

ACONDICIONAMIENTO ACÚSTICO DE SALAS PARA VIDEOCONFERENCIAS

PACS: 43.55.Br

Manuel A. Sobreira-Seoane¹, Fernando J. Mato-Méndez²
Departamento de Teoría de la Señal y Comunicaciones. Grupo Sonitum ®
E.T.S.I. de Telecomunicación, Universidad de Vigo
Campus de la Universidad de Vigo,
36200 Vigo
E-mail: ¹ msobre@gts.tsc.uvigo.es ; ² fmato@tsc.uvigo.es

ABSTRACT

During the last five years many new videoconference facilities are being installed. Many videoconference technicians rely on the echo cancellers to avoid the effect of feedback. The installation of new high bandwidth facilities for large-scale distributed meetings implies the use of large rooms which often present acoustic problems that cannot be skipped just relaying on the echo cancellers performance: the high audio quality of these new facilities make evident possible acoustic problems in the remote rooms that are masked in low quality systems. The authors have been dealing with this kind of design and concluded that this purpose the acoustic conditions and treatment for classical classrooms should be re-defined in several senses. In this paper the considerations for acoustic design for this kind of rooms are discussed.

RESUMEN

En los últimos años se están instando nuevas salas de videoconferencia de altas prestaciones. Los problemas acústicos de muchas instalaciones de este tipo estaban enmascarados por la baja calidad de la señal de audio en los sistemas tradicionales. Los sistemas de alta calidad ponen en evidencia los posibles problemas acústicos que no pueden ser resueltos confiando en las prestaciones de los canceladores de ecos. Los autores han tratado con el diseño de este tipo de salas, y concluido que las condiciones acústicas para aulas y salas de conferencias tradicionales deben ser reconsideradas para este tipo de entornos. En este artículo se discute algunas de las consideraciones realizadas para el diseño de aulas para videoconferencia de alta calidad. Estas aulas están agrupadas con la denominación Access Grid®, una agrupación que da soporte de intercomunicación de altas prestaciones entre grupos de investigación e instituciones, de gran interés sobre todo para reuniones a gran escala.

INTRODUCCIÓN

Los sistemas tradicionales para videoconferencia no necesitan en general unas condiciones acústicas especiales, debido a que el bajo ancho de banda para transmitir la señal de audio enmascara ciertos problemas acústicos y además, prácticamente siempre se culpa a las limitaciones del propio sistema como responsable de la baja calidad de audio. Los canceladores de eco suelen corregir de forma adecuada el un ligero exceso de reverberación,

especialmente porque tradicionalmente estas videoconferencias se mantienen entre salas de dimensiones reducidas (y que rara vez exceden los 0.7 s de reverberación). En salas de alta calidad, donde se desarrollan varias videoconferencias simultáneas, que son proyectadas de forma sincronizada en pantallas de gran formato y con varios proyectores a la vez, las cosas cambian. Estas salas, suelen tener un gran tamaño, las audiencias suelen ser grandes, y además en algunas videoconferencias se utilizan varios idiomas, simultáneamente por lo que se debe tener especial cuidado en mantener una elevadísima inteligibilidad (bien evaluada a partir del STI o del índice C_{50}). Debe tenerse en cuenta que los valores de inteligibilidad correctos para entender un mensaje en el idioma nativo, pueden resultar insuficiente para muchos oyentes al escuchar un idioma, que aunque dominen, no sea el nativo. Los usuarios de un sistema de altas prestaciones que reciben el sonido de un aula remota con exceso de reverberación y reflexiones con alto nivel de presión sonora que alcanzan la posición de los micrófonos de toma, describen la sensación como “escuchar al conferenciante dentro de un gran túnel metálico”. Este efecto se debe a la acción combinada del exceso de reverberación y el efecto de “filtrado peine” que introduce la acción de los canceladores de eco, resultando una calidad sonora absolutamente inaceptable en un entorno de este tipo. En la referencia [1] se describe el efecto de la degradación de la inteligibilidad en las comunicaciones clase-clase.

