Intelligent Steering Control System Based on Voice Instructions

Ki-Yeol Seo, Se-Woong Oh, Sang-Hyun Suh, and Gyei-Kark Park

Abstract: The important field of research in ship operation is related to the high efficiency of transportation, the convenience of maneuvering ships and the safety of navigation. For these purposes, many intelligent technologies for ship automation have been required and studied. In this paper, we propose an intelligent voice instruction-based learning (VIBL) method and discuss the building of a ship's steering control system based on this method. The VIBL system concretely consists of two functions: a text conversion function where an instructor's inputted voice is recognized and converted to text, and a linguistic instruction based learning function where the text instruction is understood through a searching process of given meaning elements. As a study method, the fuzzy theory is adopted to build maneuvering models of steersmen and then the existing LIBL is improved and combined with the voice recognition technology to propose the VIBL. The ship steering control system combined with VIBL is tested in a ship maneuvering simulator and its validity is shown.

Keywords: Fuzzy inference, intelligent steering control system, LIBL, VIBL.

1. INTRODUCTION

The most important issue in ship operation is how to prevent man-made accidents at sea and obtain higher competitiveness in maritime transportation and fisheries by reducing labor costs. As part of the effort to reduce the occurrence of human error in ship operation and achieve labor cost reduction, studies on digital ship and computer-guided integrated control systems have been actively conducted for years [1,2]. In addition, microprocessor and sensor technology have been applied to build automated ships to offer diverse information for safe navigation. The endeavor to create the ship maneuvering system aided by voice recognition has been continued as a way to ensure safe ship operation. Against this backdrop, this study is designed to highlight case studies that have tried to adopt voice recognition technology for safe ship operation.

The first study carried out is the study on a safe operation support system that applies voice input and output to merchant ships. In case of one person-bridge or emergency situations, voice command and voice query can be used to lighten the workload of ship personnel and boost efficiency in making responses [3]. The second study takes a look at the One Person Bridge Operation (OPBO) and the study on application methods of voice input and output [4]. The case studies above are breakthroughs in that they present the idea of attaining necessary information on ship operation with simple voice commands. However, since they are controlled by simple voice commands, it is impossible for the systems to make various linguistic instructions.

The third study performed is a look at the study of applying voice recognition technology to small fishing boats. It was aimed to resolve the labor shortage of small fishing vessels or crews' avoidance of getting on board small ships. The study suggests ways to enable workers of some 5-ton fishing boats to make voice instructions. The study is divided into the navigation support system, which controls devices for maneuvering, and a fishing support system, which controls fishing nets [2]. This study made significant headway by proposing not only a first scheme to embody a real system but also a plan to enhance actual voice recognition rate. Nevertheless, voice instructions are still limited to a few phrases only so that it is unable to deliver detailed messages for various situations and accurate intentions of instructors.

In the face of the challenges above, the fuzzy theory and machine learning-based linguistic instruction-based learning (LIBL) has been proposed by Park and

Manuscript received April 21, 2006; revised March 8, 2007; accepted June 21, 2007. Recommended by Editorial Board member Eun Tai Kim under the direction of Editor Jin Young Choi. This work was supported by the Korea Ocean Research and Development Institute (KORDI) and the Korea Research Council of Public Science and Technology (KORP).

Ki-Yeol Seo, Se-Woong Oh, and Sang-Hyun Suh are with the Ocean Exploration System Research Division, Korea Ocean Research and Development Institute, 171 Jang-dong, Yuseong-gu, Daejeon 305-343, Korea (e-mails: {vito, osw, shsuh@moeri.re.kr).

Gyei-Kark Park is with the Dept. of Maritime Transportation Systems, Mokpo National Maritime University, 571 Jukgyo-dong, Mokpo-si, Chunnam 530-729, Korea (e-mail: gkpark@mmu.ac.kr).

