An Image Based Linguistic Guiding System(IBLGS)

Gyei-Kark Park* · Cheol-Yeong Lee**

畫像을 利用한 言語的 가이딩시스템

朴 桂 珏 · 李 哲 榮

〈Contents〉

요 약

從來의 내비게이션시스템은 주로 船舶을 對象으로 利用되어 왔으나, 最近에는 各種 移動로보트 및 自動車에 대한 適用事例가 增加하고 있다. 내비게이션시스템은 機械的인 移動시스템뿐만 아니라, 移動이라는 行動을 취하는 人間시스템 즉, 視覺 不自由者를 對象으로 適用될 수 있다.

從來의 機械的인 盲人用 가이딩시스템으로서 盲導犬로보트가 있으나 階段등의 複雜한 環境下에서는 案內가 不可能한 점과 盲人의 意思가 反映되기 어렵다는 점 등의 많은 問題點을 안고 있다. 現實的으로 要求되는 案內시스템은, 단지 環境에 관한 各種情報를 提供하고 意思決定과 行動은 盲人이 實施하는 시스템의 形態이다. 本論文에서는, 盲人用 내비게이션시스템의 構築을 目標로 하는 中間課題로서, 畫像을 利用한 言語的 가이드시스템을 提案한다. 當該시스템 構築에 있어서 重要한 하나의 課題는, 주어진 環境에 관한 畫像情報를 言語로 變換시키는 일이다. 本研究에서는, 가이드를 위한 言語指示의 生成에, 定性的이고 定量的인 屬性을 가지고 言語的 表現에 有效한 퍼지理論을 利用한다. 具體的인 一例로서, 公園에서 벤치까지 盲人을 가이드 하는 狀況을 설정하여 言語指示를 生成하는 알고리즘을 提案한다.

1. Introduction

An important area of application of intelligent system is that of developing user friendly intelligent aids for the handicapped. We consider here the requirements for a blind aid navigation system. The aid for the partially blind can be in the form of special equipment to enhance the visual

* Technical Laboratory of ILSAN INDUSTRIAL CO., LTD.
** Dept. of Port and Transportaion Engineering, Korea Maritime University

ability of the user. However, for the totally blind such equipment must be replaced by a guiding system. A robot system equivalent to the guide dog has been sought and attempted in [1] and [2]. However, limitations of current robotics technology have made difficult the realization of an "artificial guide dog". In addition, the task is made difficult by the desire of the user to be able to interact with the guiding system. Strictly speaking, the blind person needs a physical system only to the extent to which this system provides accurate information for his use. For example, such information could be given in natural, or specialized (Braille) language. Information of environment-spatial relation, obstacles, quality of the path, and traffic conditions (pedestrian and vehicle) adjusted to the individual needs are expected to enable the user to perform the desired actions.

In this paper, we focus on a intermediate step towards developing a Blind person Aid Navigation System(BANS) based on a linguistic interpretation of visual information. More precisely, we aim at building an Image Based Linguistic Guiding System(IBLGS), where the visual and numerical information are fusioned and linguistic instructions are generated. From this point this paper is organized as follows : In section 2 the general architecture and functions for IBLGS are introduced. In section 3, for illustration purposes we consider the example of "finding a bench in a park and describing the path to reach the bench".

2. Overview of IBLGS

Currently, we envision the architecture of the IBLGS as consisting of seven functional components(Fig. 1) : Dialogue, Target Identification, Image Processing, Image Based World Map, Image Based Controller, Language Data Base, and Generation of Linguistic Instructions Components. We now describe briefly these components.

2. 1. Dialogue

This component is responsible for the interface between the user and the IBLGS system. The dialogue takes place in order to identify the goal and to guide the user towards it. Typically it consists of the following steps :

- The user requests the IBLGS to find the target object. (A search for the desired object is triggered by a call to the Target Identification Functional Component).
- Upon successful identification of the target object the user asks the system to describe the path from the current location to the target location.
- Adequate linguistic instructions are generated taking into account the user's individual walking characteristic abilities.