Como se ha comentado en el resumen, el aula tratada pertenece a la agrupación Access Grid[®], cuyo objetivo es proporcionar al usuario un altísimo potencial de interacción con múltiples usuarios remotos, suministrando tanto video como audio de alta calidad. Cada institución tiene una o más salas de videoconferencia, dotada de recursos multimedia de alta calidad. La figura 1 muestra la distribución en planta del aula Access Grid[®] de la Universidad de Vigo. Tanto la audiencia como los conferenciantes se orientan hacia una pantalla de gran formato, donde varias videoconferencias se proyectan simultáneamente, utilizando varios proyectores sincronizados y gestionados desde un ordenador de control. Para que ambos (audiencia y conferenciantes) puedan verse mutuamente, la sala se dispuso en “V”. Esta disposición además tiene la ventaja de que evita que las reflexiones de paredes paralelas alcancen a los micrófonos de forma directa, evitando el consecuente efecto de coloración indeseada del sonido.

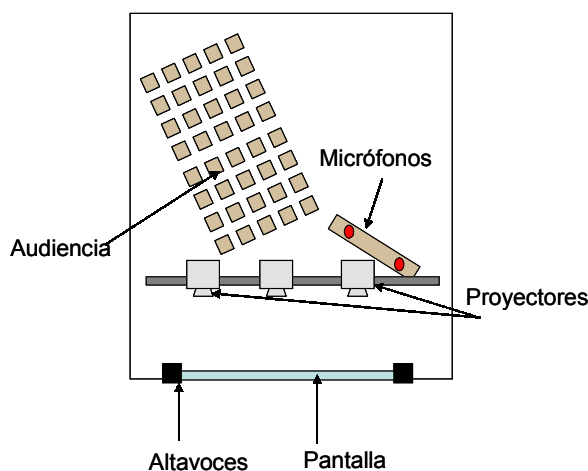


Figura 1. Distribución en planta del aula Access Grid[®] de la Universidad de Vigo.

La figura 2 muestra algunas fotografías de la posición de la audiencia, la pantalla y los conferenciantes. El color de las paredes (denominado azul francés) tiene la característica que proporciona un contraste adecuado y la coloración de la piel resulta más natural, con lo que la elección de cualquier material acústico para reforma de este tipo de sala, debe tener en cuenta la disponibilidad de este color (o la posibilidad de fabricar el material a medida).



Figura 2. Imagen de parte de la audiencia y la posición de los conferenciantes, y vista de la pantalla en el aula Access Grid © de la Universidad de Vigo.

DISEÑO ÓPTIMO DE SALAS PARA VIDEOCONFERENCIA

El diseño de espacios para palabra, debe basarse en la optimización del tiempo de reverberación y del Índice de Claridad para la Palabra. El diseño óptimo de una sala para videoconferencia no dista mucho del utilizado para el diseño de aulas:

- El nivel de ruido de fondo debe ser lo más bajo posible.
- El tiempo de reverberación: aunque en aulas con inteligibilidad adecuada se pueden medir reverberaciones que oscilan entre 0.5 y 1 segundo dependiendo de su volumen, para salas de videoconferencia el tiempo de reverberación en frecuencias medias nunca debería superar 0.7 s siempre que se haya optimizado los índices de claridad en toda la sala mediante la utilización de elementos reflectores.
- El índice de claridad para la palabra debe ser mayor de 2 dB.

Los índices basados en las relaciones energéticas, *ERL Early-to-Late*, se han propuesto para evaluar la claridad tanto de la música como de la palabra. La ecuación (1) muestra la definición general de estos parámetros, donde t_i es el tiempo expresado en milisegundos, utilizado para dividir la respuesta al impulso de la sala en dos fracciones de energía: temprana y tardía. Estos parámetros se pueden interpretar como relaciones Señal/Ruido, donde las reflexiones significativas que alcanzan al oyente en el intervalo inicial de t_i ms serían las consideradas como “señal”, mientras que la energía tardía podría ser interpretada como el “ruido” o el factor que produce una contribución negativa, disminución, de la claridad de la señal.

$$C_{ti} = 10 \log \frac{\int_0^{t_i} p^2(t) \cdot dt}{\int_{t_i}^{\infty} p^2(t) \cdot dt} \quad dB \quad (1)$$

En salas dedicadas fundamentalmente a palabra (salas de conferencia, aulas, teatros) se utiliza el índice de claridad para la palabra, y el tiempo de integración t_i se toma igual a 50 ms, valor que está relacionado con la respuesta subjetiva a la señal de voz. En el caso de salas donde predomine su utilización musical, se utiliza el índice de claridad musical con un tiempo de integración superior ($t_i=80$ ms).

