INSTRUMENTATION & TECHNIQUES

Measurement of head movement during auditory localization

SAMUEL GILMAN, DONALD D. DIRKS, and STEVEN HUNT
University of California, Los Angeles, Los Angeles, California 90024

Head movement can have a significant effect on the ability to locate the direction of a sound source. A system has been designed to track the head movement in response to sound originating at different azimuth locations with respect to the head. A videotape record is made of a light approximating a point source carried on a lightweight "beanie" mounted on the listener's head. Movement of the light is monitored by the video camera and recorded on tape, along with the sound stimulus and information concerning loudspeaker location and time. The horizontal and vertical coordinates of the light-spot image are determined in relation to the video synch pulses defining the field borders. Synch signals are available from a video monitor either in real-time or from tape replay to define each TV frame and horizontal scan line. The circuitry interfaces to a computer programmed to take the information, apply a calibration, and process the data into records of time-varying head position and velocity. Examples of both digital and graphic printouts of head movement are given. The system is capable of expansion to three-axis operation.

Head movements can supply information about acoustic fields that is helpful to a listener in localizing the position of a stationary sound source. Wallach (1940) provided a detailed and comprehensive description of the head movements that occur during sound localization. However, the complexity of such movements make them difficult to describe quantitatively. As a result, most sound-field localization experiments have been conducted with the head physically restrained, or with movements of the head forced by drive motor systems connected to lever arms (Pollack & Rose, 1967; Thurlow & Runge, 1967). Thurlow, Mangels, and Runge (1967) attempted to measure head movement during sound localization by photographing head movement with a 16-mm motion picture camera as a subject localized each of 10 sound sources located at several horizontal and vertical positions. In order to facilitate measurements of angular movement, the subject wore a circular lightweight frame on his head. The film was viewed on a standard microfilm reader, and from this image angular changes in head position were determined. Head movements from 23 subjects were described in terms of the patterns and magnitude of the maximal movements in rotation (horizontal direction), tip (vertical direction), and pivot (an arc to right or left) (see Figure 1). Thurlow et al. limited their analysis to general summaries, due apparently to the complexity of the head movements and the large amount of data obtained for each test condition.

Our interest in auditory localization developed from previous investigations (Dirks, Stream, & Wilson, 1972; Dirks & Wilson, 1969a, 1969b; Stream & Dirks, 1974)

of speech perception in relation to the sound field. In these experiments, changes in speech discrimination were measured for binaural and monaural listening conditions as a function of the azimuthal position of the speaker. These investigations were carried out with the head immobilized. To simulate the more realistic conditions where head movement is unrestrained, we conducted experiments with free head-movement capability. To facilitate these measurements, we have developed an optical method for monitoring head movement combined with a video system for recording and preparing the resulting signals for computer processing. This report details the technical description

Figure 1. Head and neck movements and their limits.

of the head-movement monitoring system, the preparation of the video signal for computer processing, and several examples of the resulting computer outputs.

SYSTEM REQUIREMENTS

Ultimately, magnitude and direction of motion of the human head in response to a stimulus is limited by the physiological capabilities of the head-to-body joints. Average limits for humans have been reported (Batch, 1955; Van Cott & Kinkade, 1972).

Typically, the human head can move rotationally in one or any combination of three motions, as shown in Figure 1. (1) *Neck Rotation*: This is motion to the right or left on a horizontal plane about a vertical axis through the cervical vertebrae and takes place primarily as an articulation between the first and second vertebrae. (2) *Neck Flexion (Tipping)*: This is motion forward (ventral) and backward (dorsal). The motion is the sum of the movements in articulation between the skull and all of the cervical vertebrae in the sagittal plane. No axis is defined for this motion, since the axis changes with head movement. (3) *Lateral Bending (Pivot)*: This motion results from articulation of all of the cervical vertebrae in the coronal plane. As in flexion, no axis is defined.

Thurlow et al. (1967) observed that most head movement is rotational regardless of loudspeaker elevation. For our studies, the loudspeakers are located on a horizontal plane; this results only in azimuth sound directionality. Thus, measurement of head movement in only the horizontal plane is sufficient for the initial investigations. Although the data indicated the accuracy of this assumption, the system is designed so that measurement capability for movement in other planes can be added.

