III. CONCLUSIONS AND RECOMMENDATIONS

The dispersion characteristics of the wave can be readily and efficiently observed by analyzing transient waveforms of an initially impulsive wave using the STFT. The MsSs provide a simple and convenient means for transmitting an initial impulse and detecting the transient waveforms. Since the MsSs require no direct physical contact, they do not obstruct the wave propagation and, thus, are particularly well suited for investigation of wave-propagation properties.

The method used in this experiment can also be applied to investigation of wave dispersion in other structures such as rods and shells with regular and irregular geometries. Additional experimental investigations of wave propagation properties in these other structures are thus recommended to validate theory and numerical calculations. Also, additional sensor and instrumentation development is recommended to further extend the frequency range of the investigation.

ACKNOWLEDGMENTS

The work performed and reported here was supported by the Advisory Committee for Research, Southwest Research Institute.

[1] E. Giebe and E. Blechschmidt, "Experimentelle und Theoretische Untersuchungen uber Dehnungseigenschwingungen von Stäben und Röhnen," Ann. Phys. (Leipzig) **18**, 417–457 (1933).
[2] J. H. Heimann and H. Kolsky, "The Propagation of Elastic Waves in Thin Cylindrical Shells," J. Mech. Phys. Solids **14**, 121–130 (1966).
[3] R. W. Mortimer, J. L. Rose, and P. C. Chou, "Longitudinal Impact of Cylindrical Shells," Exp. Mech. **12**, 25–31 (1972).
[4] W. Mohr and P. Höller, "On Inspection of Thin-Walled Tubes for Transverse and Longitudinal Flaws by Guided Ultrasonic Waves," IEEE Trans. Sonics Ultrason. **SU-23**, 369–374 (1976).
[5] H. Kwun and C. M. Teller, "Detection of Fractured Wires in Steel Cables Using Magnetostrictive Sensors," Mat. Eval. **52**, 503–507 (1994).
[6] H. Kwun and C. M. Teller, "Magnetostrictive Generation and Detection of Longitudinal, Torsional, and Flexural Waves in a Steel Rod," J. Acoust. Soc. Am. **96**, 1202–1204 (1994).
[7] C. H. Hodges, J. Power, and J. Woodhouse, "The Use of the Sonogram in Structural Acoustics and an Application to the Vibrations of Cylindrical Shells," J. Sound Vib. **101**, 203–218 (1985).
[8] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and Quadratic Time-Frequency Signal Representations," IEEE Sig. Proc. Mag. **9(2)**, 21–67 (April 1992).
[9] T. Önsay and A. G. Haddow, "Wavelet Transform Analysis of Transient Wave Propagation in a Dispersive Medium," J. Acoust. Soc. Am. **95**, 1441–1449 (1994).
[10] M. C. Junger and F. J. Rosato, "The Propagation of Elastic Waves in Thin-Walled Cylindrical Shells," J. Acoust. Soc. Am. **26**, 709–713 (1954).
[11] P. M. Naghdi and R. M. Cooper, "Propagation of Elastic Waves in Cylindrical Shells, Including the Effects of Transverse Shear and Rotatory Inertia," J. Acoust. Soc. Am. **28**, 56–63 (1956).
[12] I. Mirsky and G. Hermann, "Axially Symmetric Motions of Thick Cylindrical Shells," J. Appl. Mech. **25**, 97–102 (1958).
[13] J. E. Greenspan, "Axially Symmetric Vibrations of a Thick Cylindrical Shell in an Acoustic Medium," J. Acoust. Soc. Am. **32**, 1017–1025 (1960).

HRTF measurements of a KEMAR [43.66.Pn, 43.66.Qp]

William G. Gardner and Keith D. Martin

The Media Laboratory, Massachusetts Institute of Technology, 20 Ames Street, Room E15-401, Cambridge, Massachusetts 02139

An extensive set of head-related transfer function (HRTF) measurements of a Knowles Electronics Mannequin for Acoustic Research (KEMAR) has recently been completed. The measurements consist of the left and right ear impulse responses from a Realistic Optimus Pro 7 loudspeaker mounted 1.4 m from the KEMAR. Maximum length (ML) pseudorandom binary sequences were used to obtain the impulse responses at a sampling rate of 44.1 kHz. In total, 710 different positions were sampled at elevations from −40 deg to +90 deg. These data are being made available to the research community on the Internet via anonymous FTP and the World Wide Web.