2. 2. Target Identification

This component consists of two modes for identifying a given objective target : Target Image Memory and Pattern Matching mode. The basic method for this is the model based object recognition method of [3] which makes use of the results of the image processing component and a model stored in the Target Image Memory.

If the target object exists, the system informs the user and waits for the users's next request(e.g. to be guided to the target object or find another target). If the user wants to be guided to it, the system gives the linguistic instructions to approach to the target. The instructions are generated using the given numerical information of the target and environmental space described in the Image Based World Map component. The given numerical information in the world map is the position of the identified, the situation of obstacles, the route to the target, etc.

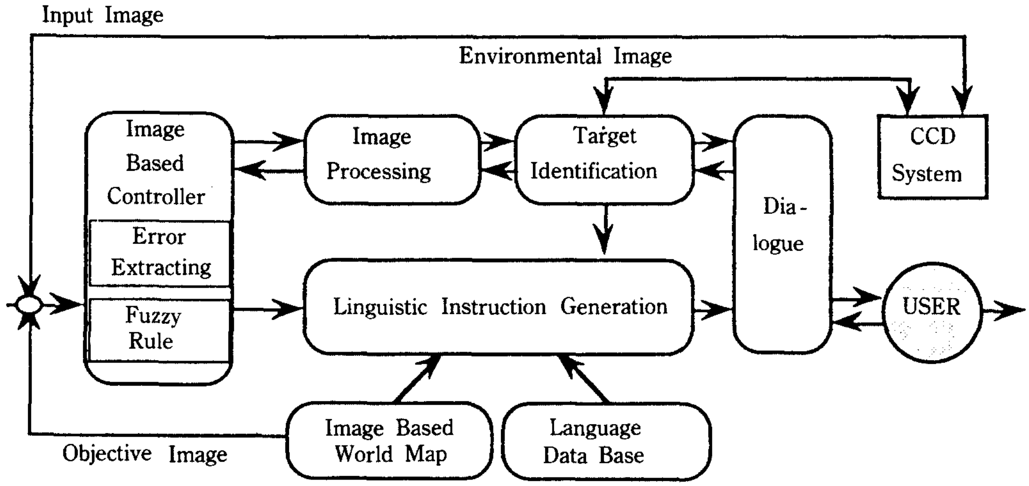

Fig. 1 General Framework for IBLGS

2. 3. Image Processing

The traditional image processing method is adopted in IBLGS. The input image is provided by a CCD system.

2. 4. Image Based World Map

A world map is required for guiding the user to the objective target. The guiding is done by comparing the input image to the image in critical points along the path. These images, or alternatively, the descriptions, are stored in a data base. The image retrieval method based on descriptions of the contents of the image developed in [4] can be extended to handle this step. The Image Based World Map component is constructed from numerical data and image data. The World Map component consists of two modes : the numerical map of a given environmental space and the guiding image data base. In IBLGS, we assume that the information necessary for these two modes is given, i.e., the route and position of obstacles are assumed known.

2. 4. 1. Numerical Map Mode

The minimum information given by this mode concerns location of target, location of obstacles, and a route to the target.

2. 4. 2. Guiding Image Data Base Mode

Two kinds of information are needed along the path : (i) information concerning the approach to a critical check point, and (ii) information on each check point corresponding to the direction from which it is approached.

2. 5. Image Based Controller

In this section, we describe two modes for Image Based Controller component : Image Variation Extracting and Fuzzy Inference mode.

2. 5. 1. Image Variation Extracting Mode

The basic mechanism for obtaining error information is a match between stored knowledge about the environment at, or around the check points and the input image. The matching is done either at the image level, or at the level of descriptions of the contents of the image. Two kinds of error values between the input picture and an element of the given image data base are obtained :

(i) Error in Variation for left or right direction(EV) : This error results from the left or right variation between the objective image of the world map and the input image at a check point.