Sugeno [5-7], where a mechanical system is controlled by linguistic instruction and its performance is improved by it. This method utilizes the fuzzy theory to interpret linguistic instructions with the assistance of existing system knowledge, and the interpretation is used to operate the system. The performance of LIBL has been proven by its application to numerous systems such as the Truck Backer Upper Control System, the Helicopter Flight Control System [8], and the Dynamic System Control such as motor control [9], the control rule management method [10] and the Fuzzy Classifier System [11]. However, this method has its own shortfalls as well. While using natural language makes it possible to communicate with systems without much effort, the interpretation process of linguistic instructions is very complicated, so that simple text-based linguistic instructions should be used. Because of such restrictions, the method's practical application doesn't seem to be available anytime soon.

To address the problems related to the search of meaning elements, it should be that when a command is made, dialogue-type sentences should be generated to ask the operator what meaning elements are contained in the concerned linguistic instruction. The text-based linguistic instruction method should be replaced with the voice recognition-based linguistic instruction to make sure of the easy linguistic instruction process. Therefore, as an upgraded version of the LIBL, this study proposes the VIBL system which tackles the problems existing with the LIBL, and desires to test the VIBL's effectiveness with the practical application to the ship steering system.

In this paper, we focus on the building of a ship's steering control system for supporting safe navigation and resolving crew labor shortage. It proposes the Voice Instruction Based Learning (VIBL) method applied to a steering control system of a ship with improved LIBL.

As a study method, the fuzzy theory is adopted to build the maneuvering models of steersmen and then the existing LIBL is improved and combined with the voice recognition technology to propose the VIBL. Lastly, the ship steering control system combined with VIBL is tested in a ship maneuvering simulator on a PC and its validity will be shown.

2. VOICE INSTRUCTION-BASED LEARNING

As set forth in Fig. 1, the Voice Instruction-based Learning (VIBL) system consists of the voice recognition system and the linguistic instruction-based learning (LIBL) system. The ship operator issues voice instructions such as steering commands, while watching the system conditions, and the system transforms the relevant voice instructions to text instructions with the voice recognition technology

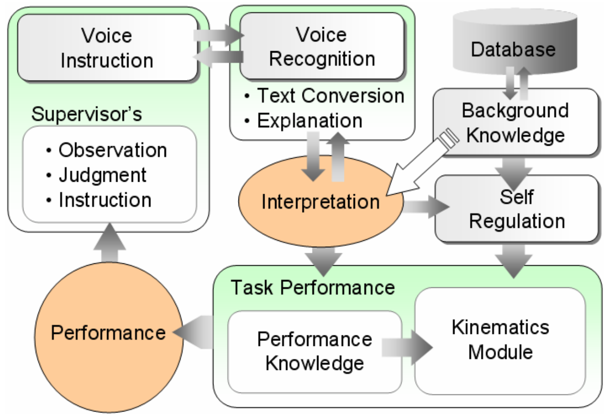


Fig. 1. Framework of VIBL system.

using the voice recognition part. The system then checks in the database of the Background Knowledge Part whether concerned instructions exist. If the same linguistic instructions are found, the background knowledge part applies the evaluation rules for revising primary control rules that have been already created by the interpretation of linguistic instructions. If the previous linguistic instructions are non-existent, the input linguistic instructions are interpreted by using the meaning elements in the background knowledge's database to figure out the meaning. Then the evaluation rules chosen using selected meaning elements are utilized to self-adjust the primary control rules of the Performance Knowledge Part and then the steering system can be operated.

2.1. Voice recognition part

2.1.1 Instruction formation

The voice instruction is performed using a microphone, and its formation is as follows.

$$L_i = [AP][HA][LH][AW],$$

where L_i is an instruction, AP means fuzzy auxiliary phrases, HA means heading angle, LH means fuzzy linguistic hedge, and AW means a fuzzy atomic word, respectively.

The Voice Recognition Part recognizes and converts voice instructions of ship operators into text linguistic instructions. When it is applied to the steering control system, the following shows how it works:

$$\boldsymbol{L_i} = [\textit{Follow at } (\boldsymbol{AP})][\textit{180 degrees } (\boldsymbol{HA})] [\textit{More}(\boldsymbol{LH})][\textit{Fast}(\boldsymbol{AW})].$$

The Background Knowledge Part is divided into the regular database and the knowledge database.

In the regular DB, the commands regularly used by the system are stored, while in the knowledge DB, the evaluation value of the meaning elements are stored to set up the control of the system.