It is generally accepted that the auditory cues for sound localization are embodied in the transformation of sound by the torso, head, and external ear. A head-related transfer function (HRTF) is a frequency response describing the pressure transformation from a specific free field source position to the eardrum. Sets of HRTFs are often needed for the study of sound localization and for the synthesis of spatial cues, but measuring HRTFs using human subjects or dummy-head microphones is a laborious task. This technical note describes a set of HRTF measurements made using a Knowles Electronics Mannequin for Acoustic Research (KEMAR) which is being made available to the research community on the Internet.

Measurements were made using an Apple Macintosh computer equipped with a Digidesign Audiomedia II DSP card, which has 16-bit stereo A/D and D/A converters that operate at a 44.1-kHz sampling rate. One of the audio output channels was sent to an amplifier which drove a Realistic Optimus Pro 7 loudspeaker, a small two-way loudspeaker with a 4-in. woofer and 1-in. tweeter. The KEMAR was Knowles Electronics model DB-4004 and was configured with two neck rings and a torso. The left pinna was the "small" model DB-061 and the right was the "large red" model DB-065. The KEMAR was equipped with Etymotic ER-11 microphones, Etymotic ER-11 preamplifiers, and DB-100 occluded ear simulators with DB-050 ear canal extensions. The outputs of the microphone preamplifiers were connected to the stereo inputs of the Audiomedia card.

From the standpoint of the Audiomedia card, a signal sent to the audio outputs results in a corresponding signal appearing at the audio inputs. Measuring the impulse response of this system yields the impulse response of the combined system consisting of the Audiomedia D/A and A/D converters and anti-alias filters, the amplifier, the speaker, the room in which the measurements are made, and most importantly, the response of the KEMAR with its associated microphones and preamps. Interference due to room reflections can be avoided by ensuring that any reflections occur well after the head response time, which is several milliseconds.

The measurements were made in MIT's anechoic chamber. The KEMAR was mounted upright on a motorized turntable which could be rotated accurately to any azimuth under computer control. The speaker was mounted on a boom stand which enabled accurate positioning of the speaker to any elevation with respect to the KEMAR. Thus the measurements were made one elevation at a time, by setting the speaker to the proper elevation and then rotating the KEMAR to each azimuth. With the KEMAR facing forward toward the speaker (0 deg azimuth), the speaker was positioned such that a normal ray projected from the center of the face of the speaker bisected the interaural axis of the KEMAR at a distance of 1.4 m. It is believed that the speaker was always within 1.5 cm of the desired position, which corresponds to an angular error of ± 0.5 deg.

The impulse responses were obtained using a maximum length (ML) sequence measurement technique.[1,2] The sequence length was 16383 samples, corresponding to a 14-bit generating register. This sequence length was chosen to yield a good signal-to-noise ratio (SNR) without excessive storage requirements or computation time. Because the measurements were performed in an anechoic chamber and the ML sequence was sufficiently long, time aliasing in the impulse responses was not significant. The measured SNR for frontal incidence was 65 dB.

The spherical space around the KEMAR was sampled at elevations from −40 deg (40 deg below the horizontal plane) to +90 deg (directly overhead) in 10-deg increments. At each elevation, a full 360 deg of azimuth was sampled in equal sized increments. The azimuth increment sizes were chosen to maintain approximately 5-deg great-circle increments. Table I shows the number of samples and azimuth increment at each elevation (all angles in degrees). In total, 710 locations were sampled.

TABLE I. Number of measurements and azimuth increment at each elevation. All angles are in degrees.

Elevation	Number of measurements	Azimuth increment
−40	56	6.43
−30	60	6.00
−20	72	5.00
−10	72	5.00
0	72	5.00
10	72	5.00
20	72	5.00
30	60	6.00
40	56	6.43
50	45	8.00
60	36	10.00
70	24	15.00
80	12	30.00
90	1	x

It was desired to obtain HRTFs for both the "small" and "large red" pinna styles. If the KEMAR had perfect medial symmetry, including the pinnae, then the resulting set of HRTF measurements would be symmetric within the limits of measurement accuracy. In other words, the left ear response at azimuth θ would be equal to the right ear response at azimuth 360-θ. It was decided that an efficient way to obtain symmetrical HRTF measurements for both the "small" and "large red" pinnae was to install both pinnae on the KEMAR simultaneously, and measure the entire 360-deg azimuth circle. This yields a complete set of symmetrical responses for each of the two pinna, by associating each measurement at azimuth θ with the corresponding measurement at azimuth 360-θ. For example, to form the symmetrical response pair for the "small" pinna (which was mounted on the left ear), given a source location at 45 deg right azimuth, the left ear response at 45 deg (contralateral response) would be paired with the left ear response at 315 deg azimuth (simulated ipsilateral response). Such a symmetrical set will not exhibit interaural differences for sources in the median plane, which has been shown to be a localization cue.[3] Assuming an HRTF is negligibly affected by the shape of the opposite pinna, these symmetrical sets should be the same as sets obtained using matched pinnae.

