TECNIACUSTICA VALENCIA · 2015

46º CONGRESO ESPAÑOL DE ACÚSTICA
ENCUENTRO IBÉRICO DE ACÚSTICA
EUROPEAN SYMPOSIUM ON VIRTUAL ACOUSTICS
AND AMBISONICS

MEJORAS Y PREFERENCIAS DE LOS USUARIOS EN LA AURALIZACIÓN PARA SISTEMAS DE MULTICONFERENCIA UTILIZANDO SONIDO BINAURAL

REFERENCIA PACS: 43.60.Dh

Aguilera, Emanuel; Lopez, Jose Javier; Gutierrez-Parera, Pablo

Universidad Politécnica de Valencia
ITEAM, Instituto de Telecomunicaciones y Aplicaciones Multimedia
Camino de Vera s/n, Edificio 8G, acceso D
46022 Valencia (España)
Tel: +34 963 877 007 (Ext. 73008) Fax: +34 963 879 583
emagmar@iteam.upv.es, jjlopez@dcom.upv.es, pabgupa@iteam.upv.es

ABSTRACT

The introduction of spatial audio in multi-party teleconferencing systems create more realistic and immersive communication environments compared to monaural systems. Moreover, the introduction of auralization can increase the immersion but at the expenses of a reduced intelligibility. In this paper, by means of subjective testing, we analyze the influence on user preferences of some specific auralization processing details (in relation to early echoes and late reverberation) and a simple near-field HRTF processing. This way, we study a trade-off between realism and intelligibility in a multi-party teleconferencing system with binaural audio developed for mobile terminals.

RESUMEN

La introducción de audio espacial en sistemas de multiconferencia produce entornos de comunicación más realistas e inmersivos que los sistemas monoaurales. Además, añadir auralización puede incrementar la inmersión a costa de reducir la inteligibilidad. En este trabajo, por medio de varios tests subjetivos, se analiza la influencia de algunos detalles más específicos de la auralización (en relación a las primeras reflexiones y reverberación tardía) y de un cálculo simple de HRTF en campo cercano en la preferencia de los usuarios. Se estudia así una solución de compromiso entre realismo e inteligibilidad en un sistema de multiconferencia con audio binaural desarrollado para dispositivos móviles.

1. INTRODUCCIÓN

Los servicios de telecomunicación que permiten celebrar reuniones han mejorado en los últimos años con sistemas como las multiconferencias. Estas tecnologías se han adoptado rápidamente en el ámbito de los negocios porque permiten ahorrar mucho tiempo y gastos en viajes [1,2], pero también son útiles para el público general. Por esta razón desde hace mucho tiempo las empresas han estado usando estos servicios a través de la red de telefonía fija.

Del mismo modo, un terminal móvil también puede usar servicios de multiconferencia [3]. La primera y más obvia ventaja es que los usuarios empresariales ganan movilidad fuera de sus oficinas. Por otra parte, con un sistema de acceso y tarifas adecuado, este servicio puede ser también atractivo para usuarios particulares (no empresariales) que podrían realizar reuniones de audio con amigos y familiares desde cualquier lugar. Además, hoy en día los terminales móviles tienen muchas más prestaciones que los terminales de telefonía fija, que han sido relegados prácticamente solo a comunicaciones de voz. Aprovechando las capacidades de los nuevos terminales móviles inteligentes es posible realizar multiconferencias más realistas, ubicuas e inmersivas. Una de las mejoras más atractivas está relacionada con la incorporación de audio espacial a la conferencia [3,4].

En anteriores trabajos se ha demostrado que el audio espacial (estéreo o binaural) es preferido por los usuarios más que el audio monoaural no espacial [4,5], mejorando la discriminación del hablante [5,6] y la inteligibilidad de las conversaciones ya que se aprovecha la habilidad del sistema auditivo humano de prestar una atención selectiva a los sonidos de una determinada dirección ("*cocktail party effect*"). Por otra parte, la incorporación del efecto de sala (auralización) en la audioconferencia produce resultados controvertidos según estudios anteriores [7], ya que, aunque incrementa la percepción de distancia, degrada la dirección de la fuente de audio y su inteligibilidad.