(ii) Error in Height(EH) : This is the distance error between a check point and the user's real position. It can be obtained from the error in height between the input image and the corresponding stored image.

2. 5. 2. Fuzzy Inference Mode

In this mode the fuzzy rules needed to correct the user's actions are constructed. Initially, these rules will encode the knowledge and experience of an expert. In a more advanced version of the system these could be, at least partially, learned. The range of each input variable, EV and EH is covered by a fuzzy partition of cardinality three. Therefore nine fuzzy IF-THEN rules are constructed. The consequent part is the correction value of the user's direction.

2. 6. Language Data Base

In this component, two kinds of linguistic expressions needed for the communication with the user are generated : linguistic instructions for attaining the goal and linguistic instructions for approaching a check point. These linguistic instructions are generated from Atomic Words(AW), Linguistic Hedges(LH) and Auxiliary Phrases(AP) using rules of combination specified in a special purpose grammar[5] [6]. The atomic words are expressed by fuzzy sets whose membership functions are obtained from local computations based on the current position and desired direction. The auxiliary phrases explain the goal and intermediate environment.

2.6.1. Instructions for a Goal

Linguistic instructions on the route and the environmental space for the user before approaching the target object are generated :

— Atomic words for explaining the approach direction : These words are the membership functions for the objective target. Examples of such words are : about left angle, medium left, straight, medium right, and about right angle.

— Linguistic hedges for explaining the approach direction : For a finer representation of the approach direction hedges are used to change atomic words. This is done by shifting the center value of the membership function of an atomic word. Each linguistic hedge for the approach direction is specified by a linguistic label and the corresponding shift : (slightly, 0.2), (rather, 0.4), (more, 0.6), (pretty, 0.8), and (very, 1.0).

— Atomic words for number of steps necessary for approach : This mode supplies two membership functions for the number of steps needed to reach an objective point : about 'n' steps, about '(n+1)' steps.

— Linguistic hedges for explaining the number of steps necessary for approach : These hedges changes the approach step word. It is specified by a linguistic label and the value for shifting the center of the fuzzy set corresponding to an atomic word. Examples of such linguistic hedges are : (slightly, 0.15), (rather, 0.3), (more, 0.5).

— Auxiliary phrases for subgoals : This mode supplies the phrases for explaining a subgoal for approaching the next check point. Examples of such phrases are : "Approach the point P1", " Turn left/right at the point P1", "Walk n steps".

— Auxiliary Phrases for situation explanation : These phrases are for explaining a given environmental space : before, will pass, will arrive, etc. For example : "You will pass a fountain on the left".

2.6.2. Instructions for a check point

Linguistic instructions for the correction of the user's approach of the target based on the values derived from the Image Based Controller are generated.

— Atomic words for explaining the left/right variation : These words express the membership functions of EV and indicate the direction in which the user erred, to the left or to the right, from the given path.

— Linguistic hedge for left/right variation : The linguistic hedges used to refine the explanation of the left/right variation specify the amount of shift of the center value and the associated label : (a little, 0.3), (rather, 0.5), (pretty, 0.75).

— Atomic words for explaining the distance error : These words express the membership functions of EH and indicate the user's advancement with respect to the objective point : near, just, far.

— Linguistic hedge for the distance error : The distance error explanation membership functions are changed by these linguistic hedges : (a little, 0.3), (rather, 0.5), (pretty, 0.75).

— Atomic words for explaining corrections in the direction : These words modify the user's action on the basis of the values obtained from the fuzzy inference rules in the Image Based Controller. Examples of these words are : left angle, medium left, straight, medium right, right angle, eight o'clock direction, and four o'clock direction.

— Linguistic hedge for corrections in direction : The membership functions of atomic words for corrections of directions are changed by the given hedge based on the shift necessary in the center value, as follows : (slightly, 0.2), (rather, 0.4), (more, 0.6), (pretty, 0.8), and (very, 1.0).

— Subgoal auxiliary phrases : This mode supplies the phrases for explaining a subgoal : "approach the point P2", "turn around the point P2 to the left".

