

Taxonomía de sonidos y su interpretación física y psicoacústica a partir de un análisis factorial.

S. Cerdá¹, A. Giménez¹; J. Romero^{1 4}; R. Cibrián²; J. V. Garriges⁵; J.L. Miralles³,

¹ Grup d'Acústica Arquitectònica, Ambiental i Industrial. Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia. agimenez@fis.upv.es

² Facultat de Medicina. Universitat de València. Blasco Ibáñez s/n, Valencia, Rosa.M.Cibrian@uv.es

³ Facultat de Psicologia. Universitat de València. Blasco Ibáñez, 21, Valencia, mirallel@uv.es

⁴ Escuela Universitaria De Gandía, Carretera de Nazaret-Oliva, S/N, 46730 Grao de Gandía, Tel +34962849333, extensión 49400, romerof@fis.upv.es

⁵ Facultat de CC.Físicas. Universitat de València. Avda. Dr. Moliner s/n

PACS 43.60.+d, 43.66.+y

Resumen

En los últimos años es creciente el interés en el desarrollo de técnicas y criterios de clasificación de los sonidos tanto para simular el procesamiento auditivo de los mismos como su reconocimiento por máquinas. Los principios de clasificación dependen evidentemente de las características extraídas de la señal y de los criterios posteriores de agrupación. En este artículo se presenta un proceso de reducción de variables novedoso que será de ayuda para la clasificación de los sonidos. A partir del análisis en tercio de octava de una colección variada de sonidos que hemos utilizado en diversos estudios[1],[2], se ha realizado un análisis factorial. Dicho análisis ha convergido a cinco factores, con una varianza total explicada del 86.5%. Esto supone una gran avance, al conseguir *reducir* las 34 bandas de frecuencia estudiadas a solamente 5 factores. Además estos factores muestran correlaciones con algunos parámetros psicoacústicos, lo que nos permite proponer una interpretación de los mismos.

Summary

In the last years it's increasing in the interest in the development of techniques and criteria of sound classification to simulate the auditory processing or make possible its recognition by machines. The classification principles depend evi-

dently on the extracted characteristics of the signal and the criteria of grouping. In this article, a novel process of reduction of variables is presented. It will be an aid for sound classification. From the analysis in third of eighth of a varied collection of sounds, that we have used in several studies [1], [2], a factor analysis. This analysis has been made. This analysis has converged to five factors, that explain the 86.5% of the total variance. Reducing the 34 studied frequency bands to only 5 factors. In addition, these factors show correlations with some psychoacoustic parameters, this allows us to propose a possible interpretation.

Introducción

Los procesos de clasificación y categorización constituyen un potente y económico mecanismo cognitivo de adaptación en la medida en que permiten organizar el conocimiento de forma simplificada, facilitan el reconocimiento de los objetos y la generalización y disminuyen la carga de memoria. Clasificación y categorización no sólo implican una ventaja para el procesamiento humano de los sonidos sino que facilitan la aproximación al estudio del reconocimiento de los mismos por máquinas. En los últimos años, estos estudios han alcanzado un notable interés [3], [4], [5]. La clasificación de los sonidos se ha realizado utilizando criterios diferentes, objetivos o subjetivos:

- acústicos
- perceptivos
- semánticos
- atendiendo a la naturaleza de la fuente.

En cualquier caso todo proceso de clasificación en base a criterios objetivos exige al menos las tres fases siguientes de procesamiento:

- a. preprocesamiento de la señal
- b. extracción de características
- c. aplicación de criterios de clasificación a las características extraídas.

El preprocesamiento supone la captura de señal acústica y la digitalización de la misma mediante su cuantización [6].

La extracción de características, es probablemente el aspecto más complejo del tratamiento de la señal en base a la cantidad y variabilidad de los parámetros candidatos a ser extraídos. En general, y teniendo en cuenta las diferencias entre sonidos ambientales y sonidos del habla, la extracción de características frecuenciales, espectrales o temporales de la señal acústica se ha realizado utilizando técnicas de análisis como FFT, STFT (Short-Time Fourier Transform), CWT (Continuous Wavelet Transform), FWT (Fast Wavelet Transform) o HCC (Homomorphic Cepstral Coefficients) entre otras.

Las características extraídas se sintetizan finalmente en unidades informativas de orden superior que, en virtud de algún criterio, permiten agrupar los sonidos en clases o categorías.