Los autores de este artículo presentaron recientemente un trabajo [8] investigando individualmente la influencia de las primeras reflexiones, la reverberación tardía y otros detalles que normalmente se emplean en los métodos de auralización clásicos. Los resultados mostraron que para este tipo de aplicación los usuarios prefieren la inteligibilidad sobre la sensación realista de distancia, de modo que en los test de usuario la señal limpia o solo con una simple auralización obtuvo mejores resultados que una compleja y completa auralización. En concreto, la reverberación tardía obtuvo mejor sensación de distancia pero peor inmersión que usando las primeras reflexiones.

En este trabajo se propone una nueva versión del sistema de multiconferencia inmersiva para dispositivos móviles que mejora el anterior introduciendo mejores algoritmos y un ajuste fino de los efectos de auralización al mismo tiempo que se presentan todos los test subjetivos realizados para llegar a estas mejoras. Una nueva característica de esta versión del software es que un usuario puede comenzar una conversación privada con cualquiera de los participantes de la multiconferencia acercando mucho el avatar del otro. Esta función se llama "modo susurro" porque ambos en la conversación privada se escuchan como si hablaran muy cerca mientras siguen escuchando al resto del grupo en el fondo.

De nuestro anterior trabajo se deduce que en una teleconferencia los usuarios prefieren una señal limpia a una señal con una fuerte auralización, pero ahora se pretende averiguar si una auralización sutil puede mejorar este tipo de aplicaciones. Usando este sistema se han analizado algunos ajustes finos de la auralización y del procesado HRTF de la audioconferencia por medio de test subjetivos a voluntarios. La idea es investigar la influencia de tres aspectos diferentes por medio de los tests:

- Diferentes valores de reflexión de las paredes virtuales para el cálculo de las primeras reflexiones.
- Diferentes valores para el cálculo la reverberación tardía.
- La introducción de un procesado HRTF de campo cercano para las fuentes virtuales que están muy cerca de la cabeza para el "modo susurro" de nuestra aplicación.

El resto del artículo está organizado de la siguiente manera: el apartado 2 describe el desarrollo del software utilizado en los experimentos y la tecnología relacionada, incluyendo los algoritmos de auralización utilizados. La sección 3 describe el experimento subjetivo y sus objetivos. El apartado 4 expone los resultados así como su análisis y discusión. Por último la sección 5 presenta las conclusiones de este trabajo.

2. IMPLEMENTACIÓN

Para la realización del test se ha utilizado una aplicación de multiconferencia desarrollada en el grupo de investigación y que funciona sobre teléfonos inteligentes y tabletas. El software permite a los participantes controlar fácilmente su uso por medio de una interfaz gráfica y táctil. Esta aplicación utiliza audio binaural usando un modelo simple de HRTF (*Head-Related Transfer Function*). Cada usuario puede mover y colocar virtualmente a los otros participantes alrededor de su propia posición en la mitad del borde inferior. Se puede encontrar una descripción completa de la aplicación en [9].

2.1. Sonido binaural

En la práctica, las funciones HRIR (*Head-Related Impulse Response*) tienen una longitud entre 128 y 512 muestras. Convolucionar cada señal mono con las dos HRIR supone un importante coste computacional. A pesar de que estas convoluciones se puedan realizar en el dominio frecuencial con multiplicaciones FFT (usando algoritmos *overlap-add* u *overlap-save*) de un modo eficiente, este proceso sigue teniendo un gran coste. Y aunque hoy en día los teléfonos inteligentes tienen suficiente potencia para realizar en tiempo real estos algoritmos de filtrado, se tendría que utilizar un gran parte de los recursos computacionales en esta etapa de procesado, reduciéndose también la duración de la batería.

Por estas razones se ha pensado en una implementación eficiente de la HRTF para esta aplicación. Para simplificar mucho el coste computacional se ha implementado una aproximación de la HRTF que proporciona una síntesis suficientemente buena.

La implementación de la HRTF se ha dividido en dos partes. Primero se ha implementado la ILD (*Interaural Level Difference*) con dos filtros IIR (*Infinite Impulse Response*). Segundo, la ITD (*Interaural Time Difference*) se ha conseguido añadiendo un retardo temporal entre las señales del oído izquierdo y derecho. El diseño de los filtros ILD IIR se ha realizado siguiendo un procedimiento similar a [10], donde los autores obtuvieron un modelo HRTF estándar promediando las respuestas de una base de datos de respuestas HRTF reales. Usando la respuesta promediada, se ha ajustado un filtro IIR de 6º orden paramétrico para cada dirección azimut. Se pueden encontrar más detalles de esta implementación en [9].

