**KEMAR HRTF:** **Full**

270°

180°

90°

0°

Lautsprecher nicht entzerrt, inverse Response Optimus Pro 7 in Datei Opti\_inverse.dat

512 Samples mono, 16 bit Integer

L, normale

Ohrmuschel

R, große

Ohrmuschel

**KEMAR HRTF: Compact**

Linke Ohrmuschel mit Opti\_inverse.dat gefaltet

128 Samples stereo, 16 bit Integer

Signal linker Kanal n=[0°,180°]: l\_full(n)

Signal rechter Kanal n=[0°,180°]: l\_full(n+180°)

Da Kopf 0°/180° Symetrisch:

HRTF linkes Ohr:

n=[0°,180°]: l\_compact(l,n)

n=[180°,360°]: l\_compact(r,n-180°)

HRTF rechtes Ohr:

n=[0°,180°]: l\_compact(r,n)

n=[180°,360°]: l\_compact(l,n-180°)

HRTF Measurements of a KEMAR Dummy-Head Microphone

MIT Media Lab Perceptual Computing - Technical Report #280

Bill Gardner and Keith Martin

MIT Media Lab

May, 1994

abstract

An extensive set of head-related transfer function (In this document,

we use the acronym HRTF to refer to head related impulse responses.

The impulse response and transfer function are related in the obvious

way by the Fourier transform.) measurements of a KEMAR dummy head

microphone 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 meters from the KEMAR. Maximum length (ML)

pseudo-random 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 degrees to +90 degrees.

Also measured were the impulse response of the speaker in free field

and several headphones placed on the KEMAR. This data is being made

available to the research community on the Internet via anonymous FTP

and the World Wide Web.

Measurement technique

Measurements were made using a Macintosh Quadra computer equipped with

an 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. This is a small two way loudspeaker with a

4 inch woofer and 1 inch tweeter. The KEMAR, Knowles Electronics

model DB-4004, was equipped with model DB-061 left pinna, model DB-065

(large red) right pinna, Etymotic ER-11 microphones, and Etymotic

ER-11 preamplifiers. 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. We can avoid interference due to room reflections by

ensuring that any reflections occur well after the head response time,

which is several milliseconds. We can compensate for a non-uniform

speaker response by measuring the speaker response separately and

creating an inverse filter. The inverse filter, when applied to an

HRTF measurement, equalizes the speaker response to be flat.

The impulse responses were obtained using ML sequences (for a detailed

description of the ML sequence measurement technique, see [2]). The

sequence length was N = 16383 samples, corresponding to a 14-bit

generating register. Two copies of the sequence were concatenated to

form a 2\*N sample sound which was played from the Audiomedia card.

Simultaneously, 2\*N samples were recorded on both the left and right

input channels (we wrote software for the Audiomedia to simultaneously

play and record stereo sounds). For each input channel, the following

technique was used to recover the impulse response. The first N

samples of the result were discarded, and the remaining N samples were

duplicated to form a 2\*N sample sequence. This was cross-correlated

with the original N sample ML sequence using FFT based block

convolution, forming a 3\*N - 1 sample result. The N sample impulse

response was extracted starting at N - 1 samples into this result.

Noise in the ML sequence impulse responses can be attributed to

measurement noise, non-linearities in the system, and time aliasing.

Measurement noise can be averaged out by using longer ML sequences.

This is completely analagous to averaging smaller length measurements.

For instance, averaging two independent N point impulse response

measurements should achieve a 3 dB signal to noise ratio (SNR)

improvement over either of the measurements considered alone.

Similarly, using a 2\*N(+1) point ML sequence should achieve a 3 dB SNR

improvement over using an N point ML sequence. However, noise caused

by non-linearities in the system will not be reduced by repeated

averaging over ML sequence measurements, because the noise is

correlated between measurements. It is necessary either to use longer

ML sequences or to average the reponses resulting from different ML

sequences (i.e. from different masks) to reduce noise caused by

non-linearities (see [3]). Time aliasing can be eliminated by using

ML sequences which are longer than the reverberation time of the

measurement space. Since the measurements were done in an anechoic

chamber and the ML sequences were sufficiently long, time aliasing was

not a problem. We chose 16383 point measurements to give good signal

to noise ratios without excessive storage requirements or computation

time. The measured SNR was 65 dB, as discussed later.

Measurement procedure

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 degrees 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 meters. This was accomplished using a tape

measure, plumb line, calculator, a 1.4 meter rod, and a fair amount of

eyeballing. We believe the speaker was always within 0.5 inch of the

desired position, which corresponds to an angular error of +/- 0.5

degrees.

The spherical space around the KEMAR was sampled at elevations from

-40 degrees (40 degrees below the horizontal plane) to +90 degrees

(directly overhead). At each elevation, a full 360 degrees of azimuth

was sampled in equal sized increments. The increment sizes were

chosen to maintain approximately 5 degree great-circle increments.

The table below shows the number of samples and azimuth increment at

each elevation (all angles in degrees). A total of 710 locations were

sampled.