— Auxiliary phrases for situation explanation : These phrases explain a given environmental space using terms such as : in front of, will pass, will arrive, etc.

2.7. Linguistic Instruction Generation

In this component, linguistic instructions are generated using the image based world map, the language data base controller.

2.7.1. Format of an Instruction

A linguistic instruction consists of three kinds of elements : Linguistic Hedge(LH), Atomic Word(AW) eventually modified by a linguistic hedge and Auxiliary Phrases(AP). A dictionary of words and the grammar rules needed for combining these are also supplied. For example, a linguistic instruction might be as follows : "walk straight, turn a little to the right after five steps", In this example 'walk' and 'turn' are auxiliary phrases, 'straight', 'right', and 'five steps' are atomic words, and 'a little' is a linguistic hedge.

The general algorithm for generating a linguistic instruction has the following steps :

Step.1 : Matching with atomic word.

Step.2 : Matching with linguistic hedge.

Step.3 : Construction of linguistic instructions taking into account the following :

3.1 the user's characteristics

3.2 output of the Image Based Controller

3.3 the user's request.

The following section will illustrate this algorithm on a simple test problem.

3. Instruction Generating Example

In this section we explain the process of generating the linguistic expression on the test problem of finding a bench in a park. We assume that the following are known :

– the Image Based World Map, and all numerical data for the environmental space ;

– the bench and its location(identified by the Target Identification component in Fig. 2) ;

– the error between an objective image and the current input image(obtained by Image Variation Extracting Mode and Image Processing component).

3.1. World Map for the "bench problem"

This map consists of the image of two check points P1, P2 and numerical data in the X-Y euclidean space in Fig. 2. The data for the route are as follows :

– Distances(in meters) between two points as follows : dist(P0, P1)=25, dist(P1, P2)=7.5, dist (P2, P3)=12.5.

– The User's characteristics : length(1 step)=60(cm) and walking speed=1.5(step/sec).

– Approaching angle : $E\Delta_1=20°$, $E\Delta_2=70°$, and $E\Delta_3=40°$.

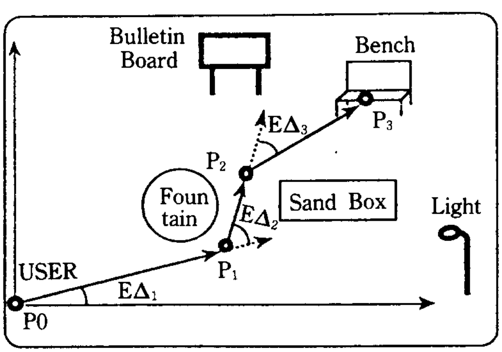

Fig. 2 Image Based World Map

3. 2. Linguistic Instruction Generation

Two kinds of linguistic instruction are generated : a global linguistic instruction(for the entire approach to the before beginning to start) and local linguistic instructions(for each check point).

3. 2. 1 Global Instruction Generation

The process of generating linguistic instruction is explained by four steps.

– Step.1(matching with atomic words) : The atomic words are chosen according to the highest membership value corresponding to the current input. Thus for the situation in Fig. 3. we have the following choices :

i) "Medium Left" is selected for the input angle 20°from P0 to P1(see Fig. 3, for direction of approach).

ii) "About 42 Steps" is selected for the distance between P0 and P1(see Fig. 4.).

– Step.2(choosing linguistic hedges) : For deciding the direction of approach the difference between the center of "medium left" and the input is RD=(–45)–(–20)=25. The portion of the support of 'Left Medium' in the positive direction is obtained : PS=45. The shifted rate of the input from the the Left Medium is RD/PS=0. 56.