2.2. Procesado binaural de la distancia y el campo cercano

La sensación de distancia se consigue aplicando una atenuación a la señal directa dependiente de la posición de la fuente sonora y añadiendo las técnicas de auralización que se describen en la siguiente subsección. Para distancias más allá de 1 m, la HRTF prácticamente no varía para un ángulo de llegada dado ya que las ondas sonoras se pueden considerar planas. Pero para distancias más bajas (campo cercano), Brungart y Rabinowith [11] mostraron que la HRTF es dependiente de la distancia debido a la interacción de la curvatura del frente de onda con la cabeza, el torso y el pabellón auricular. De esta forma la ILD (diferencia de nivel interaural) aumenta para cortas distancias, habiendo mayor diferencia de nivel en altas frecuencias.

Varios trabajos, como [12,13], describen métodos para sintetizar HRTFs de campo cercano a partir de HRTFs de campo lejano obteniendo una gran fidelidad en la localización sonora. En esta aplicación de multiconferencia no es necesario conseguir una correspondencia exacta de la distancia, es más importante la percepción de distancias relativas, por eso se han evitado algoritmos complejos para campo cercano y se aplica una aproximación mucho más simple.

El algoritmo que se utiliza cuando las fuentes sonoras están muy cerca del sujeto aplica una atenuación de ganancia al oído opuesto dependiendo de la distancia y el ángulo, obteniendo la mayor diferencia cuando la fuente sonora está pegada a un oído. En esta simple aproximación no se está teniendo en cuenta el comportamiento para diferentes frecuencias que podría mejorar la sensación de distancia.

2.3. Auralización

La aplicación simula que los participantes están en una habitación virtual cuyo tamaño se extiende por la pantalla. El algoritmo de auralización tiene que simular las reflexiones y otros efectos acústicos que suceden en una sala teniendo en cuenta la posición de los participantes y del oyente. Existen muchos métodos precisos para modelar la respuesta al impulso de una sala, RIR (*Room Impulse Response*), entre dos puntos de la sala, muchos de ellos basados en modelos físicos de la propagación del sonido. Para esta aplicación no es necesario un método demasiado preciso, basta con uno que produzca una sensación realista.

La RIR se describe generalmente como la composición de una señal directa, primeras reflexiones y una cola de reverberación. Nuestra aplicación calcula las primeras reflexiones usando el método de la imagen. Se tienen en cuenta los cuatro ecos que proceden de las reflexiones de primer orden de las paredes laterales. Adicionalmente la aplicación procesa las reflexiones binauralmente considerando la dirección de llegada. Para la cola de reverberación se utiliza un reverberador clásico de Schroeder [14] ajustando los coeficientes al tiempo de reverberación deseado.

3. DESCRIPCIÓN DEL EXPERIMENTO

Se ha diseñado el experimento para encontrar las mejores técnicas y parámetros de auralización y procesado de campo cercano para un sistema de multiconferencia. Para medirlo, los participantes han probado el sistema con diversas opciones y han puntuado su percepción de la inteligibilidad, sensación de distancia, inmersión y preferencia general.

El test se ha realizado con una tableta iPad usando auriculares Sennheiser HD 439. Estos auriculares son de tipo cerrado para tener un buen aislamiento del ruido externo y circumaurales porque producen menos fatiga y son más cómodos de llevar puestos. En el experimento han participado un total de 10 personas (6 hombres y 4 mujeres), todos sin problemas de audición con edades comprendidas entre los 23 y 42 años, siendo estudiantes y personal de diferentes grupos de investigación de la universidad.