El objetivo de esta investigación es encontrar una serie de parámetros que permitan la clasificación de los sonidos e interpretarlos. Para ellos hemos optado por un análisis espectral de una larga serie de sonidos que hemos utilizado en estudios sobre psicoacústica [1,2]. Se grabaron 60 sonidos ambientales y musicales según los criterios de van Derveer [7] con un DAT TASCAM a 44.1k Hz de frecuencia de muestreo. De cada uno de ellos se obtuvieron descriptores psicoacústicos relativos a: su intensidad, aspereza, tonalidad, fuerza de la fluctuación y agudeza, de acuerdo con Zwicker [8] analizados con el sistema Sound Quality de Brüel & Kjaer. El análisis espectral, en tercio de octavas, ha sido realizado con el analizador Artemis.

A partir del análisis espectral se ha realizado un análisis factorial cuyo resultado se presenta en la sección 5. La interpretación de los Factores se ha realizado buscando correlaciones con los parámetros psicoacústicos obtenidos para los sonidos estudiados.

Análisis espectral y parámetros psicoacústicos

Análisis se refiere a la acción de descomponer lo complejo en simple o identificar en ese algo complejo los componentes simples que lo forman. Como se sabe, hay una base o fundamento físico para modelar la luz, el sonido o las ondas de radio en superposición de diferentes frecuencias. Un proceso que cuantifique las diversas intensidades de cada frecuencia se llama análisis espectral.

Matemáticamente el análisis espectral está relacionado con una herramienta llamada transformada de Fourier o análisis de Fourier. Ese análisis puede llevarse a cabo para pequeños intervalos de tiempo, o menos frecuentemente para intervalos largos, o incluso puede realizarse el análisis espectral de una función determinista. Además la transformada de Fourier de una función no sólo permite hacer una descomposición espectral de los formantes de una onda o señal oscilatoria, sino que con el espectro generado por el análisis de Fourier incluso se puede reconstruir (sintetizar) la función original mediante la transformada inversa. Para poder hacer eso, la transformada no solamente contiene información sobre la intensidad de determinada frecuencia, sino también sobre su fase. Esta información se puede representar como un vector bidimensional o como un número complejo. En las representaciones gráficas, frecuentemente sólo se representa el módulo al cuadrado de ese número, y el gráfico resultante se conoce como espectro de potencia o densidad espectral de potencia.

Es importante recordar que la transformada de Fourier de una onda aleatoria, mejor dicho estocástica, es también aleatoria. Un ejemplo de este tipo de onda es el ruido o sonido ambiental. Por tanto para representar una onda de ese tipo se requiere cierto tipo de promediado para representar adecuadamente la distribución frecuencial. Para señales estocásticas digitalizadas de ese tipo se emplea con frecuencia la transformada de Fourier discreta. Cuando el resultado de ese análisis espectral es una línea plana la señal que generó el espectro se denomina ruido blanco por la analogía con la luz blanca.

Descripción de los parámetros psicoacústicos

Los parámetros psicoacústicos son magnitudes sensoriales que se toman como características del sonido de la audición humana. Constituyen un suplemento útil a la medida física objetiva y a la evaluación subjetiva del sonido. Los descriptores considerados son:

- **Loudness.** “Intensidad Sonora, Fuerza, volumen”, en *sonios. ISO 532B (ISO 1966)*. Es la magnitud más básica, describe generalmente mejor el nivel de sonido percibido que el nivel de presión sonora. El nivel de intensidad sonora en dBA se utiliza como pauta en todos los reglamentos. Sin embargo la intensidad psicoacústica de este parámetro o Loudness es una medida más conveniente para evaluar la opinión real de la intensidad, ella se adapta mejor a la audición humana (e.g. considerando el enmascaramiento).
- **Roughness.** “Aspereza, rugosidad” en *asper. Zwicker, 1982, Aures, 1984, Schiffbänker 1991*. En un ruido de banda ancha una modulación con dos tonos cuya diferencia se mantenga entre 13 hertzios y 300 hertzios, la sensación de la magnitud de la fluctuación da la impresión de la aspereza. La impresión de la aspereza maximiza en una frecuencia de modulación de cerca de 70 hertzios. Un carácter áspero de un sonido causa generalmente una impresión desagradable de la audición. Si la diferencia de la frecuencia de los dos tonos se excede 300 hertzios se consigue gradualmente la impresión de dos tonos distintos.
- **Tonality.** Tonalidad. La tonalidad de un ruido significa que un sonido contiene uno o más componentes prominentes tonales. Los sonidos tonales intensifican generalmente la impresión desagradable de un sonido. Por esta razón en pautas y regulaciones estos sonidos conllevan una penalización de 3 hasta 6 dBA. En los ruidos de trenes usualmente no aparecen tonalidades.
- **Sharpness.** “Nitidez, agudeza”, en *acum .von Bismarck 1974*. Sharpness representa un atributo para la evaluación del timbre. Se define en función de la sensación de placer o de desagrado de un sonido referido a toda la envolvente del espectro. Se determina por medio del centroide de la distribución espacial en el espectro de Zwicker, 1982. Cuanto mayor es la parte de alta frecuencia dentro de un ruido mayor es la impresión de la agudeza del sonido. La agudeza es el parámetro psicoacústico más importante debido a su influencia considerable en el desagrado “unpleasantness” de sonidos.

