Human Spatial Cognition Using Visual and Auditory Stimulation

Mi Yu[1], Yong Jun Piao[1], Yong-Yook Kim[2], Tae-Kyu Kwon[3], Chul-Un Hong[3] and Nam-Gyun Kim[3#]

1 Department of Biomedical Engineering, Graduate School, Chonbuk National University, South Korea
2 Center for Healthcare Technology Development, Chonbuk National University, South Korea
3 Division of Bionics and Bioinformatics, Chonbuk National University, South Korea
Corresponding Author / E-mail: ngkim@chonbuk.ac.kr, TEL: +82-63-270-2246, FAX: +82-63-270-2247

KEYWORDS : Spatial Cognition, Auditory Stimulation, Visual Stimulation, Frequency, Pinna, LabVIEW

This paper deals with human spatial cognition using visual and auditory stimulation. More specially, this investigation is to observe the relationship between the head and the eye motor system for the localization of visual target direction in space and to try to describe what is the role of right-side versus left-side pinna. In the experiment of visual stimulation, nineteen red LEDs (Luminescent Diodes, Brightness: 210 cd/m²) arrayed in the horizontal plane of the surrounding panel are used. Here the LEDs are located 10 degrees apart from each other. Physiological parameters such as EOG (Electro-Oculography), head movement, and their synergic control are measured by BIOPAC system and 3SPACE FASTRAK. In the experiment of auditory stimulation, one side of the pinna function was distorted intentionally by inserting a short tube in the ear canal. The localization error caused by right and left side pinna distortion was investigated as well. Since a laser pointer showed much less error (0.5%) in localizing target position than FASTRAK (30%) that has been generally used, a laser pointer was used for the pointing task. It was found that harmonic components were not essential for auditory target localization. However, non-harmonic nearby frequency components was found to be more important in localizing the target direction of sound. We have found that the right pinna carries out one of the most important functions in localizing target direction and pure tone with only one frequency component is confusing to be localized. It was also found that the latency time is shorter in self moved tracking (SMT) than eye alone tracking (EAT) and eye hand tracking (EHT). These results can be used in further study on the characterization of human spatial cognition.

Manuscript received: April 22, 2005 / Accepted: October 26, 2005

1. Introduction

Human spatial cognition not only depends on the properties of actual stimulation but also on the frequency of sound as well as the morphology of sensory organ in which stimuli are embedded. Various conditions that influence perceptual judgments triggered by a preceding or concurrently presented stimuli have been demonstrated.

During the past 25 years, behavioral, clinical, and physiological observations have demonstrated the influence exerted by the arm motor system on the oculomotor system during target tracking tasks. When a human, or a trained monkey, tracks with his eyes a visual target moved by the observer's arm (this condition will be called self-moved target tracking), the smooth pursuit (SP) system performance is enhanced, as compared with the condition in which the observer tracks with his eyes an externally driven visual target (eye-alone tracking). This was evidenced at the end of the 1960s by Steinbach[1], who showed that a subject could track a visual target more accurately if the target was attached to the observer's hand. Angel and Garland[2] attributed this enhanced performance to information transfer between the arm motor system and the oculomotor system. Gauthier and Hofferer[3] showed that the interaction between the moving hand and the eyes was preserved even in total darkness: when the subject moved his or her finger and was instructed to track his non-viewed fingertip with his eyes, SP could be produced without vision, though, with lower gain than in the experiments where the finger was visible. On the contrary, no SP movements could be produced during the tracking of either imaginary or acoustic targets.[4]

The major purpose of this study was to examine spatial cognition by visual stimulation and the mechanism of how combinative information from eyes, head, and arms settle in the brain. Human gaze saccades are investigated in two dimensions (2-D), in order to obtain more insight into the signals controlling the eye and the head motor systems. In addition, eye head coordination was investigated under conditions in which the eye and head were not initially aligned.