Each HRTF measurement yielded a 16 383-point impulse response at a 44.1-kHz sampling rate. Most of these data are irrelevant. The 1.4-m air travel corresponds to approximately 180 samples, and there is an additional delay of 50 samples inherent in the playback/recording system. Consequently, in each impulse response, there is a delay of approximately 230 samples followed by the head response, which persists for several hundred samples and is in turn followed by reflections off objects in the anechoic chamber (such as the KEMAR turntable). In order to reduce the size of the data set without eliminating anything of potential interest, the first 200 samples of each impulse response were discarded and the next 512 samples were saved. Each HRTF response is thus 512 samples long.

The impulse response of the Optimus Pro 7 speaker was measured in the anechoic chamber using a Neumann KMi 84 microphone at a distance of 1.4 m. The measurement technique was exactly the same as the HRTF measurements. The speaker response was used to create an inverse filter, which can equalize the HRTF measurements to compensate for the nonuniform speaker response, particularly at low frequencies. However, this equalization will not compensate for other nondirectional elements of the measurements, including the ear canal resonance and the response difference between the Etymotic and Neumann microphones. To prevent excessive noise amplification, the gain of the inverse filter was limited to +20 dB at very low frequencies.

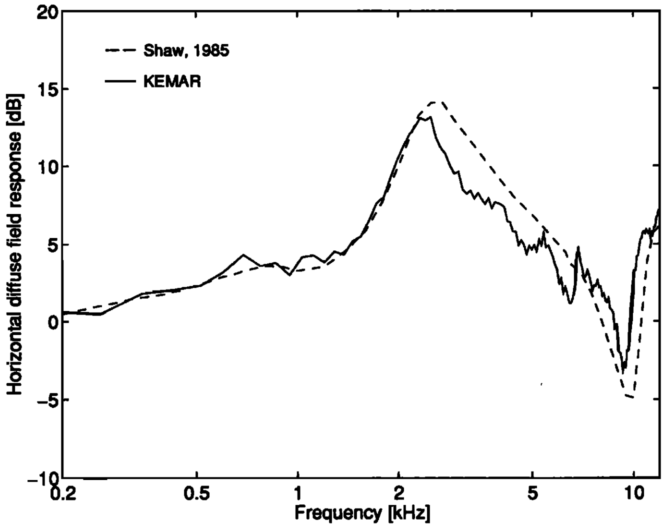

FIG. 1. Comparison of horizontal diffuse field response from Shaw's data and KEMAR data.

Figure 1 shows the average magnitude response of all the HRTF measurements made in the horizontal plane using the "small" pinna, and includes the speaker equalization described above. This planar diffuse field response is compared with that derived from Shaw's data, obtained from HRTF measurements of human subjects.[4] The two curves have the same general shape, exhibiting a peak at 2–3 kHz and a notch near 10 kHz, and are in close agreement for frequencies below 2 kHz.

The HRTF data are available on the Internet via anonymous FTP from the machine "sound.media.mit.edu" (Internet address 18.85.0.105) in the directory "pub/Data/KEMAR". The data are organized into binary archives accompanied by a document that describes in detail the format of the data. The data may also be retrieved via the World Wide Web page "http://sound.media.mit.edu/KEMAR.html". Any correspondence regarding the data may be sent to the authors via the Internet mail address "kemar@media.mit.edu".

ACKNOWLEDGMENTS

This research was supported by the MIT Media Laboratory and National Science Foundation grant GER-9253879. The authors are very grateful to W. M. Rabinowitz for the use of the KEMAR equipment and the anechoic chamber. Thanks are extended to J. G. Desloge and Abhijit Kulkarni.

[1]D. D. Rife and J. Vanderkooy, "Transfer-Function Measurements using Maximum-Length Sequences," J. Audio Eng. Soc. **37**, 419–444 (1989).

[2]J. Vanderkooy, "Aspects of MLS Measuring Systems," J. Audio Eng. Soc. **42**, 219–231 (1994).

[3]C. L. Searle, L. D. Braida, D. R. Cuddy, and M. F. Davis, "Binaural pinna disparity: another auditory localization cue," J. Acoust. Soc. Am. **57**, 448–455 (1975).

[4]E. A. G. Shaw and M. Vaillancourt, "Transformation of sound-pressure level from the free field to the eardrum presented in numerical form," J. Acoust. Soc. Am. **78**, 1120–1123 (1985).