Análisis factorial

El Análisis Factorial es una técnica estadística que consiste en resumir la información contenida en una matriz de datos con V variables. Para ello se identifican un reducido número de factores F , siendo el número de factores menor que el número de variables. Los factores representan a las variables originales, con una pérdida mínima de información

Podemos distinguir entre Análisis Factorial Exploratorio, donde no se conocen los factores “a priori”, sino que se determinan mediante el Análisis Factorial y, por otro lado estaría el Análisis Confirmatorio donde se propone “a priori” un modelo, según el cual hay unos factores que representan a las variables originales, siendo el número de estos superior al de aquellos, y se somete a comprobación el modelo

Para que el Análisis Factorial tenga sentido deberían cumplirse dos condiciones básicas: Parsimonia e Interpretabilidad, Según el principio de parsimonia los fenómenos deben explicarse con el menor número de elementos posibles. Por lo tanto, respecto al Análisis Factorial, el número de factores debe ser lo más reducido posible y estos deben ser susceptibles de interpretación sustantiva. Una buena solución factorial es aquella que es sencilla e interpretable.

Colección seleccionada de sonidos

Se grabaron 60 sonidos naturales con un DAT TASCAM a 44.1k Hz de frecuencia de muestreo. Los sonidos fueron analizados con el software Sound Quality de Brüel & Kjaer para determinar los descriptores psicoacústicos: intensidad (loudness), agudeza (sharpness), aspereza (roughness) y tonalidad (tone to noise ratio y prominent ratio)

Los sonidos fueron igualados en nivel equivalente. Al igualarlos en nivel equivalente los parámetros relativos a la intensidad diferían poco entre sí (specific loudness media = 11 son; desviación típica = 2,5), en cualquier caso lejos de los 20 sonos que pueden ser considerados como molestos. Los demás parámetros psicoacústicos, independientes en mayor o menor medida de la intensidad, variaban de valores mínimos a elevados.

Los sonidos grabados corresponden a objetos y fuentes sonoras naturales de la experiencia cotidiana (motores, electrodomésticos, transportes, ruido blanco y rosa, instrumentos musicales).

Resultados

Las principales aplicaciones del análisis factorial son la reducción del número de variables y la detección de estructura en las relaciones entre variables. El análisis factorial que hemos realizado consiste en la extracción de las componentes principales, analizando la matriz de correlaciones, para valores propios superiores a uno. Hemos completado el proceso rotando los factores mediante el procedimiento *varimax*. Con ello hemos obtenido tres factores que explican el 86.5% de la varianza total. La tabla muestra las componentes obtenidas en función de los parámetros estudiados y sus correlaciones.

Matriz de componentes rotados(a)

	Componente				
	1	2	3	4	5
B10	,872	,154	,121	-,006	-,297
B12.5	,872	,154	,121	-,006	-,297
B16	,924	,147	,108	,041	-,237
B20	,908	,171	,122	,141	-,017
B25	,913	,172	,110	,157	-,005
B31.5	,912	,182	,087	,189	,020
B40	,934	,195	,060	,166	,029
B50	,933	,207	,034	,141	,065
B63	,917	,235	,027	,104	,120
B80	,891	,319	,040	,062	,190
B100	,865	,371	,076	,066	,224
B125	,798	,452	,022	,026	,310
B160	,747	,547	,022	,056	,272
B200	,605	,676	,000	,035	,210
B250	,548	,737	,013	,066	,243
B315	,624	,666	,071	,112	,261
B400	,331	,811	-,035	-,202	-,131
B500	,174	,765	-,080	-,006	-,059
B630	,443	,780	,035	,000	,271
B800	,213	,736	-,080	-,200	-,177
B1000	,194	,790	-,057	,104	,119
B1250	,238	,752	,028	,181	-,248
B1600	-,121	,559	-,056	,498	-,411
B2000	,158	,355	,044	,721	-,257
B2500	,099	-,132	-,062	,814	-,193
B3150	,145	,071	,131	,870	,024
B4000	,140	,007	,352	,866	,184
B5000	,121	-,125	,488	,781	,176
B6300	,123	-,093	,569	,723	,210
B8000	,146	-,059	,732	,554	,195
B10000	,222	-,047	,846	,321	,050
B12500	,065	-,003	,961	,082	,002
B16000	,064	-,015	,965	,070	-,056
B20000	,039	-,031	,944	,062	-,095