Por medio de nuestra aplicación en la tablet se presentó a los participantes una grabación de voz. Se invitó a los participantes a mover interactivamente el avatar con la pantalla táctil a diferentes posiciones de la sala virtual todo lo que quisieran durante el test para que notaran como la percepción espacial cambia con la posición y con las diferentes técnicas de procesado. La aplicación muestra unos botones (figura 1) en la pantalla para cambiar de forma ciega entre las diferentes opciones de procesado. Además cada test fue repetido dos veces por cada participante con dos voces grabadas diferentes (una voz de hombre y otra de mujer). El experimento se dividió en dos partes A y B.

3.1. Experimento A

El propósito de este primer experimento es encontrar la técnica de auralización más adecuada teniendo en cuenta los resultados obtenidos en el trabajo anterior de estos autores [8]. En dicho trabajo se descubrió que los participantes rechazaban una auralización fuerte al no ser confortable para este tipo de aplicación, pero los resultados mostraron que una auralización menor podía servir para conseguir cierta sensación de distancia y presencia sin variar mucho la inteligibilidad. Por ello, para este test es interesante ajustar una auralización más sutil. Se les pidió a los participantes puntuar de forma ordenada cada opción técnica evaluando 4 parámetros: inteligibilidad, sensación de distancia, inmersión y preferencia general.

Se usan las dos partes principales del RIR: las primeras reflexiones y la reverberación tardía. Para las primeras reflexiones se había encontrado previamente que una pequeña sala virtual

(5x3.75m) conseguía mejores resultados que una sala más grande, pero el alto coeficiente de reflexión (r=0.9) de las paredes usado en ambas salas era perjudicial para la sensación de distancia debido a la alta potencia de las reflexiones. Así que en este experimento se han usado las mismas dimensiones de la sala pequeña con un menor coeficiente de reflexión (r=0.3) de las paredes para compararlo con el previo. Para la reverberación se encontró previamente que era la mejor técnica para la percepción de distancia si se usaba aisladamente, pero parecía exagerada (se usó un tiempo de reverberación de 340 ms). Por eso en este experimento se ha usado un tiempo de reverberación menor (190 ms) comparándolo para lograr una experiencia más realista.

Al haber primeras reflexiones con baja (EE_L) y alta reflexión (EE_H), y reverberación tardía corta (R_S) y larga (R_L), y considerando que se pueden usar solos o combinados hay 8 posibles opciones. Las pruebas internas preliminares mostraron que las diferencias entre opciones eran demasiado pequeñas para compararlas todas a la vez. Así que se decidió dividir este test en dos fases más fáciles de comparar para los participantes:

- Fase A-1: en esta primera fase se pretende encontrar los mejores parámetros dentro de cada grupo de opciones similares. Por tanto los participantes tuvieron que comparar de forma separada los dos primeras reflexiones sin reverberación (EE_L y EE_H) entre ellas, las dos reverberaciones tardías sin primeras reflexiones (R_S y R_L) entre ellas, y las 4 combinaciones (EE_L+R_S, EE_L+R_L, EE_H+R_S y EE_H+R_L) entre ellas. Cuando todos los participantes acabaron, se evaluó sus puntuaciones en cada aspecto para elegir la mejor opción de primeras reflexiones, reverberación y combinación de ambas, obteniendo así tres técnicas para la siguiente fase.
- Fase A-2: en la segunda fase tuvieron que comparar únicamente las tres mejores opciones de la fase anterior (se analiza en siguiente sección) de forma que las diferencias entre ellas fueran más fáciles a la hora de puntuar cada parámetro.

3.2. Experimento B

En esta parte se ha evaluado el procesado de HRTF para campo cercano. Es importante potenciar el efecto de voz cercana para la aplicación desarrollada porque hay un “modo susurro” donde dos personas del grupo hablan privadamente como si estuvieran muy cerca.

Se pidió a los participantes que compararan la sensación de distancia cercana usando el algoritmo simple explicado en la sección 2.2 y sin él. Como en el experimento A, los participantes repitieron el test con dos voces grabadas distintas de forma ciega, es decir, sin saber que técnica se escondía detrás de cada botón de opción.

4. RESULTADOS Y DISCUSIÓN

4.1. Experimento A-1

Los participantes evaluaron cuatro parámetros (inteligibilidad, sensación de distancia, inmersión y preferencia general) para 3 grupos de comparación.