The so-called ventriloquism effect in auditory spatial perception is a well-known phenomenon on which the perceived position of an auditory stimulus appears to be shifted in the direction of a simultaneous, spatially disparate visual stimulus.[5]

The external ear (pinna) modifies the frequency spectrum of sounds in a fashion that depends on the direction of the sound source. This filtering action of the pinna produces spectral cues for source location, which are most important in determining the elevation of the source and its front-back position relevant to the human observer. Auditory space is also constituted by the sound of wind, a voice, and the roar of a car heard from back.[6,7]

Accurate spatial perception across a wide frequency range is a prerequisite for music and speech perception, for instance, in melodic, harmonic, and prosodic processing. In all natural pitched sounds, the

Table 1 Comparison of measured roughness data

Sound source type	
Pure tone :	1000Hz sine wave
Harmonic tone :	
	1000Hz + 2000Hz
	1000Hz + 2000Hz +4000Hz
	1000Hz + 2000Hz +4000Hz + 8000Hz
Non-Harmonic tone :	
	1000Hz + 1100Hz
	1000Hz + 1100Hz +1200Hz
	1000Hz + 1100Hz +1200Hz + 1300Hz
White noise	

sound spectrum consists of a time-varying pattern of multiple harmonic partials across a large frequency spectrum. However, the overwhelming majority of the experiments on pitch perception in psychoacoustics and auditory neuroscience used sinusoidal tones consisting of one harmonic partial (fundamental) only. For instance, Wier et al.[8] compared the accuracy in the frequency discrimination sinusoidal tones with eight different frequency ranges from 200 to 8000 Hz by presenting sound with the duration of 500ms in one-second interval against the background of a low-level broadband noise. More recently, Sek and Moore[9] employed three different psychoacoustic methods (difference limen for single tones, in paired sounds, and for frequency-modulation) at six different frequencies ranging from 250 to 8000 Hz. Both experiments showed that frequency discrimination is most accurate up to 2000 Hz, with the accuracy deteriorating thereafter at a rate depending on the method used. Kishon-Rabin et al.[10] compared the frequency discrimination of 300ms of sinusoidal tones at three different frequencies 250, 1000, and 1500 Hz with each other by using 2 and 3-interval forced-choice methods. They found that the higher the frequency was, the more accurately subjects detected the frequency differences.

In this paper, we studied the relationship between the sound color and the accuracy in the cognition of target localization and found the exact role of pinna in localizing target direction.

2. Methods

2.1 Visual Stimulation

Three conditions of gaze arm tracking coordination were considered: eye-alone tracking (EAT), where the subject follows an external target moving in a sinusoidal (peak-to-peak amplitude 20°, frequency: 0.1~2.0Hz) manner with the eyes only; eye and hand tracking (EHT), where the subject follows the external target moving in a similar way as in the case of EAT with both the eyes and the arm; and self-moved target tracking (SMT), where the subject moves the target with his hand in a sinusoidal wave at a trained amplitude, frequency and tracks the self-moved target with his or her eyes.

2.2 Auditory Stimulation

Computer generated sound sources, which include various sound color types, were used for the experiment. Each sound wave had 1 sec of duration with a 3 sec interval and was delivered from a random position by LabVIEW software program. The sound sources were classified into four groups as shown in Table 1. The first group had a pure sine wave with the frequency of 1000Hz. The second group contained harmonics waves that were generated by summing the components of sine waves. The third group had non-harmonic waves that were generated by summing sine waves with nearby frequencies of sine waves. Thus, the third group does not contain any harmonic components at all. The fourth one was white noise that contains all the possible frequencies. Sound pressure level was 57dB as measured at the head position of the subject.

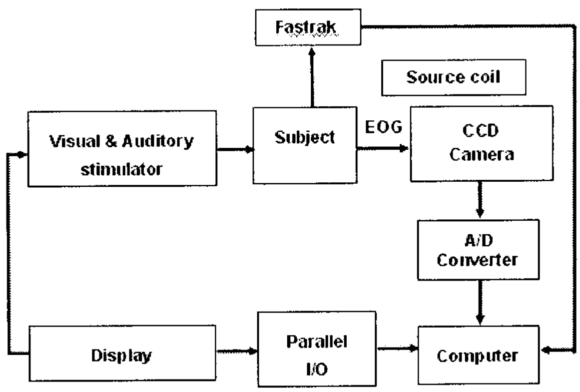

Fig. 1 Block diagram of stimulation system

2.3 Experimental Instrumentation

Fig. 1 shows the block diagram of the experimental system to investigate human spatial cognition using visual and auditory stimulation. In investigating of human spatial cognition by visual stimulation, the measuring units for eyes, head and arms were composed of two devices. One was an EOG. The other was 3SPACE FASTRAK. The EOG passed through a polygraph was sent to a computer with an A/D converter. The head and arm movements were measured by 3SPACE FASTRAK that used an electromagnetic search-coil technique. The FASTRAK calculated three-dimensional position and orientation from the relative positions of a source coil and a sensor coil. The source coil was fixed on a wooden pole and a sensor coil was placed on the head of a subject. The calculated data were sent to a computer by using RS-232C. In investigation of the human's spatial cognition by auditory stimulation, a laser pointer(error 0.5%) was used for the pointing task since it showed much less error in localizing target position especially for the left side angle of the subject position than FASTRAK(error 30%).