Método de extracción: Análisis de componentes principales.
Método de rotación: Normalización Varimax con Kaiser.

coeficientes ^a

coef. No estandarizados coef. estandarizados

MODELO	B	Error típ.	Beta	t	Sig.
constante	-1,329	,087		-15,361	,000
B12.5	-,026	,016	-,422	-1,654	,105
B16	,058	,019	,928	3,009	,004
B20	-,089	,062	-1,709	-1,3435	,158
B25	,153	,087	2,910	1,755	,086
B31.5	-,071	,032	-1,340	-2,205	,032
B40	,023	,039	,431	,579	,565
B50	-,007	,043	-,132	-,164	,871
B63	,023	,032	,443	,716	,478
B80	-,001	,015	-,020	-,067	,947
B100	-,010	,009	-,194	-1,133	,263
B125	,034	,009	,674	3,554	,001
B160	-,031	,007	-,600	-4,530	,000

a Variable dependiente: REGR factor score 1 for análisis 1

coeficientes ^a

coef. No estandarizados coef. estandarizados

MODELO	B	Error típ.	Beta	t	Sig.
constante	-3,847	,241		-15,966	,000
B200	-,025	,010	-,467	-2,497	,016
B250	,038	,013	,714	2,800	,007
B315	-,018	,006	-,334	-2,934	,005
B400	,020	,010	,325	1,975	,054
B500	,013	,009	,164	1,507	,138
B630	-,002	,008	-,029	,210	,835
B800	,004	,008	,055	,517	,607
B1000	,035	,01	,395	3,421	,001
B1250	,016	,007	,196	2,298	,020

a Variable dependiente: REGR factor score 2 for análisis 1

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Valor Mínimo	Tono complejo	Tono complejo tipo chirrido	Grillos	Nota de afinación	Nota de afinación fluctuante
Valor Máximo	Tren	Sirena barco	Chasquidos tipo eléctrico	Ruido fluctuante	Máquina rotora.

coeficientes ^a

coef. No estandarizados coef. estandarizados

MODELO	B	Error típ.	Beta	t	Sig.
constante	-,912	,063		-14,486	,000
B8000	-,004	,004	-,080	-1,126	,265
B10000	,013	,004	,251	2,956	,005
B12500	,026	,010	,536	2,713	,009
B16000	-,005	,015	,099	-,333	,740
B20000	,022	,008	,407	2,737	,008

a Variable dependiente: REGR factor score 3 for análisis 1

coeficientes ^a

coef. No estandarizados coef. estandarizados

MODELO	B	Error típ.	Beta	t	Sig.
constante	-5,761	,269		-21,406	,000
B2000	,017	,006	,159	3,020	,004
B2500	,048	,008	,336	5,799	,000
B3150	,021	,008	,206	2,547	,014
B4000	,033	,010	,411	3,151	,003
B5000	,021	,011	,315	1,971	,054
B6300	-,020	,008	-,330	-2,447	,018

c Variable dependiente: REGR factor score 4 for análisis 1

Para interpretar estos factores se han buscado correlaciones con los parámetros psicoacústicos fundamentales. Se ha encontrado que:

1. El factor 1 muestra correlación negativa con la Tonalidad (-0.83)
2. El factor 2 muestra correlación positiva con el Loudness (0.83)
3. El factor 3 muestra correlación positiva con el Sharpness (0.68)
4. El factor 4 no muestra correlaciones con los parámetros psicoacústicos estudiados.
5. El factor 5 muestra correlación positiva con Roughness (0.77) y Fluctuation (0.71).