La primera comparación se hizo aplicando solo primeras reflexiones a la señal con coeficientes de reflexión bajo (r=0.3) y alto (r=0.9) de las paredes. La figura 2 muestra como al usar bajo coeficiente de reflexión (EE_L) se obtiene mejor resultados para todos los parámetros. Es lógico puesto que un sonido reflejado alto puede enmascarar el sonido directo y producir una señal más sucia y confusa a los receptores.

La segunda comparación se hizo aplicando solo la reverberación tardía a la señal con un tiempo de reverberación más corto (190 ms) y otro más largo (340 ms). La figura 3 muestra que una reverberación más corta (R_S) obtiene mejores resultados excepto para la inmersión.

En la tercera comparación se evaluaron las 4 combinaciones de primeras reflexiones y reverberación tardía. La figura 4 muestra que los mejores resultados para todos los parámetros se obtuvieron con la combinación de baja reflexión y larga reverberación (EE_L+R_L). Si se observan las combinaciones a pares, las combinaciones con alta reflexión tienen siempre los peores resultados. Significa que los malos resultados que dan la alta energía reflejada por las paredes virtuales en la primera comparación cuando estaban a solas se mantienen cuando se añade reverberación. Contrariamente, el comportamiento de la reverberación tardía es diferente cuando se le añaden las primeras reflexiones, reduciéndose mucho las diferencias percibidas entre colas de reverberación cortas y largas. Es más, la cola larga de reverberación es puntuada ligeramente mejor que la corta cuando se mezcla.

Por último, se puede señalar que la inmersión no tiene resultados muy claros, las puntuaciones son muy parecidas en todas las opciones. Al preguntar a los participantes qué habían considerado al evaluar la inmersión mostró que significaba algo distinto para cada uno porque cada uno esperaba estar en un diferente tipo de sala virtual, así que lo que percibían no siempre se correspondía con lo que esperaban perdiendo por tanto inmersión.

4.2. Experimento A-2

Se han evaluado los mismos cuatro parámetros (inteligibilidad, sensación de distancia, inmersión y preferencia general) pero ahora comparando únicamente las tres mejores opciones anteriores: solo primeras reflexiones con bajo (r=0.3) coeficiente de reflexión (EE_L), solo corta (190 ms) reverberación tardía (R_S) y la combinación EE_L+R_L (r=0.3 y 340 ms).

La figura 5 muestra como usando únicamente la reverberación tardía (R_S) se obtienen los mejores resultados. Es la mejor opción no solo para tener una buena sensación de realismo y distancia sino también por ser la más inteligible y preferida por los participantes.
Mezclar ambas técnicas de primeras reflexiones y reverberación tardía es la opción menos preferida por los participantes teniendo menos inteligibilidad e inmersión. Pero la percepción de la distancia mejora cuando la reverberación tardía se añade a las primeras reflexiones. No es un comportamiento nuevo; lo mismo se obtuvo en el trabajo anterior de estos autores [8].

4.3. Experimento B

Este experimento fue realmente simple: se preguntó por la mejor sensación de distancia para posiciones de campo cercano, es decir, cerca de la cabeza. La figura 6 muestra que aplicando el algoritmo binaural simple para campo cercano obtuvo apenas un poco mejor resultado que sin aplicarlo a cortas distancias. En cualquier caso, considerando el intervalo de confianza, no es una gran diferencia aunque quizás interesante al usar tan poca CPU para este algoritmo.

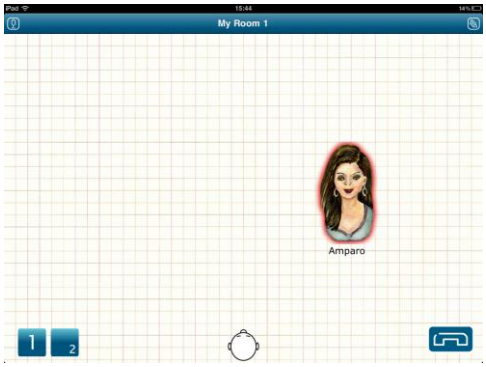

Figura 1. Pantalla principal de la aplicación en tableta

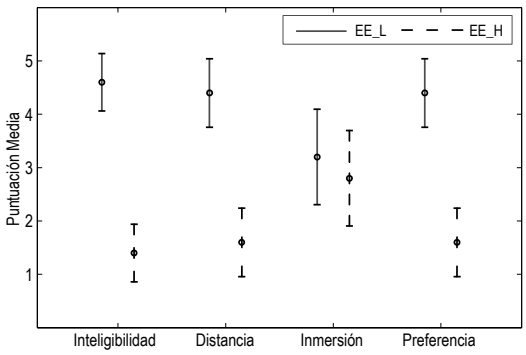

Figura 2. Comparación coeficiente reflexión bajo (EE_L) y alto (EE_H) usando primeras reflexiones (IC 95%)