Based on the described head-movement capabilities, the system is designed to monitor horizontal rotational movement of ± 70 deg with a resolution capability of 2 deg. The head-movement indicator is constructed to have a minimal effect on the acceleration and mobility of the subject's head and to avoid acoustic reflections or interference. A video-optical method was used because it appeared to have a number of advantages over other systems. The system components can be obtained as commercial equipment at comparatively low cost; the results can be easily monitored while recording; the complete sequence of subject responses can be reconstituted (not just the data); and the record is permanent and instantly usable. The data can be processed on-line or from the videotape and can be stored indefinitely in its original form. Similar systems have been reported for monitoring lip and jaw movement during speech (McCutcheon, Fletcher, & Hasegawa, 1977).

SYSTEM DESCRIPTION

The complete system block diagram is shown in Figure 2 and is grouped into four major subsystems.

Figure 2. Head-movement measuring system and subsystem block diagram.

Subsystems

Stimulus generator and loudspeaker selector. A Grason-Stadler Series 1200 logic and stimulus program generator delivered the stimulus signal randomly to each of 11 loudspeakers. Each subject was exposed to the same set of sequences the same number of times for each complete test program. The stimulus consisted of high- and low-frequency bands of noise, delivered as a train of 10 pulses. Each pulse was 200 msec in duration separated by an off-time of 200 msec. The train of pulses was followed by a subject-response period of 7 sec, during which time the subject verbally indicated which speaker was stimulated. Each series of pulses and the response period required 10.8 sec.

Test chamber and loudspeaker. Figure 3a shows a subject and the arrangement of the equipment in the test room. The test chamber was a modified double-wall audiometric booth with internal dimensions of 8 ft x 10 ft x 6 ft 6 in. Additional sound-absorption material in the form of 4-in. mats of Type PF105 sound-absorbing fiberglass, enclosed in porous cloth, was applied to all room surfaces. At a distance of 1 m from the loudspeaker, the environment approximated a free field for frequencies above 500 Hz and was usable with some compromise down to 200 Hz.

The 11 loudspeakers were small (6 x 8 x 10-1/2 in.), acoustic suspension, two-element units selected for their uniform output from about 100 Hz to 12 kHz. Loudspeakers were located on 15-deg centers on a horizontal circular arc with 1-m radius, aligned so that the center lines of all the speakers were on the same level.

Optical signal generation. Experience with other applications for optical measurement indicated the need for a high-intensity point source of light to obtain an adequate signal-to-noise ratio for video recording. A small, lightweight skullcap ("beanie") made of thin nylon sheet molded to approximate the contour of the top of a human head was used as the mount for the lamp. Four narrow Velcro straps were used to hold the "beanie" to a padded headstrap firmly fastened around the subject's head, as shown in Figure 3b.

Figure 3. Head-movement measuring system equipment and arrangements: (a) testing room arrangement, (b) close-up of subject during test showing position of "beanie" and monitoring lamp, (c) monitoring and recording system (the monitor displays actual image), (d) monitor image for signal processing.

The light source was a Model 30-30 Pinlite incandescent filament lamp whose overall (glass) dimensions were 1/8 in. long and 1/32 in. in diameter. The lamp was mounted on an adjustable bracket supported by an arm that could be moved in a line along the fore-and-aft axis of the subject's head. A mark on this arm was located at 120 mm from the plane of the lamp. The assembly, weighing 98 g is shown in Figure 4.

The subject was seated so that the head's center of rotation was at the center of the 1-m radius formed by the speakers (Figure 3b). The "beanie" was fastened on the head with the mark on the arm at the center of rotation of the head, so that the lamp was 120 mm from the center of rotation of the subject's head. Light intensity from the lamp could be adjusted from outside the chamber.

The movement of the point of light was monitored and recorded using the videotape system shown in Figure 3c, consisting of a Sony Model AVC-3260 Vidicon camera, a Sony Model AV-3650 videocorder, and a Sony Model CVM-194 video monitor. The camera was located outside the test chamber close to the window and was aligned horizontally and vertically with the lamp and lamp support in line with the subject's

head and the zero-azimuth loudspeaker. The size and location of the total image was adjusted so that a scale at the plane of the lamp (not shown) occupied a preset position and length on the monitor screen. The lamp image was then focused for minimum size. Time was recorded from an electronic counter readout, and the loudspeaker stimulated was indicated by a bank of lights. Figure 3c shows a photograph of the normal image as seen on the monitor. The spot image of the light source produced by adjusting the monitor control for coordinate determination is shown in Figure 3d.