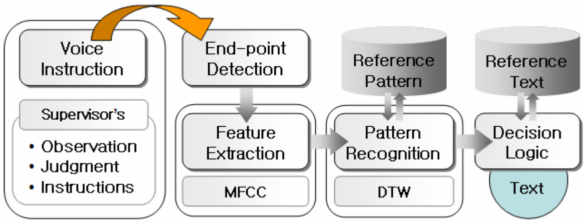


Fig. 2. Block diagram of voice recognition system.

2.1.2 Voice recognition

As set forth in Fig. 2, the voice recognition system consists of the voice instruction and the end-point detection, the feature extraction, the pattern recognition, and the decision logic [12]. If the operator's voice instruction is input, it detects the end-point of the voice signal. And it extracts the feature of voice information using the MFCC (Mel Frequency Cepstral Coefficients) method. The DTW (Dynamic Time Warping) algorithm using the pattern matching method is based on the templates adopted for the pattern recognition, and it selects a text that corresponds to the reference text by the decision logic rules. So it is formed into a linguistic instruction by combining the selected texts.

2.1.3 Text conversion

Fig. 3 displays a flowchart for converting the operator's voice instruction into text. If the operator's voice instruction is input, it detects the end-point of the voice signal. It also extracts the feature of voice information using the MFCC method. For distinguishing sharply between the voice and the noise,

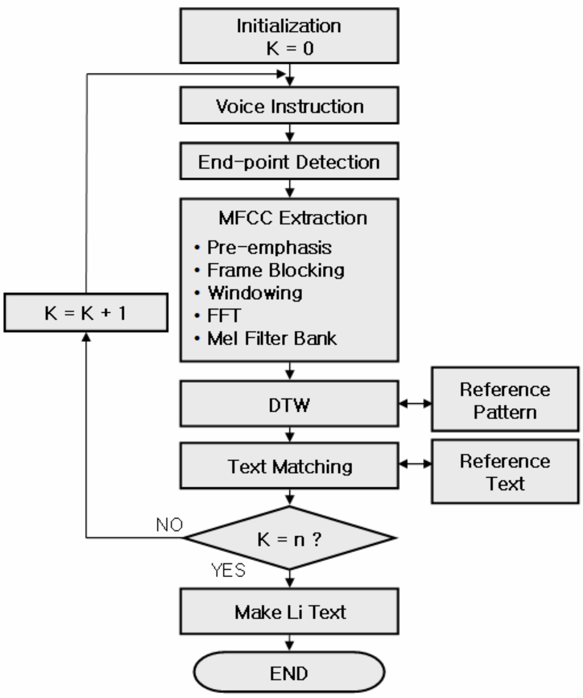


Fig. 3. Flow chart for the text conversion.

the pre-emphasis is implemented for highlighting the elements of high frequency, and it divides the emphasized voice signal into the sample blocks. The hamming windows are used for minimizing the discontinuity of the frame.

Using the windowing function, it analyses the voice signal, and then determines the spectrum using the FFT (Fast Fourier Transform). Corresponding with the filter bank by the mel-scale, it determines the size and the summation in each band, and extracts the last MFCC converting the output values of the filter bank. Then using the DTW algorithm, it selects the nearest frame, operating the feature vector of a standard voice signal with an input signal. When the input voice is recognized, it extracts the feature vector analyzing the input voice, and then finds the member that gets the nearest distance using the DTW algorithm respectively, comparing with the standard model set, and selects the text.

2.2. The interpretation part

2.2.1 Selection of meaning elements

In order to make a decision on system response results, the Steering Gear Control System has three meaning elements and three trends each. The three meaning elements include ΔmR_θ referring to other rudder angle; ΔmT_S referring to the arrival time; and ΔmS_θ referring to stability status. The three trends include increase (+), maintenance (0), and decrease (-). Equation (1) shows the selection result of meaning elements. Refer to the following:

$$L_i = (LH_i)(\Delta mR_\theta(+)\ \ and\ \ \Delta mT_S(-)), \quad (1)$$

where LHi refers to linguistic hedge, ΔmR_θ (+) refers to increase in rudder angle and ΔmT_S (-) refers to decrease in arrival time respectively.

