OBJECTIVE EVALUATION IN VIRTUAL HALLS USING MUSIC FEATURES AND OBJECTIVE PARAMETERS

PACS no. 43.55.Cs, 43.55.Gx

Cerdá, Salvador[1]; Segura, Jaume[2]; Barba, Arturo[3]; Cibrián, Rosa[3]; Montell, Radha[4]; Giménez Alicia[4] .
1 E.T.S.I. del Diseño, Univ. Politècnica de València, Camí de Vera s/n. salcerjo@mat.upv.es
2 IRTIC – Universitat de València, C/ Catedràtic Dr José Beltrán, 2. jsegura@uv.es
3 Facultat de Medicina – Universitat de València. rosa.m.cibrian@uv.es
4 E.T.S.I. Industriales, Univ. Politècnica de València, Camí de Vera s/n. arbarse@doctor.upv.es , radmonse@upvnet.upv.es , agimenez@fis.upv.es

ABSTRACT

In this work, we study the relations between musical characteristics obtained with MIRToolbox and objective parameters from virtual acoustics models. Four channels of the music signal from different positions and different halls are auralized. The usual objective parameters are computed and, by statistical analysis, they are compared with musical characteristics. Room acoustics models have been calibrated with Odeon software.

RESUMEN

El presente trabajo está dedicado a estudiar las relaciones entre las características musicales obtenidas con el paquete MIRToolbox, y los parámetros objetivos en los modelos acústicos virtuales de salas. Para ello se auralizan cuatro canales de señal musical en diferentes lugares de diferentes salas y, mediante análisis estadístico, se comparan los parámetros objetivos habituales de la salas con las característica musicales. Los modelos acústicos han sido calibrados utilizando el programa ODEON.

1. INTRODUCTION

Auralizations are commonly used today by architectural acousticians as a tool to acoustically model sensitive spaces. The procedure involves inputting the room geometry and material properties, assigning the source and receiver characteristics and locations, and then simulating the room impulse responses (RIR). The predicted impulse responses can then be used to calculate room acoustic parameters and the corresponding auralizations for each source-receiver combination [1]. The directional characteris- tics of the source influence the room acoustic measurements and subsequent objective parameter calculations, as documented by Prince and Talaske [2] and San Martın et al. [3]. Studies on how source directivity impact on the results of room acoustic computer simulations and the subjective perception of auralizations have also been conducted [4,5]. Recently, studies on multichannel auralization techniques applied to solo instruments and room acoustic investigations of actor positions and orientations [6,7] have used simulations as a tool to evaluate the effects of different source variables on objective parameters. These recent studies serve as evidence of the reliability of virtual models.

The MIRtoolbox [8] is an integrated set of functions written in Matlab for extracting musical features from audio files. These features are organized into five main musical dimensions: (1) dynamics field as related to temporal changes in energy (2) rhythm field viewed from the musical point of view (3) timbre field in reference to the spectrum that is computed and analyzed by auditory models (4) pitch field related to fundamental frequency and harmonicity (5) tonal field computing features related to energy and its time evolution when associated with musical keys. A detailed description of all audio features is found in the MIR- Toolbox manual [9].

In this way, the auralizations are the link that enable us to study the existence of relationships between audio features, as calculated by the MIRToolbox, and the objective acoustic parameters of the room. The acoustic parameters considered in this study are:

(1) reverberation parameters: reverberation time (RT30), and early decay time (EDT)
(2) energy parameters: sound pressure level (SPL), clarity (C80), and center time (Ts)
(3) intelligibility parameters: definition (D50) and speech transmission index (STI)
(4) spatial parameters: early lateral energy fraction (LF) and late lateral sound level (LG)

Other works have also extracted room acoustic parameters from recorded signals. In Ref. [12], Kendrick et al. use music and speech signals to estimate parameters that represent in-use room characteristics. In Ref. [13], the author estimates the reverberation time from signals captured in a room that is in use. In Ref. [14], the authors present a performance comparison of three methods of blindly estimating reverberation time from speech signals. In a recent work we have studied the relationship between the audio features of auralizations and the acoustic parameters obtained from simulations [15]. In the present work we first considerer a factor analysis reduction of MIRToolboox features. From the results correlations are studied between the reduced variables and objetvie parameters.