Banda de Frecuencia	250	500	1K	2K	4K
SPL	22%	46%	20%	3%	2%
Contribución a la inteligibilidad de la palabra	5%	13%	20%	31%	26%

Tabla 1. Contribución a la Inteligibilidad de la Palabra en cada banda de frecuencia.

La tabla 1 muestra la contribución de cada banda de frecuencia a la inteligibilidad global de la palabra. Esta información la utiliza Marshall [2] para definir el Índice de Claridad para la Palabra como una media ponderada de las distintas bandas de frecuencia normalizada en frecuencias superiores a los 500 Hz, como indica la ecuación (2).

$$C_{50}(\text{Speech Average}) \equiv 0,15 \cdot C_{50}(500\text{Hz}) + 0,25 \cdot C_{50}(1\text{KHz}) + 0,35 \cdot C_{50}(2\text{KHz}) + 0,25 \cdot C_{50}(4\text{KHz}) \quad (2)$$

Siempre que se mantenga el nivel de ruido de fondo suficientemente bajo, la optimización del C_{50} conduce a valores elevados del STI. Los autores prefieren la utilización del C_{50} como descriptor de la inteligibilidad de la palabra, ya que puede darse el caso de que en determinadas salas un valor bajo de STI no se corresponda con una inteligibilidad inadecuada. Pensemos que la medición del STI se basa en la pérdida del índice de modulación [5-8]. Esta pérdida va a ser elevada siempre que tengamos reverberaciones elevadas (superiores a 1s), sin embargo, reforzando las primeras reflexiones puede incrementarse el C_{50} mejorando la inteligibilidad. Técnicas de dimensionado adecuado, utilizando los contornos de la sala como superficies reflectores y/o difusoras como el planteado por Arau [9] pueden dar lugar a inteligibilidades muy altas, con valores del STI relativamente bajos y tiempos de reverberación en torno a los 2s. La ecuación (1) podría ser interpretada en el siguiente sentido: para dos salas con la misma reverberación y por tanto con la energía tardía, la que tenga mayor fracción de energía temprana tendrá un valor superior de Claridad y por tanto de Inteligibilidad. Si se utilizan las paredes laterales y parte del techo para controlar la reverberación (en el caso de corrección acústica de una sala mal dimensionada, con exceso de reverberación), se pueden localizar las superficies que proyectan la fracción de energía temprana y situar en estas zonas elementos reflectantes y de esta forma optimizar la relación entre la energía temprana y la tardía.

Para aplicaciones de Videoconferencia se debe considerar los puntos dónde están situados los micrófonos y evitar que a estos lleguen reflexiones con elevada energía, bien orientando adecuadamente los micrófonos (que suelen ser cardioides) o ayudando con la colocación de superficies absorbentes en las zonas donde tengan lugar estas reflexiones potencialmente dañinas.

La Utilización de Reflectores

Los reflectores pueden definirse como elementos de acabado suave, rígidos (no deben actuar como membranas) y no porosos (su coeficiente de absorción debe ser lo más bajo posible, idealmente muy próximo a cero). Nos planteamos como objetivo incrementar la fracción de energía temprana, manteniendo la energía tardía prácticamente inalterada en el aula. Si se consigue este objetivo, el índice de claridad para la palabra debería incrementarse de forma ostensible. Para reflectores planos, debe considerarse su respuesta en frecuencia y el comportamiento descrito por Rindel [13], que muestra la ineficiencia en baja frecuencia de estos elementos. Esta ineficiencia por debajo de cierta frecuencia de corte, puede ser utilizada a favor para incrementar la fracción de energía temprana por encima de los 500 Hz y así orientar el diseño a optimizar explícitamente el valor de la ecuación (2).