The experiment was performed in a soundproof chamber as shown in Fig. 2. For the localization test, 19 numbered loud speakers were arrayed in the horizontal plane of 1.15m radius and placed 10 degree apart from each other. In this experiment, one side of the pinna function was distorted intentionally by inserting a short tube in one of the ear canal. The localization error caused by the right and the left side pinna distortion was investigated.

The experimental data were obtained from ten subjects who were four women and six men with normal visual, head and arm motor function. Their ages ranged from 24 to 34 years. All the subjects had clinically normal hearing in both ears as determined by a pure-tone audiogram. Subjects were asked to participate in these experiments with full knowledge that they could withdraw at any time and singed a consent form.

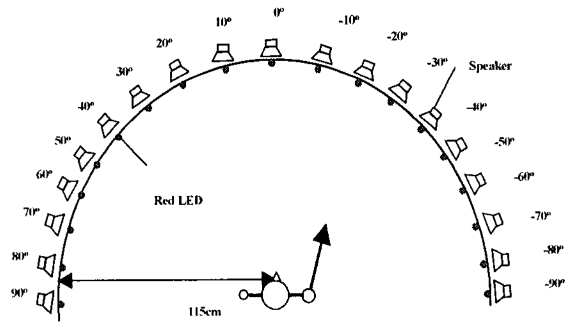

Fig. 2 Surrounding panel for auditory stimulation

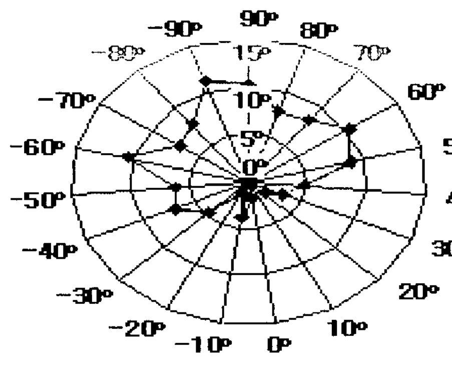

Fig. 3 Angle of error by the location of LED

3. Results

3.1 Visual Stimulation-Gaze arm tracking coordination

Fig. 3 is the angle of error by the location of LED during the experiment of gaze arm tracking coordination. The above arm motor system model (Fig. 1) cannot exhibit a known property of the human arm tracking system, which seems to behave like a servo-controller. In fact, the tracking of a slowly moving target was performed with a sequence of single but fast movements instead of a continuous movement. This intermittency is attributed to visual correction of arm trajectory.[11-12] The angle of error by the location of LED in the center is lower than the periphery.

Fig. 4 illustrates the mean latency time in the three conditions. We measured SP, arm and head latency in EAT, EHT and SMT. SP means the latency time from stimulation to the start of eye movements during the pursuit of the target. Saccade is the latency time during feedback to focus on the target after SP.

The quantitative evaluation of the model's performance was based on harmonic analysis.[13] To calculate the SP, saccades are removed from eye movement first. Saccades were detected using an acceleration threshold that was removed from the eye velocity curve and was replaced by the average velocity of the SP component just before and after the saccades.

In EHT, the mean latency time of SP, Saccade, Arm and Head movement is shorter than that in EAT. In SMT, the latency time is shorter than in EHT. Most of the previous work on gaze arm tracking coordination is based on the analysis of SP latency changes in different experimental conditions. With the dynamic blocks and the internal delays set as above, the simulations showed that the response of the SP system to an external visual target (condition EAT) in an onset latency of 160 ms, while in the SMT condition, the eye-to-arm latency was -5 ms. These values may be compared to the average latencies observed in previous studies (for instance, Vercher et al.[14] showed latencies of 150± 29 ms and -5± 35 ms in the EAT and SMT conditions, respectively). This result stems from the contribution of the arm motor command to the eye movement control via CCS (Fig.2: Coordination Control System). It

Table 2 Average RMS cognition errors for various sound sources

Sound source	Average RMS cognition error (degree)
Pure tone	10.49 ± 1.05
Non-Harmonic tone	8.54 ± 0.76
Harmonic tone	10.45 ± 0.47
White noise	11.10 ± 1.08