2. MODELED AND STUDIED ROOMS

Objective parameters were computed from impulse responses at three different points in each room using the ODEON software v.10.1. To calculate a representative value for each parameter, merit figure definitions and frequency averaging were used. The equations for the merit figures can be found in other works from the same authors [16]. The ODEON software was also used for the room acoustic simulations. This software uses a hybrid algorithm in which two geometrical methods are combined to predict the impulse response of a virtual room. The simulation of the room impulse response (RIR) is performed in two steps. The first part, which contains information about early reflections, is calculated by com- bining the image source method and early scattered rays. The du- ration of the early part can be chosen by the user via the so-called transition order (TO). This is the maximum number of image sources taken into account per initial ray [16]. The second part of the RIR (i.e., the late reflections) is calculated by a modified ray-tracing algorithm that also takes into account the scattering coefficient of the surfaces. At every reflection event, local diffuse secondary sources are generated that radiate sound with a directiv- ity that is in accordance with Lambert's cosine-law [17,18]. The detailed algorithm of the ODEON software has been described by the developers [19, 20].

This study required a suitable range of values of the acoustic parameters. Therefore, a group of rooms with different acoustics have been selected. The simulated rooms are:

(1) La Llotja de la Seda (the Silk Exchange). It is a late Valencian Gothic building in Valencia,

Spain. It was built between 1482 and 1548 and is one of the principal tourist attractions in the city. The UNESCO named it a World Heritage Site in 1996. Its RT at mid-frequencies is 3.1 s with a volume of 12,100 m3. We measured 25 points. In Ref. [21], the virtual model is presented in detail.

(2) Principal Theater of Valencia. It is Italian-style theater that is horseshoe shaped with boxes on various floors. It is used for plays, orchestral and soloist concerts, opera, chorus, and dance. It was built in 1832 (refurbished in 1991) and has 1224 seats and a volume of 6986 m3. The RT at mid-frequencies is 1.5 s and the V/seat is 5.7. We measured 53 points. An exhaustive study of this theater and its model is shown in Ref. [22].

(3) Paranimf UPV. It is a rectangular room used for conferen- ces, soloist concerts, chamber orchestras, and choral per- formances. Built in 1978 with 385 seats and a volume of 2700 m3, it has an RT at mid frequencies of 1.3 s and a V/seat of 7. We measured 24 points. In Ref. [23], the room and a detailed model is presented.

(4) Modified Paranimf UPV. It is a modification of the original Paranimf room. It is a very dry room with an RT at mid- frequencies of 0.5 s.

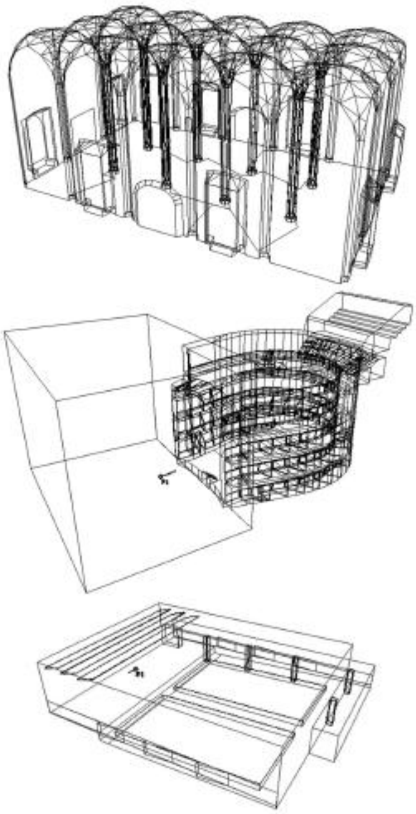

Figure 1. 3D models in ODEON.

2.1 Description of the Acoustic Models.

A spatial geometrical model of the rooms was virtually constructed in the ODEON software (see Fig. 1). The sound absorption and diffusion properties of the surfaces in each room's computer model were cali- brated following an iterative process in which these properties were adjusted until the mean measured reverberation time RT30 and the mean simulated time were as similar as possible. The just noticeable difference (JND) index is used to know how well adjusted the model is. Figure 2 shows the mean measured values together with their range and their cor- responding mean simulated values of RT30. It can be seen that, in all frequencies, the mean simulated values are inside the range.