Elevation Number of Azimuth

Measurements 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.xx

Table 1: Number of measurements and azimuth

increment at each elevation

If the KEMAR was perfectly symmetrical and its ear microphones were

identical, we would only need to sample either the left or right

hemisphere around the KEMAR. However, our KEMAR had two different

pinnae (the left pinna was ``normal'', the right pinna was the ``large

red'' model), and consequently the responses were not identical. This

was actually a bonus, because by sampling the entire sphere we

obtained two complete sets of symmetrical HRTFs.

Speaker and headphone measurements

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 meters. The measurement technique was exactly the same as the

HRTF measurements. The speaker impulse response can be used to create

an inverse filter to equalize the HRTF measurements, as will be

discussed later.

In addition to measuring the speaker response, we also measured a

variety of headphones placed on the KEMAR. The headphones measured

are listed in Table 2.

AKG K240 Circumaural, closed earcups, but

not well isolated.

Sennheiser HD480 Supraaural, open air.

Radio Shack Nova 38 Supraaural, walkman style.

Sony Twin Turbo Intraaural, earplug style.

Table 2: Description of headphones measured

It is possible the HRTF data will be used to create a spatial auditory

display, in which case the frequency response of the headphones used

to render the display is important. The above headphone responses may

be useful to create appropriate inverse filters. We did not gather

data on the repeatablitity of such measurements (i.e. how much

variation in the frequency response is expected each time the

headphones are placed on the head).

The data

As described earlier, each HRTF measurement yielded a 16383 point

impulse response at a 44.1 kHz sampling rate. Most of this data is

irrelevant. The 1.4 meter 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 before the

head response occurs. The head response persists for several hundred

samples (subject to interpretation) and is followed by various

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, we decided to discard the

first 200 samples of each impulse response and save the next 512

samples. Each HRTF response is thus 512 samples long. Most

researchers will no doubt truncate this data further.

The impulse responses are stored as 16-bit signed integers, with the

most significant byte stored in the low address (i.e. Motorola 68000

format). The dynamic range of the 16-bit integers (96 dB) exceeds the

signal to noise ratio of the measurements, which we conservatively

measured to be 65 dB. Using the 0 degree elevation, 0 degree azimuth,

left ear, 16383 point measurement, we compared the energy in 100

samples centered on the head response to the first 100 samples of the

response (these should ideally be zero) which yielded the 65 dB SNR.

The HRTF data is stored in directories by elevation. Each directory

name has the format ``elevEE'', where EE is the elevation angle.

Within each directory each filename has the format ``XEEeAAAa.dat''

where X is either ``L'' or ``R'' for left and right ear response,

respectively, EE is the elevation angle of the source in degrees, from

-40 to 90, and AAA is the azimuth of the source in degrees, from 0 to

355. Elevation and azimuth angles indicate the location of the source

relative to the KEMAR, such that elevation 0 azimuth 0 is directly in

front of the KEMAR, elevation 90 is directly above the KEMAR,

elevation 0 azimuth 90 is directly to the right of the KEMAR, etc.

For example, the file ``R-20e270a.dat'' is the right ear response,

with the source 20 degrees below the horizontal plane and 90 degrees

to the left of the head. Note that three digits are always given for

azimuth so that the files appear in sorted order in each directory.

To select a pair of HRTF responses, we recommend using symmetrical

responses obtained from one of the KEMAR ears. For instance, for the

HRTF responses for a source 45 degrees to the right of the head at 0

degrees elevation, use ``L0e045a.dat'' for the left ear and

``L0e315a.dat'' for the right ear, or use ``R0e315a.dat'' for the left

ear and ``R0e045a.dat'' for the right ear. Note that this approach

eliminates binaural localization cues in the median plane.

The maximum sample value in the left ear HRTF data is -26793 in file

``L40e289a.dat''. In the right ear HRTF data the maximum value is

29877 in the file ``R40e039a.dat''.

The speaker impulse response and headphone impulse responses are

stored in the directory ``headphones+spkr''. An inverse filter for

the Optimus Pro 7 speaker is included. The inverse filter was

designed by zero-padding the measured impulse response and taking the

DFT of the zero-padded sequence. The resulting complex spectrum was

inverted by negating the phase and inverting the magnitude. This was

done over the range from DC to 18 kHz; beyond 18 kHz the inverse

spectrum was made flat by repeating the 18 kHz magnitude value. The

inverse filter was obtained by computing the inverse DFT of this

spectrum. A minimum phase version of this inverse filter was also

computed using the real cepstrum (see [1]). The files in the

``headphones+spkr'' directory are listed in Table 3.

filename description

Optimus.dat Optimus Pro 7 impulse response

Opti\_inverse.dat Inverse filter for Optimus Pro 7

Opti\_minphase.dat Minimum phase inverse filter

AKG-K240-L.dat AKG headphone impulse response

AKG-K240-R.dat

Senn-HD480-L.dat Sennheiser headphone impulse response

Senn-HD480-R.dat

RS-Nova38-L.dat Radio Shack headphone impulse response

RS-Nova38-R.dat

Sony-TwinTurbo-L.dat Sony headphone impulse response

Sony-TwinTurbo-R.dat

Table 3: Contents of ``headphones+spkr'' directory

The 512 point impulse responses and speaker and headphone data may be

found in the tar archive ``full.tar.Z''.