Fig. 4 displays the process of selecting trends of the meaning elements to evaluate linguistic instructions of operators [5]. In Fig. 4, n means the number of meaning elements. In case of the steering control system, if n=3, LHM means a pre-defined linguistic hedge module. First of all, when the linguistic instruction of the ship operator is entered through the dialogue box, the comparing and checking process is launched to decide whether the concerned linguistic instruction exists in the database.

If the matching linguistic instruction is not found, the conversation with the operator is conducted to determine meaning elements, and the new linguistic instruction is added to the database. Likewise, when meaning elements are determined, a search is conducted to find out whether the entered linguistic instructions contain linguistic hedge. If it is found that the linguistic hedge exists, weight value is allocated according to concerned linguistic hedge. If the

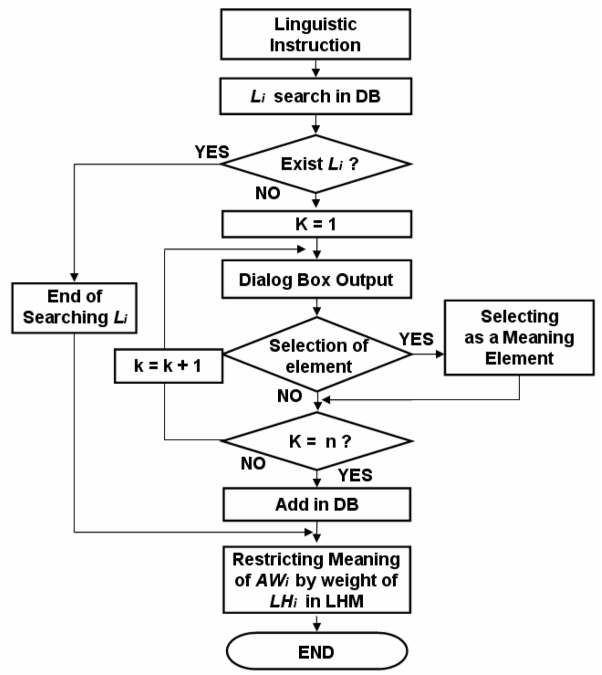


Fig. 4. Selection process of the meaning elements.

concerned linguistic instruction exists in the existing database, searching of linguistic instructions is stopped, and weight value is allocated depending on linguistic hedge. As described above, the meaning elements and trends of the linguistic instructions of the ship operator are searched and compared, and the evaluation rules are created accordingly.

2.2.2 Generation of evaluation rules

When each meaning element and its trend are found for linguistic instruction, the Background Knowledge Part is utilized to generate evaluation rules by element and trend. The fuzzy membership function of the premise used by a meaning element includes *SMALL*, *MED*, and *BIG*. As a consequence, three fuzzy membership functions by the trend of the meaning element should be ready in response to membership functions of the premise. Fig. 5 displays an example of evaluation rule when ΔmR_θ (+), a meaning element referring to the increase of heading angle to starboard, is chosen.

Meaning elements can be limited by linguistic hedge in the way of reflecting the value of W_{LHi} on the evaluation rule. Use Equation (2) to calculate ΔH, the value of movement of parameter in the consequence, and prepare the evaluation rule limited by linguistic hedge.

$$\Delta H = W_{LHi} \cdot \Delta R, \tag{2}$$

where the maximum movement of the consequence of the evaluation rules by linguistic hedge is set at ΔR = 5.0.

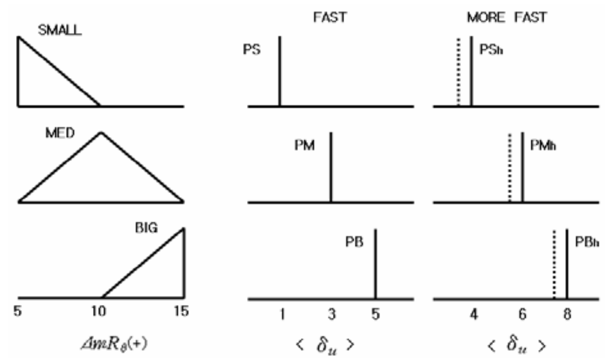


Fig. 5. Composition of the fuzzy evaluation rules.