2.2 Source Ensemble and Receivers.

In order to have a large number of audio files, the source distributions shown in Fig. 3 have been used. Five combinations have been set: d =0, a unique source at the center of the stage; and four length combinations associated with d = .75 m, d = 1.5 m, d = 2.25 m, and d = 3 m and corresponding to half of the edge of a square centered at the position of d = 0 case. This spatial distribution was used in Ref. [24] in order to study the relationships between audio features and the width of the source ensemble. Although the models were calibrated using the results in all measured positions, auralizations have only been calculated in three positions (see Fig. 4).

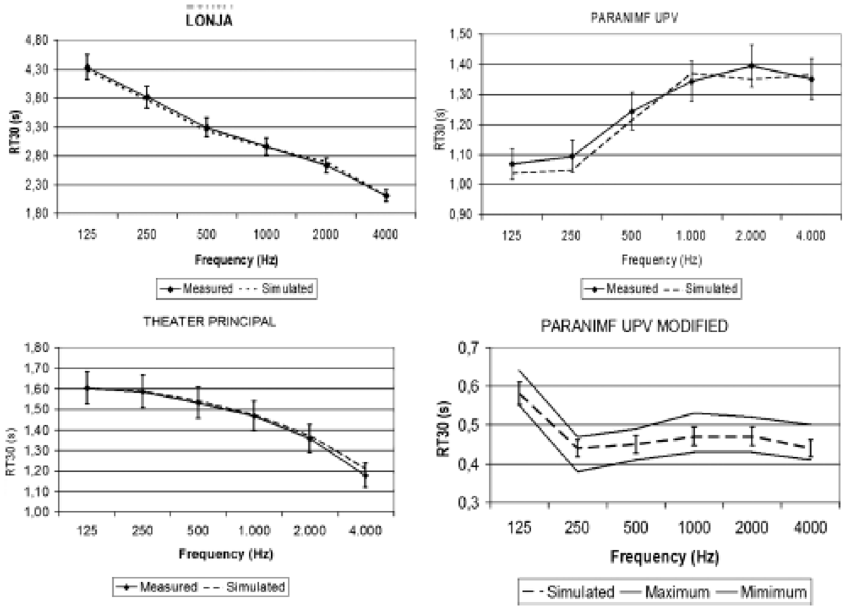

Figure 2. RT30 calibration in each room. In the modified Paranimf, the simulated RT30, together with its range, is shown.

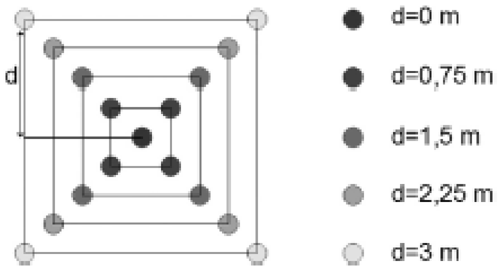

Figure 3. Spatial distribution of sources.

3. AURALIZATION PROCEDURE

3.1 Synthetic Signals as Source Signals.

One of the main difficulties in auralization is obtaining anechoic recordings [1]. As a novelty, we used synthesized signals (using Finale Notepad software) as source signal instead of anechoic recordings. There have been a few attempts to use synthetic signals as source signals in auralizations [25]. To the best of our knowledge, commercial products for music editing have not been used to produce source signals for auralizations. Finale Notepad 2012 (free distribution) has been used, which includes high quality built-in software instrument sounds, to produce separate instrument recordings as source signals.

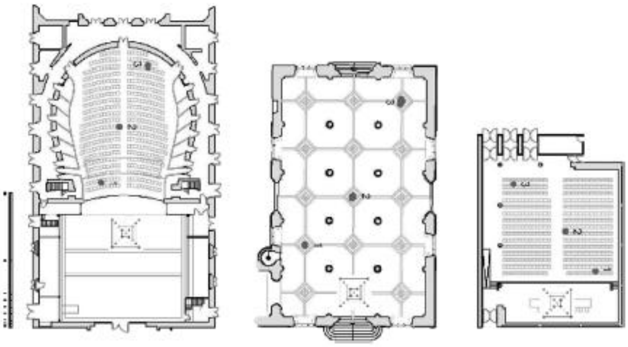

Figure 4. Receivers where auralizations were computed together with source position (represented as shown in Fig. 3). From top to bottom: Paranimf, Lonja, and Principal Theater of Valencia.