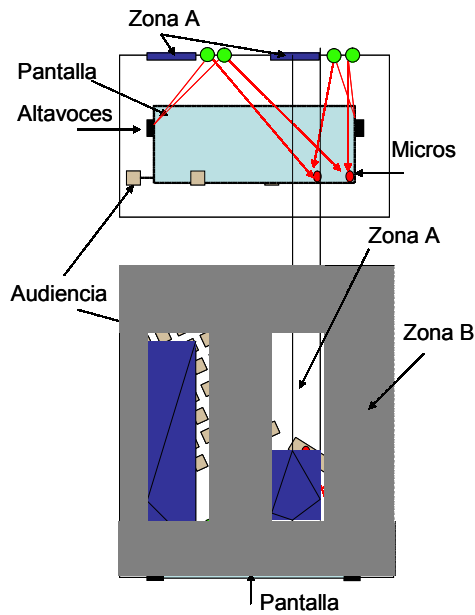


Figura 3. Esquema de perfil y planta de la sala de videoconferencias: Zona A, difusor; Zona B, alta absorción.

La figura 3 muestra el procedimiento para obtener la posición y el tamaño mínimo de la zona reflectante. Se asume que las últimas fila necesitan el aporte de reflexiones para mejorar tanto la sonoridad como la inteligibilidad de la palabra. Apoyándonos en la teoría de la imagen virtual, podemos obtener tanto la forma como la cobertura del reflector, seleccionando la parte de la audiencia que deseamos mejorar acústicamente. Debido a la distribución de esta sala en concreto, las zonas reflectantes son grandes y se obtienen reflexiones útiles efectivas a partir de frecuencias superiores a los 250 Hz. En la figura, los círculos indican dónde tienen lugar las reflexiones que alcanzan a los micrófonos desde el techo, y puede comprobarse que se ha tenido el cuidado de que estos puntos se sitúen sobre la zona B.

Mejora de las condiciones acústicas de la sala.

En la tabla 2 se muestran los valores medidos antes de la reforma. Los resultados evidencian unas condiciones acústicas muy deficientes, con reverberación excesiva y un índice de claridad inadmisibles. La acción de los canceladores de eco y el acoplamiento con la sala remota (que tendrá su propia acústica) implica que en estas condiciones el oyente de la sala remota tendrá que hacer grandes esfuerzos para poder entender mínimamente al hablante, sin hablar de la malísima calidad del audio reproducido,

<i>Tiempo de Reverberación</i>							
	125	250	500	1000	2000	4000	8000
P1	1.42	1.62	1.37	1.22	1.23	1.2	0.9
P2	1.56	1.25	1.39	1.16	1.23	1.15	0.85
<i>Índice de Claridad para la Palabra</i>							
	P1	-1.54					
	P2	-2.18					

Tabla 2. Tiempo de reverberación y claridad medidos antes de la reforma en dos puntos diferentes de la sala.

Para evaluar los cambios de reverberación y claridad antes de la reforma, se utilizó el Software Odeón, debido a la experiencia previa de los autores en la validación de resultados para este programa [14]. Como objetivo se planteó un tiempo de reverberación no superior a 0.7 segundos en frecuencias medias y un valor del Índice de Claridad de la Palabra superior a 2. Se planteó la utilización de material absorbente en las zona marcadas como B en la figura 3, y en la parte superior de las paredes laterales. Con esta estrategia se seleccionaron 80 m² de superficie que podría ser tratada acústicamente, y se proyectó sobre ella 1 cm de un spray de celulosa absorbente. Para los cálculos de reverberación previos, se tuvo en cuenta que en las paredes laterales (superficie rígida) el coeficiente de absorción de la celulosa es menor que en el proyectado sobre el techo (PYL suspendidas). La tabla 3 muestra el tiempo de reverberación y claridad medidos después de la reforma conjuntamente con los valores estimados en la fase de proyecto. Puede observarse cómo la acción combinada de la reducción de reverberación y la utilización de superficie reflectora (que no es más en este caso

que dejar de proyectar la parte calculada), conduce a un incremento muy importante de la Claridad, mostrando la sala ahora valores de STI entre 0.7 y 0.85 en distintos puntos de la sala. El exceso de reverberación (esperado) en baja frecuencia tiene muy poca incidencia sobre la inteligibilidad de la palabra y se podría corregir utilizando un resonador de membrana en una de las paredes laterales. Esta corrección no se estima necesaria, ya que los actuales usuarios de la sala se muestran altamente satisfechos con los resultados y les gusta el carácter cálido que toma la voz en la sala. Los oyentes situados en aulas remotas han comunicado un excelente sonido, definido como limpio y con muna “gran presencia”, frente a la descripción anterior (hablante en un gran tanque de lata). Se considera por tanto el problema acústico totalmente resuelto.