The movement of the consequence caused by linguistic hedge [*MORE*] (W_{LHi}=0.6) is $\Delta H = 0.6 \cdot \Delta R$, and [*MORE FAST*], the changed membership function of the consequence is prepared. As Fig. 3 indicates, the final consequences of the evaluation rule, which is reflected by the linguistic hedge effects, are *PS**, *PM**, and *PB** respectively. When weight value is allocated, the weight value per linguistic hedge is the following:

[(None, 0.0), (A Little, 0.2), (Some 0.4), (Further, 0.6), (More, 0.8), (Very, 1.0)].

The evaluation rule applicable to meaning elements has a single input/output fuzzy inference engine. For instance, increase in rudder angle, ΔmR_θ (+) has the following rules.

If ΔmR_θ (+) is SMALL, then δ_u is PS,
If ΔmR_θ (+) is MED, then δ_u is PM,
If ΔmR_θ (+) is BIG, then δ_u is PB,

where MED means medium, PS means positive small, PM means positive medium, and PB means positive big.

2.3. Self-regulation part

This part uses the additional rudder angle δ_u by applying the linguistic instruction earned according to evaluation rules, and calculates the steering angle δ and δ_u by applying the primary control rule indicating a steering control model by steersman and combines them to conduct self-regulation.

When a linguistic instruction is made, it is calculated as the final rudder angle $\delta*$ based on (3).

$$\delta^* = \delta + \delta_u. \tag{3}$$

3. STEERING CONTROL SYSTEM

In this paper, the system's task is actually performed, and it includes Performance Knowledge

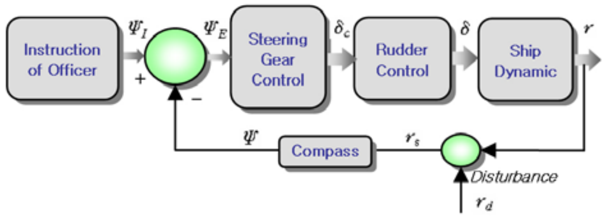


Fig. 6. Block diagram of steering control system.

and Kinematics Module corresponding to a controller and a plant respectively in a control system. In case of the steering control system, the Performance Knowledge and Kinematics Module are in response to the steersman control model and ship dynamics respectively.

In general, the steersman control model utilizes the results of the research and analysis of the experiences of skilled steersmen and creates a rudder angle control model based on the results and the fuzzy interference model. In general, the block diagram of a regular steering control system is as shown in Fig. 6.

Equation of ship's heading angle applied in this study is as follows:

$$r = R\sqrt{\upsilon_t} \cdot \delta^*,$$

$$\Psi = \int_0^t r^* dt, \; r^* = r + r_d,$$

$$\delta^* = \delta + \delta_u.$$

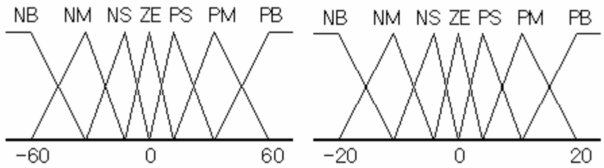
It determines the rate of turn and heading value by adding the result of fuzzy inference through linguistic instruction to rudder angle. In addition, *r* represents the rate of turn by rudder bending, *R* is calibrating constant (0.01), *vt* means ship speed, δ means inferential rudder angle steered by steersman model, δ_u means rudder angle by linguistic instruction, $\delta *$ refers to aggregated rudder angle of δ and δ_u rate of yaw due to the interference of wind or wave, and ψ means current direction, respectively.

The error (ψ_E) between the set direction (ψ_I) and the current direction (ψ) is used as the premise input value. In this way, the rudder angle (δ) of the consequence can be inferred, and the controlled output (*r*) is calculated. The membership functions used for the premise and consequence for simulation are as shown in Fig. 7.