When selecting the pieces, the main criteria was the availability of a version for four instruments. These pieces were found in Ref. [26]. In order to have a wide range of instruments we have selected an instrument from each family for each piece: wind, percussion, string, and piano. The pieces for quartets (or reduced to four instruments) were:

(1) Claude Debussy (1862–1918): golliwog's cakewalk, quartet for clarinets
(2) Claude Debussy (1862–1918): the little black, quartet for mixed clarinets; first 60 s
(3) Johann Sebastian Bach (1685–1750): in dulci jubilo choral prelude for wind quintet, bwv608; first 60 s
(4) Alexander Glazunov (1865–1936): oriental reverie for clarinet and string quartet; first 60 s

Auralizations with five different square distributions of sources and an omni directivity pattern were computed for music at three locations in each of the four rooms. 240 auralized signals have been obtained using this procedure.

4. STATISTICAL ANALYSIS: RESULTS

To study the relationships among the features provided by the MIRToolbox and the acoustic parameters in the rooms, SPSS software [27] has been used. Although correlations among the variables can be studied directly, we chose to reduce the data using factor analysis (FA). The application of this method in the statistical reduction of acoustical parameters in rooms [28,29] enables the extraction of the principal components that explain the statistical variance in the data. Once the components have been obtained, the FA provides the weights or loads for each variable in each factor and this facilitates the identification of correlations between variables. As an initial result of the application of the FA, a reduction of 182 statistical parameters of the features provided by the MIRToolbox has been obtained. These parameters do not show variation in the 240 studied cases and, therefore, they cannot be included in the FA [30]. Although this fact supposes that 194 variables remain, this quantity of statistical parameters corresponds to just 24 MIRToolbox features.

A more efficient reduction is made applying factor analysis to these 194 variables in MIRToolboox. We obtained 27 factors. Corresponded values were saved using regression method. From these new variables we have studied correlations with objective room acoustic parameters. Correlations appear with the third and sixth factors. In the first case, Factor 3 correlates with C80, STI, D50, Ts, EDT and RT. In Table 1, the parameters and their correlations can be seen. These parameters were already grouped in previous papers on room parameter reduction [28,29] and they had the same sign in the correlations. This confirms that auralizations maintain the correlations that appear in the measurements. This fact supports the reliability of virtual modeling for rooms.

Room Parameter	Correlation with Factor 3
C80	0.87
STI	0.85
D50	0.85
Ts	-0.76
EDT	-0.78
RT	-0.79

Table 1 Room parameters and correlations with Factor 3

It can be seen that the main MIRToolbox feature that appears in Factor 3 is the spectral_dmfcc_std number. The audio feature in factor 3 is a temporal differentiation of order 1 of the cepstral coefficients, also termed the delta_MFCC (dmfcc). Finally, the statistical value shown is the standard deviation of the dmfcc. we have calculated the mean of the components 7–13, naming it as dmfcc_std_mid; that is, what would correspond to the mid-frequencies in the Mel scale [31]. This result is the same that was obtained in [15]. Since the heaviest component is the acoustical parameter C80avg , we have studied the statistical relationship between C80avg and the dmfcc_std_mid, and by using the average values in each room the following formula has been obtained:

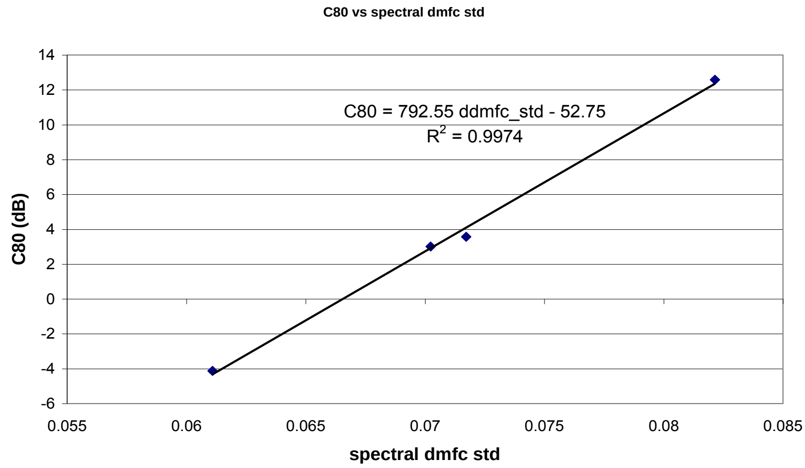

Figure 5. C80 versus dmfcc_std_mid at each room

$$C80avg = -52.75 + 792.55\ \text{dmfcc std mid}\ \ (r=0.9974) \tag{1}$$

The correlation obtained shows that the studied models exhibit a linear relationship between the musical clarity parameter (C80avg) and the standard deviation of the first temporal derivative of the cepstral spectrum in mid-frequencies dmfcc_std_mid.