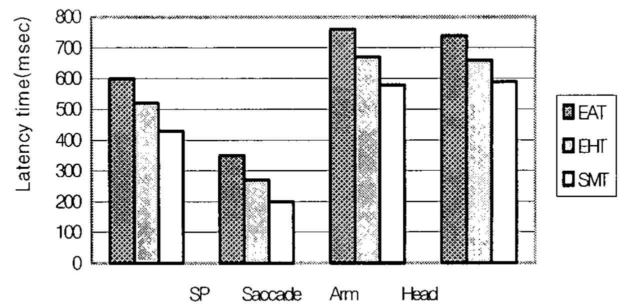

Fig. 4 Latency time of gaze arm tracking coordination

is worth recalling that SMT condition improves the onset of the eye motor response by reducing its latency. In fact, when the self-moved target stops, the eye movement shows the same pattern as that during EAT tracking and one or two resetting saccades are observed before the pursuit velocity declines to zero. This may be viewed as supporting the hypothesis that eye-arm interaction has no effect on SP prediction.

3.2 Auditory Stimulation

3.2.1 Harmonic vs. Non-harmonic

Table 2 shows the average RMS cognition errors for various sound sources. Fig. 5 shows pointing errors (root mean square error, RMS error) in locating target direction that was measured for the groups of various sound sources. It shows quite a large dependency on sound types. The RMS pointing error for the pure tone of 1000 Hz sine wave appeared to be large error. The pure tone and white noise were not very good sound sources in localizing the direction of targets.

The harmonic group shows better results and non-harmonic group shows much better results. Sound source type A (1000Hz+1100Hz) showed the best value, which shows a less than 8 degrees of error.

In general, the function of cochlear is a harmonic analyzer [15]. However, the harmonic components were not helpful in localize sound direction. A pure sine wave was even more confusing. This result reveals that cochlea-brain interaction to localizing target direction has an independent path from harmonic recognition. The directional filtering action of pinna for non-harmonic nearby frequency sound source was more important in localization of auditory target.

3.2.2 Right pinna vs. Left pinna

Fig. 6 shows the effect of the pinna distortion. The localization error of the distorted right pinna was higher than that of the distorted left pinna. However, left pinna distortion did not show large error compared to normal condition with no pinna distortion.

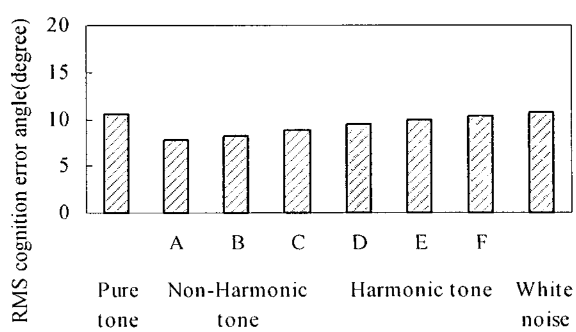

Fig. 5 Cognition errors for various sound source type

It also showed that the error did not depend on the type of sound source very much. Thus, we reached the conclusion that the right pinna performs one of the most important function in localizing target direction.

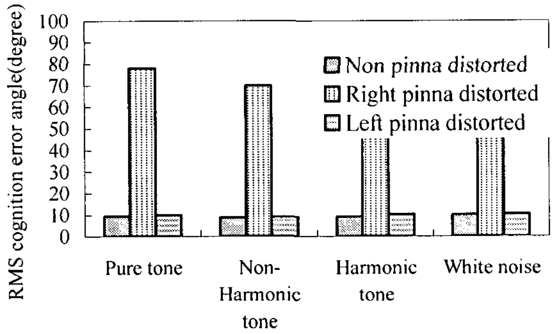

Fig. 6 Pinna distortion effect

Table 3 shows localization error for various sound types in the right, left and center area. The result showed that the total average of RMS cognition error at the center was the best (8.05°), The next best one was at the right side (11.29°). The error in the left side showed the worst (13.56°).

This hierarchy is also shown in the Fig. 7. Fig. 7 illustrates cognition error angles for various sound source types classified by the directions. In the right side, the direction-dependent cognition error in the angle was extremely higher in white noise. In left side, the cognition error of direction dependency was higher in harmonic tone. In center, the cognition error of direction dependency was highest in pure tone.