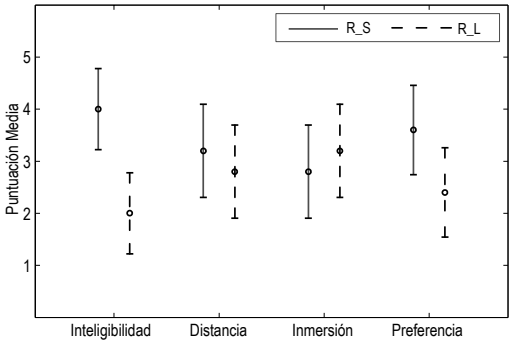

Figura 3. Comparación cola de reverberación corta (R_S) y larga (R_L) usando reverberación tardía (IC 95%)

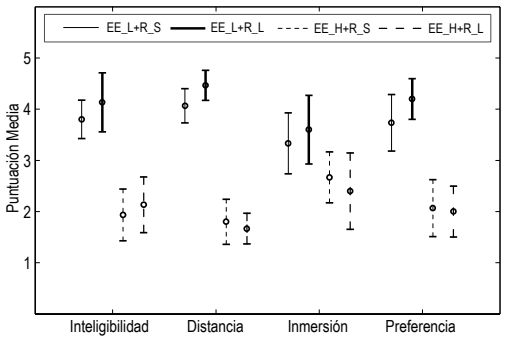

Figura 4. Comparación combinando primeras reflexiones y reverberación tardía (IC 95%)

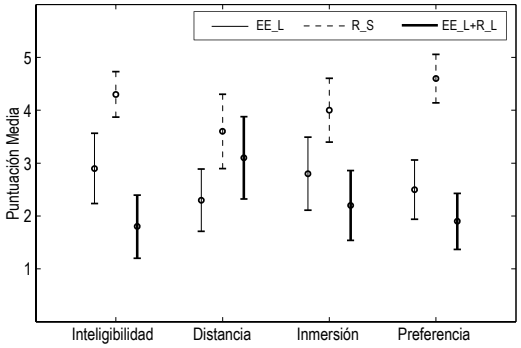

Figura 5. Comparación de las tres técnicas seleccionadas (IC 95%)

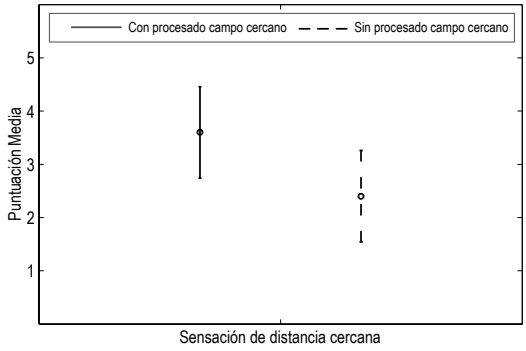

Figura 6. Sensación de proximidad para procesado binaural de campo cercano (IC 95%)

5. CONCLUSIÓN Y TRABAJO FUTURO

En este trabajo se han realizado varios experimentos subjetivos en el contexto de un sistema de teleconferencia binaural donde se aplica auralización. El principal objetivo ha sido evaluar y hacer un ajuste fino de variaciones sutiles relacionadas con la composición de la respuesta al impulso de la sala, como las primeras reflexiones y la reverberación tardía. Adicionalmente se ha evaluado la percepción de la distancia en campo cercano con un algoritmo muy simple. En las secciones anteriores se analizan en detalle la inteligibilidad, sensación de distancia, inmersión y preferencia general obtenidas pudiéndose extraer dos conclusiones principales:

- La sensación de distancia en un sistema de multiconferencia binaural se puede mejorar y mantener a la vez una buena inteligibilidad añadiendo una auralización suave, siendo la mejor técnica sintetizar únicamente una cola moderada de reverberación tardía de unos 200 ms. Se puede realizar utilizando muy poca CPU siendo suficientemente bueno para los usuarios de este tipo de servicios.
- Cuando percibir cortas distancias es importante en una aplicación (como en el “modo susurro” para tener conversaciones privadas), se pueden aplicar técnicas binaurales de campo cercano. Un algoritmo muy simple aumentando el nivel de diferencia interaural al acercarse una fuente sonora al oído mejora el efecto de fuentes cercana con muy poco coste computacional.