6. CONCLUSIONS

In this paper, the statistical relationships between the musical features of auralizations and the acoustical parameters of the corresponding impulse responses are determined. The musical characteristics have been calculated by using the MIRToolbox [9]. As these features are a large set of parameters, a reduction procedure by factor analysis was made. The auralizations and the acoustical parameters have been obtained with ODEON [20]. Four rooms have been modeled with different reverberation times. Different source positions and dif- ferent musical motifs have been used and three positions in each room have been auralized. In total, 240 auralizations have been obtained. As a novelty, synthesized, rather than anechoic signals, were used.

Using correlation analysis and factor analysis, results obtained in [15] were corroborated. Note the relation:

C80avg = -53.5 + 804.1 dmfcc std mid (r=0.998),.

6. ACKNOWLEDGMENT

This study was supported with FEDER funds, the Spanish Ministry of Science and Innovation supported the present study in a Coordinated Investigation Project Framework (Grant Nos. BIA2003-09306-C04 and BIA2008-05485).

7. REFERENCES

[1] Vorlander, M., 2008, Auralization: Fundamentals of Acoustics, Modelling, Sim- ulation, Algorithms and Acoustic Virtual Reality, Springer-Verlag, Berlin.

[2] Prince, D., and Talaske, R., 1994, “Variation of Room Acoustic Measurements as a Function of Source Location and Directivity,” W. Clement Sabine Centen- nial Symposium, Cambridge, MA, June 5–7, pp. 211–214.

[3] Martin, R. S., Witew, I., Arana, M., and Vorlander, M., 2007, “Influence of the Source Orientation on the Measurement of Acoustic Parameters,” Acta Acust., 93, pp. 387–397.

[4] Otondo, F. and Rindel, J., 2004, “The Influence of the Directivity of Musical Instruments in a Room,” Acta Acust., 90, pp. 1178–1184.

[5] Dalenback, B., Kleiner, M., and Svensson, P., 1993, “Audibility of Changes in Geometric Shape, Source Directivity, and Absorptive Treatment—Experiments in Auralization,” J. Audio Eng. Soc., 41, pp. 905–913.

[6] Vigeant, M., Wang, L., and Rindel, J., 2011, “Objective and Subjective Evalua- tions of the Multi-Channel Auralization Technique as Applied to Solo Instruments,” Appl. Acoust., 72, pp. 311–323.

[7] Gil-Reyes, B., Jeong, C.-H., and Brunskog, J., 2011, “Room Acoustic Investiga- tion of Actors’ Positions and Orientations for Various Theatre Configurations in a Moderate-Sized Drama Theatre,” Appl. Acoust., 72, pp. 48–58.

[8] Lartillot, O., Toiviainen, P., and Eerola, T., 2008, “A MATLAB Toolbox for Music Information Retrieval,” Data Analysis, Machine Learning and Applica- tions (Studies in Classification, Data Analysis, and Knowledge Organization), C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker, eds., Springer, Berlin/Heidelberg, pp. 261–268.

[9] Lartillot, O., 2011, “MIRtoolbox 1.3.4 software,” Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyaskyla, Finland, https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox

[10] EN ISO 3382-1, 2009, "Acoustics—Measurement of Room Acoustic Parame- ters—Part 1: Performance Spaces," International Organization for Standardiza- tion, Brussels, Belgium.

[11] EN ISO 3382-2, 2008, "Acoustics—Measurement of Room Acoustic Parame- ters—Part 2: Reverberation Time in Ordinary Rooms," International Organiza- tion for Standardization, Brussels, Belgium.

[12] Kendrick, P., Cox, T. J., Li, F. F., Zhang, Y., and Chambers, J. A., 2008, "Monaural Room Acoustic Parameters From Music and Speech," J. Acoust. Soc. Am., 124(1), pp. 278–287.