Conclusiones

En esta investigación hemos presentamos un método reduccionista de variables que mejora la clasificación de los sonidos. A partir del análisis en tercio de octava de una colección variada de sonidos que hemos utilizado en diversos estudios, hemos realizado un análisis factorial. Di-

cho análisis ha convergido a cinco factores, con una varianza total explicada del 86.5%. Esto supone una gran simplificación, al conseguir *reducir* las 34 bandas estudiadas a solamente 5 factores. La interpretación de estos factores la hemos buscado comparándolos con los parámetros psicoacústicos más representativos, obteniendo que el Factor 1 es un factor relacionado con la Tonalidad (bandas de baja frecuencias). El Factor 2 con el Loudness (bandas 200 a 1250 Hz). El Factor 3 correlaciona con el Sharpness (bandas de alta frecuencia) y el Factor 5 correlaciona con el Roughness y la Fluctuación (aunque no correlaciona con ningún intervalo de bandas concreto). Por otro lado el Factor 4, de las bandas centrales no muestra correlación con ninguno de los parámetros psicoacústicos estudiados.

Bibliografía

- [1]. Jose Luis. Miralles, Jose.Vicente. Garrigues, Alicia. Gimenez, José. Romero, Joaquin. Navasquillo, Salvador. Cerdá. Medida de la calidad sonora en función de la duración. 36º Congreso Tecniacústica 2005.
- [2]. Miralles, José Luis; Garrigues, José Vicente; Giménez, Alicia; Romero, José; Cibrián, Rosa; Cerdá, Salvador. Effect Of Duration In The Perception Of Pleasantness Of Sound. 19th ICA 2007.
- [3]. Fuginaga, I. (1998). Machine recognition of timbre using steady-state tone of acoustical instruments. En ICMC, Ann Arbor, pp.207-210.
- [4]. Wold, E.; Blum, T.; Keislar, D. and Wheaton, J. (1999). Classification, research and retrieval of audio. En B. Furth (Ed.) CRC Handbook of Multimedia Computing, pp. 207-226. Boca Raton, FLA: CRC Press.
- [5]. Peeters, G.; McAdams, S. and herrera, P. (2000). Instrumento sound description in the context of mpeg-7. En ICMC, Berlin.
- [6]. Kefauver, A. The digital encoding process. Fundamentals of digital audio. (vol.14. The computer music and digital audio series). A-R Editions, Inc. Madison, Wisconsin.
- [7]. Van Derveer, N. (1979). Ecological acoustics: human perception of environmental sounds. Ph. Thesis, Cornell University,
- [8]. Zwicker, E. and Fastl, H. (1999). Psycho-acoustics Facts and Models. Ed. Springer, Berlin.

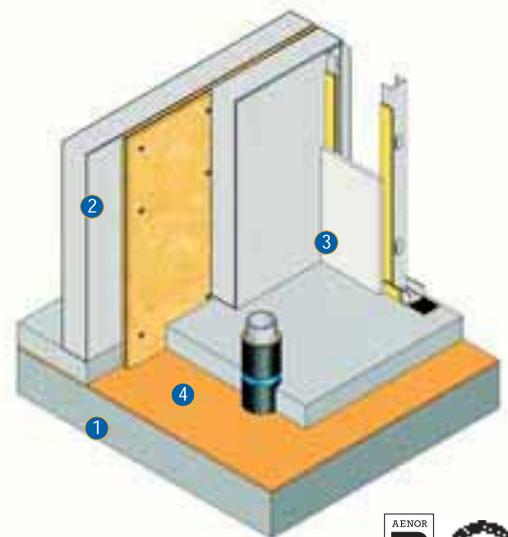
Bienvenido al silencio

Soluciones de Aislamiento Acústico para la Edificación

Nuestras soluciones superan las exigencias del nuevo Código Técnico. 15 años de experiencia en la búsqueda del confort acústico y más de 5000 obras, avalan la tecnología desarrollada por nuestro Departamento de I+D.

- **IMPACTODAN**
Aislamiento acústico de **forjados** > 50 dBA a ruido aéreo*
< 60 dBA a ruido de impacto*
- **DANOFON**
Aislamiento acústico de **medianeras** > 50 dBA a ruido aéreo*
- **FONODAN**
Aislamiento acústico de **tabiques de cartón yeso.** > 4dB de incremento de aislamiento*
- **ACUSTIDAN**
Aislamiento acústico de **bajantes.** < 30 dBA a ruido inmisión*

* Mediciones "in situ" para soluciones constructivas completas. Consulte nuestros catálogos técnicos.



UNE-EN ISO 9001: 2000

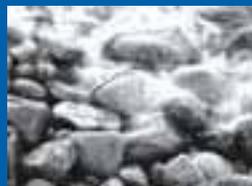
Confiamos tu tranquilidad.



Impermeabilizantes



Aislamiento Acústico



Drenajes y Geotextiles

danosa

www.danosa.com 900 211 081