<i>Tiempo de Reverberación</i>						
	125	250	500	1000	2000	4000
E	1.23	1.22	0.81	0.63	0.49	0.4
M	1.34	1.06	0.76	0.53	0.64	1.15
<i>Indice de Claridad para la Palabra</i>						
	E	6				
	M	5.5				

Tabla 3. Tiempo de reverberación y claridad Estimados (E) y medidos (M) después de la reforma.

CONCLUSIONES Y AGRADECIMIENTOS

Se ha presentado el proceso de mejora acústica de una sala de videoconferencias. La combinación de control de reverberación con la utilización de superficies para incrementar la energía útil proporciona excelentes resultados en la práctica. La sala después de la reforma muestra una excelente inteligibilidad, careciendo de coloraciones y efectos acústicos indeseados.

REFERENCIAS

- [1] Manabu Fukushima, Naoto Nakamura¹ and Hifofumi Yanagawa¹. “Evaluation of speech intelligibility for classroom-to-classroom collaborative leaning via multimedia network.” www.eecs.kumamoto.ac.jp/ITHETO1/proc/055.pdf
- [2] Marshall, L.G., “An acoustic measurement program for evaluating auditoriums based on the early to late sound energy ratio”. J. Acoust. Soc. Amer. 96, 2251=2261.
- [3] Thiele, R. “Richtungsverteilung und Zeitfolge der Schallrückwürfe in Räumen”, Acustica 3, pp. 291-302, 1953.
- [4] Houtgast, T., Steeneken, H. J. M., & Plomp, R. (1980). “Predicting speech intelligibility in rooms from the modulation transfer function”. Acustica, 16, 60–72.
- [5] Steeneken, H.J.M. and Hougast, T. “The Modulation Transfer Function in Room Acoustics as a Predictor of Speech Intelligibility”, Acustica 28, 66-73, 1973.
- [6] Hougast, T. and Steeneken, H.J.M. “A Physical Method for Measuring Speech Transmission Quality”, J. Acoust. Soc. Amer. 61, 318-326, 1980.
- [7] Hougast, T, Steeneken, H.J.M. and Plomp R. “Predicting Speech Intelligibility in Rooms from the Modulations Transfer Function. I. General Room Acoustics”, Acustica 46, 66-73, 1973.
- [8] Hougast, T. and Steeneken, H.J.M. “A Review of the MTF Concept in Room Acoustics and it’s Use for Estimating Speech Intelligibility in Auditoria”, J. Acoust. Soc. Amer. 77, 1060-1077, 1985.
- [9] Arau, H. ABC de la Acústica Arquitectónica. Grupo editorial CEAC. ISBN 84-329-2017-7.
- [10] Siebein, G. W., Lehde, M., Lee, H. W., Ashby, J., Gold, M. A, Hasell, M. J., Abbott, P., & Crandell, C. C. “ Classroom acoustics III: Acoustical model studies of elementary school classrooms”. P. K. Kuhl & L. A. Crum (Eds.), Proceedings of the International Congress on Acoustics (pp. 2725–2726). Seattle, WA: Acoustical Society of America, 1998.
- [11] Bradley, J.S., “Speech intelligibility studies in classrooms,” Journal of the Acoustical Society of America, 80(3), 1986, ppg 846-854.

- [12] Manabu Fukushima, Naoto Nakamura¹ and Hifofumi Yanagawa¹. "Evaluation of speech intelligibility for classroom-to-classroom collaborative learning via multimedia network." www.eecs.kumamoto.ac.jp/ITHET01/proc/055.pdf
- [13] Rindel, J.H. "Attenuation of sound reflexions due to diffraction". Nordic Acoustical Meeting, Aalborg, Denmark, 1986.
- [14] Sobreira-Seoane, M.A.. "On The Design Of Classrooms And Conference Rooms". 11th International Congress on Sound and Vibration, San Petersburg -2004.