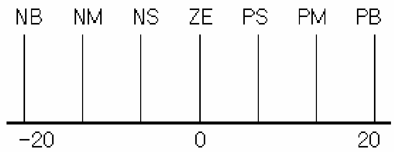
The steersman control model designed based on experiences of a steersman consist of 13 control rules like the following:

If ψ_E is PB and $\Delta\psi_E$ is ZO then δ is PB,
If ψ_E is ZO and $\Delta\psi_E$ is NB then δ is NB,
If ψ_E is NB and $\Delta\psi_E$ is ZO then δ is NB,

where PB means positive big, ZO means zero, and NB means negative big.



(a) Membership function of ψ_E and $\Delta\psi_E$ [deg].



(b) Membership function of δ [deg].

Fig. 7. Membership functions of ψ_E, $\Delta\psi_E$, δ.

The consequence of control and evaluation rules adopts the fuzzy singleton. Mamdani's inference method was used for the inference method. The center of gravity (COG) method was used as the defuzzification method.

4. SIMULATIONS

4.1. Linguistic instruction mode configuration

The linguistic instruction mode is set differently depending on situations. Normal Mode is chosen when using the general steersman control model. The linguistic instruction mode without linguistic hedge (L Mode) is chosen in case of linguistic instruction order, and the linguistic instruction mode with linguistic hedge (Lh Mode) is selected when linguistic hedge is added. Their results are displayed on the screen. Normal Mode following the set heading angle according to the regular steering control model. L Mode adds the inference results based on the linguistic instruction of the general steering control model and controls rudder angle. Lh Mode searches linguistic instructions and when linguistic hedge is found, the concerned linguistic hedge is added to control the rudder angle. The example of linguistic instructions for each mode is described in Table 1.

4.2. Simulation results

In order to build a simple ship handling simulator on a PC, M/V SAEYUDAL, a training ship of the

Table 1. Example of linguistic instructions.

Mode	Linguistic Instruction
Normal Mode	"Follow at 180°"
L Mode	"Follow at 180° in a faster speed"
Lh Mode	"Follow at 180° in an even faster speed"

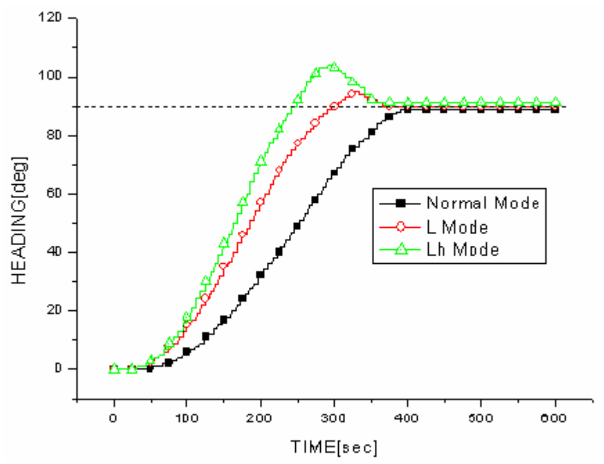


Fig. 8. Shift from 0° to 90°.

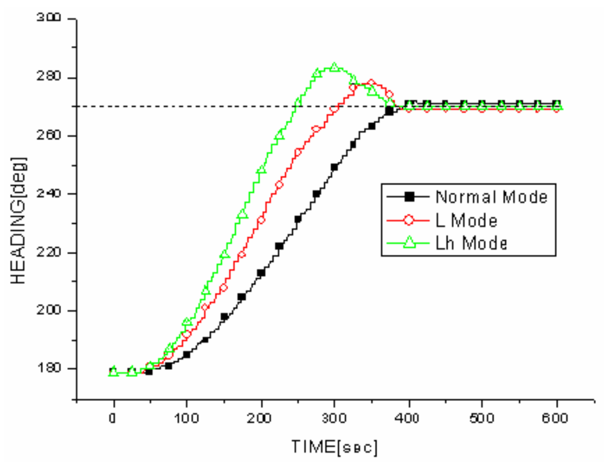


Fig. 10. Shift from 180° to 270°.