[13] Vieira, J., 2004, "Automatic Estimation of Reverberation Time," 116th Con- vention of the Audio Engineering Society, Berlin, May 8–11.

[14] Gaubitch, N. D., Loellmann, H. W., Jeub, M., Falk, T. H., Naylor, P. A., Vary, P., and Brookes, M., 2012, "Performance Comparison of Algorithms for Blind Reverberation Time Estimation From Speech," Proceedings of the International Workshop on Acoustic Signal Enhancement (IWAENC 2012), Aachen, Germany, September 4–6.

[15] Cerdá, S; Giménez, A; Montell,R; Barba, A; Lacatis, R; Segura, J and Cibrián, R.M. On the Relations Between Audio Features and Room Acoustic Parameters of Auralizations. J. Vib. Acoust. 135(6), 064501 (Jun 19, 2013)

[16] Cerdá, S., Giménez, A., and Cibrián, R. M., 2012, "An Objective Scheme for Ranking Halls and Obtaining Criteria for Improvements and Design," J. Audio Eng. Soc., 60(6), pp. 419–430.

[17] Zeng, X., Christensen, C., and Rindel, J. H., 2006. "Practical Methods to Define Scattering Coefficients in a Room Acoustics Computer Model." Appl. Acoust., 67, pp. 771–786.

[18] Christensen, C., and Rindel, J., 2005, "A New Scattering Method That Com- bines Roughness and Diffraction Effects," Proceedings of Forum Acusticum, Budapest, Hungary, August 29–September 2.

[19] Rindel, J. H., 2000, "The Use of Computer Modeling in Room Acoustics," J. Vibroeng., 2(4), pp. 219–224.

[20] Christensen, C., 2009, "Odeon Room Acoustics Program, Version 10.1, User Manual," http://www.odeon.dk/pdf/OdeonManual10.pdf

[21] Montell, R., Giménez, A., Segura, J., Planells, A., Barba, A., Cerda, S., Cibrián, R., and Lacatis, R., 2011, "Simulacion acustica de auditorios y edificios patrimoniales. Integracion con entornos de realidad virtual," International Seminar on Virtual Acoustics, Valencia, Spain, November 24–25.

[22] Barba, A. and Giménez, A., 2011, El Teatro Principal de Valencia. Acustica y arquitectura escenica, Teatres de la Generalitat y Universitat Politecnica de Valencia.

[23] Lacatis, R., Cerda, S., Giménez, A., and Romero, J., 2011, "Comparacion de los parametros acusticos obtenidos mediante dos programas de simulacion con modelos geometricos de diferente complejidad de una sala," Revista de Acustica, 42(1–2), pp. 33–42.

[24] Cerda, S., Giménez, A., Cibrián, R., Segura, J., Montell, R., and Barba, A., 2012, "Objective Evaluation of the Width of Source Ensemble in Virtual Halls," Proceedings of the Internoise 2012/ASME NCAD Meeting, New York, August 19–22.

[25] Savioja, L., Huopaniemi, J., Lokki, T., and Va¨a¨na¨nen, R., 1999, "Creating Interactive Virtual Acoustic Environments," J. Audio. Eng. Soc., 47(9), pp. 675–705.

[26] Seely, O., "Oliver's Clarinet Music Page," http://www.csudh.edu/oliver/clar- musi/clarmusi.htm

[27] "SPSS 16.0, Manual, 2007," http://min.webs.upv.es/wp-content/uploads/2011/03/Manual-SPSS16.pdf

[28] Cerda´, S., Giménez, A., Romero, J., Cibria´n, R., and Miralles, J., 2009, "Room Acoustical Parameters: A Factor Analysis Approach," Appl. Acoust., 70(1), pp. 97–109.

[29] Cerdá, S., Giménez, A., Romero, J., and Cibrián, R., 2011, "A Factor Analysis Approach to Determining a Small Number of Parameters for Characterising Halls," Acta Acust. Acust., 97, pp. 441–452.

[30] Gorsuch, R. L., 1983, Factor Analysis, Lawrence Erlbaum Associates, Hillsdale.

[31] Slaney, M., 1998, "Auditory Toolbox, Version 2.0," Interval Research Corporation, Technical Report No. 1998-010.