En cualquier caso, el realismo que se obtiene al implementar el procesado de HRTF en campo cercano se puede mejorar mucho si se usara un filtro dependiente de la frecuencia. Como

trabajo futuro para este tipo de sistemas se debería desarrollar y probar algoritmos de bajo coste cambien las diferencias de nivel interaural dependiendo de la frecuencia.

AGRADECIMIENTOS

Financiado por el Ministerio Español de Ciencia e Innovación, proyecto TEC2012-37945-C01.

REFERENCIAS

[1] G. M. Olson and J. S. Olson, “Distance matters”, Human-Computer Interaction, Volume 15, pp. 139-178, Lawrence Erlbaym Associates Inc, 2000.

[2] J. Sussman, J. E. Christensen, S. Levy, W. E. Bennett, T. V. Wolf, T. Erickson and W. A. Kellogg, "Rendezvous: Designing a VoIP conference call system", UIST, 2006.

[3] S. Deo, M, Billinghurst, N. Adams and J. Lehikioinen, “Experiments in spatial mobile audio-conferencing”. Proc. of the 4th international conference on mobile technology applications and systems and the 1st international symposium on Computer human interaction in mobile technoloy,vol. 7, ACM Press, pp. 447-451, 2007.

[4] S. Goose, J. Riedlinger and S. Kodlahalli, “3D audio conferencing and archiving services for handheld wireless devices”, Int. Journal of Wireless and Mobile Computing, vol. 1(1), pp. 5-13, 2005

[5] R. Kilgore, M. Chignell and P. Smith, Paul, “Spatialized audioconferencing: what are the benefits?”, Proc. of the 2003 conference of the Centre for Advanced Studies on Collaborative research, pp. 135-144,IBM Press, 2003

[6] J. Baldis, “Effects of spatial audio on memory, comprehension, and preference during desktop conferences”, Proc. of the SIGCHI Conf. on Human Factors in Computing Systems, pp. 166-173, 2001

[7] B. Shinn-Cunningham, “Learning reverberation: Considerations for spatial auditory displays,” in Proceedings of the ICAD, 2000, pp. 126–134.

[8] J.J.Lopez, E. Aguilera, P. Gutierrez, “A study of the influence of auralization on speech intelligibility and immersion in multi-party teleconferencing systems using binaural audio”, EAA Forum Acusticum, Krakow, 2014.

[9] E. Aguilera, J.J.Lopez, P.Gutierrez, M. Cobos, “An immersive multi-party conferencing system for mobile devices using binaural audio”, Proc. of the AES 55th Conference on Spatial Audio, Helsinki, 2014

[10] J.J. López, M. Cobos, B. Pueo, “Elevation in wave-field Synthesis using HRTF cues”, Acta Acustica, 96 (2), pp. 340-350, 2010.

[11] D. Brungart and W. Rabinowitz, “Auditory localization of nearby sources. Head-related transfer functions”, Journal Acoustic Society of America 106, pp 1465-1479, 1999.

[12] A. Kan, C. Jin, A. van Schaik, “A psychophysical evaluation of near-field head-related transfer functions synthesized using a distance variation function”, Acoustical Society of America, pp. 2233-2242, 2009.

[13] K. Nguyen, T. Carpentier, M. Noisternig, O. Warusfel, “Calculation of head related transfer functions in the proximity region using spherical harmonics decomposition: comparison with measurements and evaluation”, Proc. 2nd Int. Symp.

[14] M. R. Schroeder, ``Natural-sounding artificial reverberation,'' Journal of the Audio Engineering Society, vol. 10, no. 3, pp. 219-223, 1962.