Video signal processing. The video measurement

Figure 4. Assembly showing adjustable lamp mount on the "beanie."

system is composed of two sections. One section, an analog circuit, produces reference transistor-transistor logic (TTL) compatible pulses corresponding to the picture synchronization signals and to any intensity peaks in the video picture signal. The other section, a logic circuit, uses the pulses to compute the "x" and "y" coordinates of the intensity peak and provide them as outputs in a digital format. The system is designed to work with a standard 525-line, twin-interlaced closed-circuit television system and to interface with a Digital Equipment Corporation (DEC) PDP-11/15 minicomputer.

Three reference pulses are generated from the composite video picture signal by the analog circuitry: one corresponding to the initiation of a vertical scan, one corresponding to the initiation of a horizontal scan line, and one corresponding to any positive-going element of the composite video picture waveform that exceeds a preset threshold voltage level. The block diagram is presented in Figure 5.

The digitizing circuitry determines the horizontal position of the spot by counting the time delay between detection of the video peak corresponding to the spot and the subsequent occurrence of a horizontal line-synchronization pulse. A 10-mHz clock is gated "on" by a video peak-indicating pulse and gated "off" by the next horizontal line-synch pulse. The clock gating circuitry cannot be activated more than once during a field, giving a sampling rate of 60 samples/sec. In addition to restricting the clock gating to once every field, the horizontal and vertical areas of the field are limited to prevent any spurious video peaks from triggering the sampling circuit. The vertical position of the spot is determined by counting the number of horizontal scan lines within the field that are traversed prior to reaching the one containing the video peak.

The clock output is fed to a software-controlled 16-bit counter within the PDP-11/15 computer. The vertical line count is recorded by an onboard 7-bit counter and transmitted to the computer via a general-purpose interface. A software routine is used to direct storage of the horizontal and vertical position counts in memory and, subsequently, to route them to a disk file. After the data has been collected, the files are available for use by a set of analysis routines.

Figure 5. Signal and data processing block diagram.

DATA PROCESSING AND SOFTWARE

Software support for the video measurement system consists of a set of chainable, disk-resident routines that can collect, analyze, and display the head-movement trajectories. The programs are designed to work under a DEC RT11-2B operating system with a minimum of 24K words of memory.

One routine (VIDIN) is responsible for storing the x and y position coordinates derived from the video digitizing system and saving them in a disk file. It can also simultaneously digitize one analog channel, such as the audio channel of the videotape recorder. Data can be collected either on-line (live) or off-line from the videotape recorder. In the off-line mode, the program uses information recorded on the audio channel to synchronize the data collection process.

After the data is stored on a disk, the routine PLOTD may be used to plot the head-position trajectory vs. time on a cathode-ray tube display. The plot routine includes the capability of simultaneously plotting a number of data segments representing head-movement records of one or several subjects. These segments can be independently scaled and/or overlayed for comparison purposes. Another routine, PRINTD, may be used to dump data records to the terminal. It is also capable of simultaneously printing different data records to facilitate comparisons. In addition to these main routines, a variety of subroutines are available for performing algebraic and logarithmic transformations of the data, as well as interactive editing of data records.

SYSTEM CALIBRATION AND ACCURACY

The light source travels along the arc of a 120-mm-radius circle while the video camera image is on a flat field. As a result, the recorded movement of the light image is $120 \sin \theta$ mm, where θ is the angle from the nominal position. A manual calibration was made to compensate for the curvature of the monitor screen as well as the sine of the angle.

Angular discrimination of the system is primarily a function of the field-sampling rate (60 fields/sec). Variations in head movement that occur during one field interval, including cyclical variations, cannot be resolved. The 2-deg resolution requirement set a maximum ripple of 2 deg in 1/60 sec, corresponding to a maximum average movement of 4 deg in 1/60 sec. Thus, the maximum average head velocity during one field interval that can be recorded is 240 deg/sec. Other factors, such as Vidicon light decay time, pulse rise and decay times, and so on, are very small compared to the sampling time.