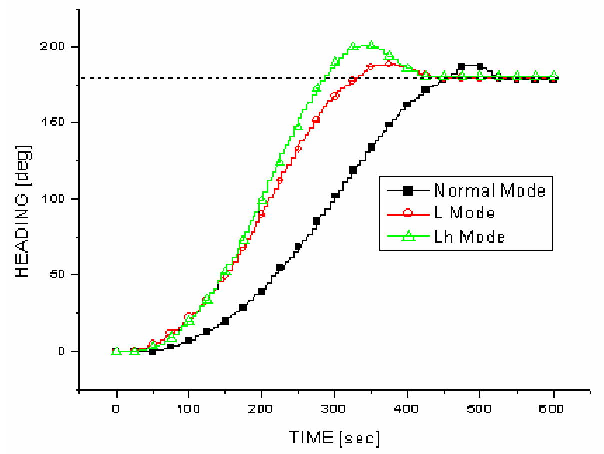


Fig. 9. Shift from 0° to 180°.

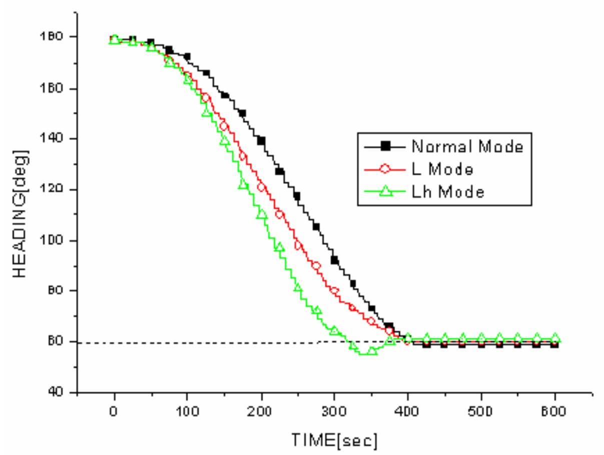


Fig. 11. Shift from 180° to 60°.

Mokpo National Maritime University was adopted as a model. For simulation, our assumption is that the ship's speed is kept constant at 200rpm and wind and waves were not taken into consideration. As indicated in Table 1, simulations were conducted separately for each mode including Normal Mode of the steersman control model, L Mode, and Lh Mode. The results were displayed according to the types of following at the same heading angle.

Figs. 8 and 9 depict what would happen in case of changing the current direction from 0° to 90° or from 0° to 180°.

The responses of each mode, including Normal Mode, L Mode and Lh Mode, have different characteristics. When it comes to the time taken to arrive at the set direction, the Lh Mode performs work at the fastest speed. However, when it comes to stable following, the L Mode shows the result with less overshoot. When a ship operator needs to have quick following at the given angle and launches a linguistic instruction in L Mode, the performance of the rudder is improved with faster arrival time. When requiring more quick following, the ship's operator can make the arrival time faster with linguistic instruction Lh Mode as in Figs. 8 and 9.

Figs. 10 and 11 show directional changes from 180° to 270° and 180° to 60°. The results were similar. When it comes to the time taken to arrive at the set direction, the Lh Mode shows the fastest response and improvement in stability and following hours. Namely, it is verified that the VIBL system indicates an improved performance accounting to the operator's linguistic instruction and its intention.

5. CONCLUSIONS

The most important issue in research of ship operation is how to prevent accidents due to human-error and maintain safe navigation at sea. For this purpose, we proposed an intelligent steering control system using the VIBL, which can replace a steersman, requiring less labor. In the system, the steering gear can be controlled by the ship operator's linguistic instruction and its intention, and it is available to reduce overload of multitasks caused from labor shortage on ships and to sustain a safe navigation environment.

This study suggested the method of recognizing and converting voice linguistic instructions of ship operators into text message, and has established a steering control model based on steersman's experiences. In order to build and realize an intelligent steering control system, it made a proposal on the meaning element for rudder angle, arrival time and stability status in rudder control respectively. The evaluation rules to regulate the fuzzy rules of the steersman control model were proposed, as well. Lastly, we built a ship maneuvering simulator to test the efficiency of the proposed system. The VIBL method was applied to the steering control system of the ship maneuvering simulator and the effectiveness was shown from some simulations. This system will be expected to be used as a supporting system for safe navigation through its improvement and tests in actual situations.

The study conducted was not practical because the effect of wind, wave, and other weather conditions which affect steering, were not taken into consideration. As such, to get a more accurate system, such factors influencing steering should be considered. For making a real system, it is necessary to build prototype hardware which can realize voice recognition and the proposed algorithm, as well as undergo outdoor tests in the future.