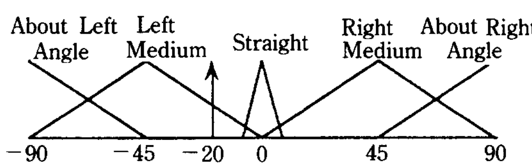

Fig. 3 Membership functions for the atomic words indicating the direction of approach

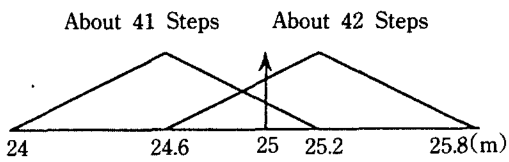

Fig. 4 Membership functions for the atomic words indicating the number of steps

According to the definition of the linguistic hedges for this mode the linguistic hedge 'more' that has nearest weight from 0.56 is selected. The truth modifier 'not' is used if the input is negative and right from the center of a selected atomic word, or if the input is positive and left of the center of a selected atomic word. Thus if the input is 20, the selected linguistic hedge would be 'not more'.

For deciding the number of approach steps needed for approach a similar procedure is used : The membership function shifting rate is obtained to be : RD= | 25.2−25 | = 0.2, PS=0.6(the given step), and RD/PS=0.33. The linguistic hedge 'rather' is selected. The modifier 'not' is used for input values to the left of the center of the atomic word selected. Thus when the input is 25, the selected linguistic hedge is 'not rather'.

– Step.3(subgoal auxiliary phrases combination) : The subgoal auxiliary phrase is constructed from the Language Data Base and the world map. Examples of subgoals phrases combinations are :

i) "Approach the point P1(AP), not more(LH) than left medium(AW)".

ii) "Walk(AP) about 42 steps(AW) not rather(LH)".

– Step.4(situation explanation instruction) : If any obstacles exist to the left or to the right of the P0-P1 route, a situation explanation instruction is constructed by combining the world map and the situation auxiliary phrases mode of from the language data base. On the route P0-P1, the combined instruction is : "You will pass a fountain to the left".

The entire linguistic instruction to P3 from P1 is generated by combining the approach direction instruction and the situation explanation instruction. All the linguistic instructions are conveyed to the Dialogue component.

3.2.2. Instruction for Check Point Approach

The check point linguistic instructions are generated from the given variation error(EV) and height error(EH) by applying the Fuzzy Inference mode. The fuzzy rules for correcting the user's approach are given in Fig. 5. : If EH < 0, the user is to the left of the approach direction, and if EH > 0, the user is to the right of the approach direction. If EV < 0, the user has not arrived at a check point.

