reality Software and Technology, 1–9.

Congdon, B. J., & Steed, A. (2019b). Sensitivity to rate of change in gains applied by redirected walking. Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST. https://doi.org/10.1145/3359996.3364277

Cuevas-Rodríguez, M., Picinali, L., González-Toledo, D., Garre, C., de la Rubia-Cuestas, E., Molina-Tanco, L., & Reyes-Lecuona, A. (2019). 3D Tune-In Toolkit: An open-source library for real-time binaural spatialisation. PLOS ONE, 14(3), e0211899. https://doi.org/10.1371/journal.pone.0211899

Engel, D., Curio, C., Tcheang, L., Mohler, B., & Bülthoff, H. H. (2008). A psychophysically calibrated controller for navigating through large environments in a limited free-walking space. Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology - VRST '08. https://doi.org/10.1145/1450579

Feigl, T., Kőre, E., Mutschler, C., & Philippsen, M. (2017). Acoustical manipulation for redirected walking. Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST, Part F131944. https://doi.org/10.1145/3139131.3141205

Goldstein, E. (2010). Sensation and Perception. Wadsworth Publishing Company.

Jaekl, P. M., Jenkin, M. R., & Harris, L. R. (2005). Perceiving a stable world during active rotational and translational head movements. Experimental Brain Research, 163(3), 388–399. https://doi.org/10.1007/s00221-004-2191-8

Nilsson, N. C., Suma, E., Nordahl, R., Bolas, M., & Serafin, S. (2016). Estimation of detection thresholds for audiovisual rotation gains. Proceedings - IEEE Virtual Reality, 2016-July, 241–242. https://doi.org/10.1109/VR.2016.7504743

Oculus. (n.d.). Oculus Rift: visor de VR para PC optimizadas para VR | Oculus. Retrieved April 29, 2021, from https://www.oculus.com/rift/?locale=es_LA

Oculus. (2017). Oculus Best Practices. Retrieved from https://developer.oculus.com/documentation/unity/unity-best-practices-intro/

Posselt, J., & Gosselin, R. (2020). Protocols of use for immersive platformsunder covid19. InEuroVr 2020 Application, Exhibition & Demo Track: Pro-Ceedings of the Virtual EuroVR Conference, 19–21.

Razzaque, S., Kohn, Z., & Whitton, M. C. (2001). Redirected Walking. Eurographics 2001 - Short Presentations. https://doi.org/10.2312/egs.20011036

Sargunam, S. P., Moghadam, K. R., Suhail, M., & Ragan, E. D. (2017). Guided head rotation and amplified head rotation: Evaluating semi-natural travel and viewing techniques in virtual reality. Proceedings - IEEE Virtual Reality, 19–28. https://doi.org/10.1109/VR.2017.7892227

Serafin, S., Nilsson, N. C., Sikstrom, E., De Goetzen, A., & Nordahl, R. (2013). Estimation of detection thresholds for acoustic based redirected walking techniques. Proceedings - IEEE Virtual Reality, 161–162. https://doi.org/10.1109/VR.2013.6549412

Steinicke, F., Bruder, G., Jerald, J., Frenz, H., & Lappe, M. (2009). Estimation of detection thresholds for redirected walking techniques. IEEE Transactions on Visualization and Computer Graphics, 16(1), 17–27.

Unity. (2021). Plataforma de desarrollo en tiempo real de Unity | Motor de VR y AR en 3D y 2D. Retrieved October 28, 2021, from https://unity.com/es

Wallach, H. (1987). Perceiving a Stable Environment When One Moves. Http://Dx.Doi.Org/10.1146/Annurev.Ps.38.020187.000245, 38(1), 1–29. https://doi.org/10.1146/ANNUREV.PS.38.020187.000245

Warusfel, O. (2003). LISTEN HRTF DATABASE. Retrieved January 14, 2020, from http://recherche.ircam.fr/equipes/salles/listen/

Woodworth, R. S. (1938). Experimental Psychology. Holt, New York.