Compact data files

For those interested purely in 3-D audio synthesis, we have included a

data-reduced set of 128 point symmetrical HRTFs derived from the left

ear KEMAR responses. These have also been equalized to compensate for

the non-uniform response of the Optimus Pro 7 speaker. The 128 point

responses may be found in the tar archive ``compact.tar.Z''. The

data-reduced impulse responses are stored in directories by elevation

as described above. Within each directory each filename has the

format ``HEEeAAAa.dat'' where EE is the elevation angle of the source

in degrees, and AAA is the azimuth angle of the source in degrees.

Each file contains a stereo pair of 128 point impulse responses

corresponding to the left and right ear responses for the given source

position. For instance, the file ``H0e090a.dat'' contains the left

and right ear impulse responses for a source directly to the right of

the listener. The left response was derived from the 512 point file

``L0e090a.dat'' and the right response was derived from the 512 point

file ``L0e270a.dat''. The data is stored as 16-bit integers and the

stereo samples are stored in (left, right) interleaved order. Each

128 point response was obtained by convolving the appropriate 512

point impulse responses with the minimum phase inverse filter for the

Optimus Pro 7 speaker. The resulting impulse responses were then

cropped by retaining 128 samples starting at sample index 26. The

maximum sample value in the 128 point data is 30496 in the file

``H-10e100a.dat''.

Accessing the data on the Internet

The data is organized into two tar archives, this document (postscript

and plain text) and a text README file. The structure of the tar

archives is described in the previous sections.

To retrieve the HRTF data by anonymous FTP, your FTP session would

look something like the following:

kdm@eno:~ > ftp sound.media.mit.edu

Connected to sound.media.mit.edu.

220 sound.media.mit.edu FTP server (ULTRIX Version 4.1 Tue Mar 19 00:38:17 EST 1991) ready.

Name (sound.media.mit.edu:kdm): anonymous

331 Guest login ok, send ident as password.

Password: {Type your User ID here}

230 Guest login ok, access restrictions apply.

ftp> cd pub

250 CWD command successful.

ftp> cd Data

250 CWD command successful.

ftp> cd KEMAR

250 CWD command successful.

ftp> ls

200 PORT command successful.

150 Opening data connection for /bin/ls (18.85.0.105,3975) (0 bytes).

README

compact.tar.Z

full.tar.Z

hrtfdoc.ps

hrtfdoc.txt

226 Transfer complete.

60 bytes received in 0.42 seconds (0.14 Kbytes/s)

ftp> binary

200 Type set to I.

ftp> get README

200 PORT command successful.

150 Opening data connection for README (18.85.0.105,3806) (417 bytes).

226 Transfer complete.

local: README remote: README

952 bytes received in 0.043 seconds (22 Kbytes/s)

etc.

Please note that there are no files shared between the two tar archive

files. To expand the tar archives, use:

kdm@eno:~ > uncompress full.tar.Z

kdm@eno:~ > tar xvf full.tar

kdm@eno:~ > uncompress compact.tar.Z

kdm@eno:~ > tar xvf compact.tar

This will create the directories ``full'' and ``compact''.

To retrieve the HRTF data via the WWW, use your browser to open the

following URL:

http://sound.media.mit.edu/KEMAR.html

Simply follow the directions found in the html document.

Usage restrictions

This HRTF data is Copyright 1994 by the MIT Media Lab. It is provided

without any usage restrictions. We request that you cite the authors

when using this data for research or commercial applications.

Correspondence

All correspondence regarding this data should be directed to:

Keith Martin Bill Gardner

MIT Media Lab, E15-401D MIT Media Lab, E15-401B

20 Ames Street or 20 Ames Street

Cambridge, MA 02139 Cambridge, MA 02139

kdm@media.mit.edu billg@media.mit.edu

Acknowledgements

The successful completion of this project would not have been possible

without the help and support of W. M. Rabinowitz, J. G. Desloge,

Abhijit Kulkarni, and the MIT Media Lab Machine Listening Group. This

research is supported in part by the MIT Media Laboratory and the

National Science Foundation.

Bibliography

[1] A. V. Oppenheim and R. W. Schafer. Discrete-Time Signal Processing.

Prentice-Hall, Englewood Cliffs, NJ, 1989.

[2] D. D. Rife and J. Vanderkooy. ``Transfer-Function Measurements using

Maximum-Length Sequences''. J. Audio Eng. Soc., 37(6):419-444,

June 1989.

[3] J. Vanderkooy. ``Aspects of MLS Measuring Systems''.

J. Audio Eng. Soc., 42(4):219-231, April 1994.