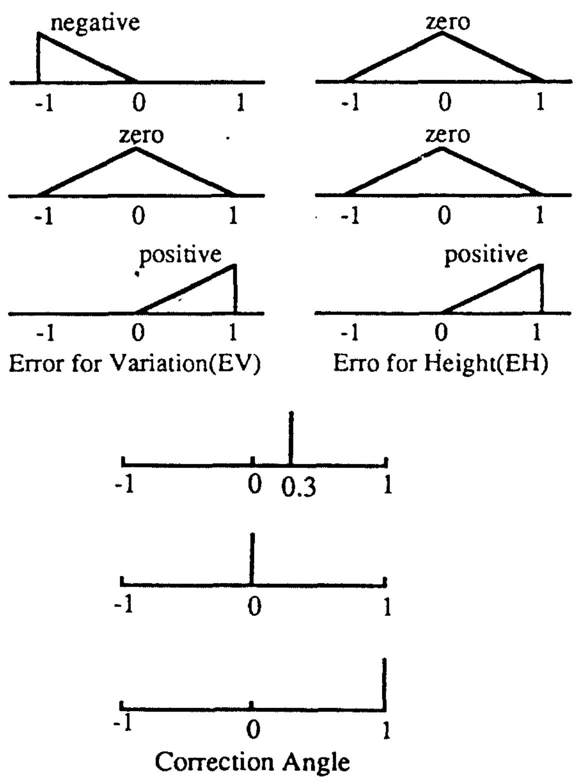

Fig. 5 Fuzzy inference rules for correcting the angle of approach

If EV > 0, the user has already passed the check point. The universe of discourse for the consequent part is the correction value to the user's direction of approach of the next check point. The value −1 is corresponding to −20° and the value 1, 20°. The left direction is negative and the right direction, positive. The user's approach direction is obtained from the inference value and the approach angle to the next check point on a given route in Figure 1. If the output is −10° in approaching the point P1, the user's approach angle is (70+(−10))=80° to the left from the direction of walking. If the output is 10°, the approach angle is 60° to the left. The correcting value 14.2° for the input EH=0.5 and EV=0.8 is obtained from the fuzzy inference rules of Figure 5. as follows : The obtained value is $y^* = (\Sigma\omega_i/\Sigma\omega_i) \cdot 20 = 14.2$, where y_i is a support of the consequent part and $\omega_i = A_{i1} \cdot A_{i2}$, and A_{i1}, A_{i2} are the degree to which the input matches the premises.

The process of linguistic instruction generation is explained using the output 14.2°of the Image Based Controller.

– Step.1(matching with an atomic word) : The approaching direction 55.8° to the left is obtained from the output 14.2° and the approach angle −70°. The obtained value is matched with the given approach direction atomic words(Fig. 3.). The membership function 'Medium Left' is selected as the direction for the entire approach.

– Step.2(matching with linguistic hedge) : The rate of difference RD between the center of Left Medium and the input is calculated RD= | (−45) −(−55.8) | = 10.8. The support for the positive direction of the Left Medium is PS=45. The shifted rate of the input from the Left Medium is RD/PS= 0.24. The weight of linguistic hedge 'slightly' is nearest to 0.24 and therefore is selected. However, since the value 0.24 is greater than the center of "Left Medium" the modifier 'not' is added to the linguistic hedge, to obtain 'not slightly'.

– Step.3(subgoal auxiliary phrases combination) : The subgoal auxiliary phrase is constructed from the Language Data Base and the World Map : "approach the point P2(AP), left medium(AW), not slightly(LH)".

– Step.4(situation explanation) : A situation explanation is constructed by combining the world map and situation auxiliary phrases mode of the Language Data Base. On the route P_1-P_2, the combined instruction is as follows : "you will pass a fountain to the left and a sand box to the right".

4. Conclusion

A general framework for IBLGS(Image Based Linguistic Guiding System) as an intermediate step towards building a Blind person Aid Navigation System(BANS) has been proposed. The assumption underlying this proposal is useful and information extracted form the environment can be conveyed to the blind person in a useful linguistic form enabling the user to have initiative in the interaction with the system and to perform desired actions. The algorithms necessary for generating linguistic instruction make use of the input image and an image based world map and were illustrated by the problem of finding a bench in a park. The framework of IBLGS will be effective in building some non-invasive monitoring systems and the intelligent bridge system for automating vessel.

References and related bibliography

1) S.Tachi et al. : "Study on guide dog(Seeing-eye) robot (1)", Bulletin of Mechanical Engineering Laboratory, 32, (1978).
2) H. Mori et al. : "A mobile robot strategy-Stereotyped motion by sign pattern", in Robotics Research The fifth International Symposium,

The MIT Press, 161−172, (1990).

3) K. Miyajima, A. Ralescu : "Fuzzy logic approach to model based image analysis", OE/TECHNOLOGY '92, Vol. 1827, (1992).

4) M. Nakayama, K. Miyajima, H. Iwamoto, T. Norita, "Interactive Human Face Retrieval System Based on Linguistic Expression", Proc. of 2nd International Conference on Fuzzy Logic and Neural Networks, IIZUKA '92, Vol. 2, pp. 683−686.

5) M. Sugeno, G.K. Park, "Learning Based on Linguistic Instructions Using Fuzzy Theory", J. of Japan Society for Fuzzy Theory and Systems, Vol. 4, No. 6, 1164−1181, (1992).

6) M. Sugeno, G.K. Park, "An Approach to Linguistic Instruction Based Learning", Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 1, No. 1(1993), pp. 19−56.