The size of the light spot can also affect accuracy of the monitoring system. Since the left edge of the spot is first detected by the horizontal scan, the spot location is determined by the edge rather than the center. Spot location is not corrected for spot size, so that minimum

spot size is desirable for maximum accuracy. Under test conditions, the spot image size was about .1 in. in diameter.

EXAMPLES OF HEAD-MOVEMENT RESPONSES

Preliminary results are shown in Figure 6. The results were selected to illustrate representative responses for binaural (A_1 and A_2) and monaural (B_1 , B_2 , and C_1 , C_2) listening conditions. The sampled data provide information on head movement from which position, velocity, and acceleration can be obtained. Each dot represents an interval of 1/60 sec, and variation of dot spacing gives immediate comparative rate and position information. The filled triangles indicate the location of the activated loudspeaker and the solid arrowheads identify the loudspeaker verbally selected by the subject.

The binaural listener (A_1 and A_2) apparently requires a minimum amount of head movement to arrive at a correct localization judgment. In comparison, the monaural listeners (B_1 , B_2 , and C_1 , C_2) exhibit substantially more erratic head movement for longer periods of time, indicating a much higher level of uncertainty. Despite the several reversals in head movement, the loudspeaker selected by the monaural listener is usually toward the side of the open ear, regardless of the actual position of the activated loudspeaker.

Some subjects rotate their heads very quickly, particularly when one ear is occluded. Results C_1 and C_2 illustrate such an effect and the degree to which the monitor system can follow, record, and analyze these movements. The maximum velocity, as shown in C_2 , in the interval between 2.5 and 3.5 sec is nearly 300 deg/sec. This fast rate of movement indicates, among other effects, that the "beanie" does not significantly limit mobility.

Data for horizontal head movement from groups of binaural and monaural listeners have been recorded and processed using the described technique. The results are currently being analyzed and the examples in the figure represent only a few of the major head-movement characteristics. The procedure appears to be adequate for single-axis movement. Provisions for recording and analyzing two- and three-axis movement has been anticipated in the system design but has not as yet been incorporated.

REFERENCES

- BATCH, J. W. Measurements and recording of joint function. *U. S. Armed Forces Medical Journal*, 1955, **6**, 359-382.
- DIRKS, D., STREAM, R., & WILSON, R. Speech audiometry: Earphone and sound field. *Journal of Speech and Hearing Disabilities*, 1972, **37**, 162-176.
- DIRKS, D., & WILSON, R. Binaural hearing of speech for aided and unaided conditions. *Journal of Speech and Hearing Research*, 1969, **12**, 650-664. (a)
- DIRKS, D., & WILSON, R. The effect of spatially separated sound sources on speech intelligibility. *Journal of Speech and Hearing Research*, 1969, **12**, 5-38. (b)
- MCCUTCHEON, M. J., FLETCHER, S. G., & HASEGAWA, A. Video scanning system for measurement of lip and jaw motion. *Journal of the Acoustical Society of America*, 1977, **61**, 1051-1055.
- POLLACK, I., & ROSE, M. Effect of head movement on the localization of sounds in the equatorial plane. *Perception & Psychophysics*, 1967, **2**, 591-596.
- STREAM, R., & DIRKS, D. Effects of loudspeaker position on differences between earphone and free field thresholds. *Journal of Speech and Hearing Research*, 1974, **17**, 549-568.
- THURLOW, W. R., MANGELS, J. W., & RUNGE, P. S. Head movement during sound localization. *Journal of the Acoustical Society of America*, 1967, **2**, 489-493.
- THURLOW, W. R., & RUNGE, P. S. Effect of induced head movements on localization of direction of sounds. *Journal of the Acoustical Society of America*, 1967, **42**, 480-488.
- VAN COTT, H. P., & KINKADE, R. G. *Human engineering guide to equipment design*. Washington, D.C.: U. S. Government Printing Office, 1972. P. 544.
- WALLACH, H. The role of head movements and vestibular and visual cues in sound localization. *Journal of Experimental Psychology*, 1940, **27**, 339-346.

Figure 6. Graphic printout of head-movement data after processing. A_1 and A_2 show representative responses from binaural listeners to extreme-azimuth stimuli. B_1 and B_2 show representative responses from monaural listeners with the left ear open to similar extreme-azimuth stimuli. C_1 and C_2 are similar responses from monaural listeners with the right ear open.

(Received for publication June 7, 1978;
revision accepted August 15, 1978.)