REFERENCES

[1] Y. K. Lim, J. W. Park, and Y. C. Choi, "A research for digital-ship," *Journal of Ships & Ocean Engineering* (Korean), vol. 33, no. 1, pp. 15-20, 2002.

[2] K. Karasuno, H. Oda, K. Igarashi, and T. Nitta, "Support system voice recognition for maneuvering and working on fishing boats," *Journal of Japan Navigation*, vol. 92, no. 1, pp. 57-64, 1995.

[3] N. S. Son and S. Y. Kim, "Development of a supporting system for safe navigation using voice control and guidance," *Journal of Ships & Ocean Engineering* (Korean), vol. 33, no. 1, pp. 117-125, 2002.

[4] J. Fukuto, M. Numano, N. Shimono, and K. Matsuda, "Countermeasures against malfunctions of a one person bridge operation support system by using speech communication," *Journal of Japan Navigation*, vol. 102, no. 1, pp. 63-70, 1995.

[5] G. K. Park and K. Y. Seo, "Intelligent ship's steering gear control system using linguistic instruction," *Korea Fuzzy and Intelligent Systems Society* (Korean), vol. 12, no. 5, pp. 417-423, 2002.

[6] G. K. Park and M. Sugeno, "Learning based on linguistic instruction using fuzzy theory," *Journal of Japan Society for Fuzzy Theory and Systems*, vol. 4, no. 6, pp. 1164-1181, 1992.

[7] G. K. Park and M. Sugeno, "An approach to linguistic instruction based learning," *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems*, vol. 1, no. 2, pp. 19-56, 1993.

[8] G. K. Park and M. Sugeno, "An approach to linguistic instruction based learning and its application to helicopter flight control," *Proc. of the 5th International Fuzzy Systems Association World Congress, in Korea*, vol. 5, no. 1, pp. 1082-1085, 1993.

[9] G. K. Park and M. Sugeno, "Linguistic instruction based learning and its applications," *Proc. of International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and the Second International Fuzzy Engineering Symposium*, vol. 4, no. 1, pp. 1891-1896, 1995.

[10] A. Miyamoto, K. Goto, O. Itoh, H. Migita, and M. Sugeno, "Adjusting method of rules by linguistic instructions," *Journal of Japan Society for Fuzzy Theory and Systems*, vol. 8, no. 5, pp. 928-935, 1996.

[11] M. Fujii and T. Furuhashi, "A proposal of extraction of intentions in linguistic instructions using fuzzy classifier system," *Journal of Japan Society for Fuzzy Theory and Systems*, vol. 11, no. 1, pp. 92-98, 1998.

[12] W. C. Chun, *Surveys on the Domestic Voice Information Industries and Development of Business Mode* (Korean), *Acoustical Society of Korea*, 2002.

Ki-Yeol Seo received the Ph.D. degree in Marine Information System Engineering from Mokpo National Maritime University, Korea in 2003. He is a Research Engineer in the Ocean Exploration System Division at the Korea Ocean Research & Development Institute. His research interests include maritime information systems, intelligent control systems, and GNSS systems & technology.

Se-Woong Oh received the M.S. degree in Logistics Engineering from Korea Maritime University in 2001. He is a Research Engineer in the Ocean Exploration System Division at the Korea Ocean Research & Development Institute. His research interests include intelligent information systems, marine GIS, ECDIS, and maritime information systems.

Sang-Hyun Suh received the B.S. and M.S. degrees in Ship Engineering from Seoul National University in 1979 and 1982, respectively, and the Ph.D. degree in Ship Engineering from the University of Michigan, U.S.A. in 1991. He is a Director in the Ocean Engineering Research Department at the Korea Ocean Research & Development Institute. His research interests include GNSS systems & technology, marine GIS, and maritime information systems.

Gyei-Kark Park received the Ph.D. degree in Systems Science from Tokyo Institute of Technology, Japan in 1993. He is a Professor in the Department of Maritime Transportation System at Mokpo National Maritime University. His research interests include fuzzy control, intelligent control systems, and decision making systems.