Table 3 Direction dependent cognition errors for various sound source types

Sound type	RMS cognition error (degree)		
	Center	Right side	Left side
White noise	15.25 ± 0.79	9.86 ± 0.81	10.61 ± 0.75
Pure tone	5.02 ± 0.39	7.37 ± 0.42	11.19 ± 0.58
Non-Harmonic tone	6.06 ± 0.46	13.14 ± 0.41	17.25 ± 0.56
Harmonic tone	5.69 ± 0.24	14.78 ± 0.35	15.19 ± 0.27
Average	8.05 ± 0.47	11.29 ± 0.42	13.56 ± 0.62

4. Conclusions

In this paper, the human spatial cognition using visual and auditory stimulation was investigated. Our results clearly indicated that the spectral response, due to directionally different pinna resonance effects, was the basic mechanism in the auditory cognition of the direction.

In conclusion to the human spatial cognition using visual stimulation, we have found the following factors.

1. The responsible mechanisms for visual stimulation and motor systems during eye-arm tracking may involve (1)the command addressed to the arm muscles (outflow). (2)an outflow copy. and (3)afferent signals originating from the arm muscles (inflow).

2. Human spatial cognition depends on information by human body than external sensory information. Thus the information of coordination and motor system (eyes, head and arms) interacts each other.

In conclusion to human spatial cognition using auditory stimulation, we have found the following factors.

1. It was found that harmonic components are not essential in the auditory target localization. However, non- harmonic nearby frequency components were more important in localizing target direction of sound.

2. The right pinna has one of the most important functions in localizing the target direction and pure tone with only one frequency component is confusing to be localized.

ACKNOWLEDGEMENT

This research was supported by the Korean Ministry of Education & Human Resources Development through the Center for Healthcare Technology Development.

REFERENCES

1. Bizzi, E., Kalil, R.E. and Morasso, P., "Two modes of eye-head coordination in monkeys," Brain Res., Vol. 40, pp. 45-48, 1972.

2. Angel, R. W. and Garland, H., "Transfer of information from manual to oculomotor control system," J. Exp Psychol, Vol. 96, pp. 92–96, 1972.

3. Gauthier, G. M. and Hofferer, J. M., "Eye tracking of self-moved targets in absence of vision," Exp Brain Res, Vol. 26, pp. 121–139, 1976.

4. Buizza and Schmid, R., "Velocity characteristics of smooth pursuit eye movements to different patterns of target motion," Exp Brain Res, Vol. 63, pp. 395–401, 1988.

5. Radeau, M., "Auditory-visual interaction and modularity," Curr. Psychol. Cogm., Vol. 13, pp. 3-51, 1994.

6. Middlebrooks, J. C. and Green, D. M., "Sound localization by human listeners," Ann. Rev Psychol., Vol. 42, pp. 135-159, 1991.

7. Perrett, S. and Noble, W., "The effect of head rotations on vertical plane sound localization," J. Acoust. Soc. Am., Vol. 102, pp. 2325-2332, 1997.

8. Kishon-Rabin, L., Amir, O., Vexler, Y. and Zaltz, Y., "Pitch discrimination: are professional musicians better than non-musicians," J. Basic. Clin. Physiol. Pharmacol., Vol. 12, pp. 125- 143, 2001.

9. Wier, C. C., Jesteadt, W. and Green, D. M., "Frequency discrimination as a function of frequency and sensation level," J. Acoust. Soc. Am., Vol. 61, pp. 178-184, 1997.

10. Sek, A. and Moore, B. C.," Frequency discrimination as a function of frequency, measured in several ways," J. Acoust. Soc. Am., Vol. 97, pp. 2479- 2486, 1995.

11. Miall, R. C., Weir, D. J. and Stein, J. F., "Intermittency in human manual tracking tasks," J Mot Behav, Vol. 25, pp. 53–63, 1993.

12. Miall, R. C., Weir, D. J. and Stein, J. F., "Planning of movement parameters in a visuo-motor tracking task." Behav Brain Res. Vol. 17, pp.1–8, 1988.

13. Vercher, J. L., Volle, M. and Gauthier, G. M., "Dynamics of human visuo-oculomanual coordination control in target tracking tasks," Aviat Space Environ Med, Vol. 64, pp. 500–506, 1993.

14. Vercher, J. L., Gauthier, G. M., Blouin, J., Guedon, O., Cole, J. and Lamarre, Y., "Oculo-manual tracking in normal subjects and a deafferented patient: respective role of arm motor efference and proprioception in initiation of smooth pursuit of self-moved targets," J Neurophysiol, Vol. 76, pp. 1133–1144, 1996.

15. Dallos, P., "The active cochlea," J. Neurosci., pp. 4575-4585, 1992.