Caller Identification Based on Cognitive Robotic Engine

Sukhan Lee, Hun-Sue Lee, Seung-Min Baek,
Jongmoo Choi, and Dong-Wook Shin
School of Information and Communication Engineering
Sungkyunkwan University
300 ChunChun-Dong, JangAn-Gu, Suwon, KyungGi-Do, 440-746, Korea
lsh@ece.skku.ac.kr, smonkeys@hanmail.net

ByoungYoul Song and Young-Jo Cho
Intelligent Robot Research Division
Electronics and Telecommunications Research Institute
161 Gajeong-Dong, Yuseong-gu, Daejeon, 305-700, Korea
sby@etri.re.kr

***Abstract*— An approach to identifying a caller or callers by a service robot is presented for a natural interaction with people in a home/office environment. The problem addressed specifically in this paper is how to successfully identify a caller in a cluttered environment with large uncertainties involved in the sensed audio-visual cues. The proposed approach is based on a proposition that the dependability of perceptual recognition may come unlikely from "the effort to make individual sensing perfect," but likely from "the effort to self-generate perceptual behaviors of integrating individual sensing that lead to mission accomplishment, no matter how imperfect and uncertain individual sensing may be." We implement the above proposition in terms of a novel robotic architecture, referred to here as "Cognitive Robotic Engine (CRE)." CRE implemented for the case of a robot identifying a caller in a crowed and noisy environment, including its experimental results, are shown.**

Index Terms – Perception, Action, Dependability, Cognitive Robotic Engine, Caller identification

I. INTRODUCTION

One of the major reasons why most of service robot prototypes remain in research laboratories, instead of being transferred to commercial sectors, may be because the issue of robot dependability has not been solved yet. Perhaps, we may not have a clear idea how to solve the entire issue of robot dependability at this moment either. Compared to robots, it seems that human is quite dependable. What makes human dependable in its perception, action, and, in general, its performance for accomplishing tasks? Human dependability may be not so much to do with the quality of individual components that constitute perception or action. But, it may be much to do with an overall system behavior by which these individual components, regardless of how uncertain and imperfect they are, are coordinated into dependable system behaviors. The question is whether we can establish a formalism of system architecture and control that implements such dependable system behaviors as human. In this paper we focus on the robot's dependability in the perception field and introduce the Cognitive Robotic Engine (CRE) that takes the best out of the human approach to reliable and dependable perception.

A. Background

Dautenhahn [10] has listed different social relationships between robots and humans based on animal-human relationship. In recent years there has been a great deal of interest to develop algorithms and systems for human-robot interface, for instance see [2][8]. Most of these researches focused on the problem of proper response to the human while the reliable and dependable recognition and perception remains of a critical issue. On the other hand, the robot's hardware dependability has been under investigation for a long time, i.e. executing valid and safe commands, from the old industrial robots to the new mobile platforms. There are many examples of employing robots in public areas [7][11][12][13], such as museums and exhibition halls, in which dependability has been achieved by performing limited tasks trying to execute valid and safe commands which are verified in hard real-time. Fritsch et. al. [2] have proposed a system based on three layer architecture to provide a flexible infrastructure suitable for human-robot interaction. The simpler interactions, such as gesture detection, are handled at the reactive layer while the sophisticated interactions, such as speech recognition and understanding, have been dealt with in the deliberative layer. Although the architecture allows incorporation of different human-robot interaction modules, however, it does not address handling uncertain situations in which the robot needs further information or clue to make an action.

In this paper we present the Cognitive Robotic Engine that is regarded as a more general form of behavior based approach [14] that is extended to include perceptual behavior. In this architecture, an asynchronous and concurrent flow of perceptual building blocks are connected to the actions to be taken proactively to collect further sensing data to resolve uncertainty and provide additional evidence. This will guarantee the dependability of the robot for a natural human-robot interaction.

II. COGNITIVE ROBOTIC ENGINE (CRE) – CONCEPTUAL OVERVIEW

As an introduction of the concept of CRE, let us consider how a human identifies a caller, if there is, dependably despite the adverse condition of, say, a crowded and noisy party environment. Upon hearing a novel but highly uncertain nature of sound that may indicate someone calling, one immediately registers in his/her consciousness an ad-hoc mission of verifying if there is a caller, if any. The mission will remain in his/her consciousness till the verification is done with a sufficient level of confidence an individual set. With the registered mission producing a stress, a flurry of asynchronous and concurrent perceptual processing takes place inside in such a way as to reduce the uncertainty as efficiently as possible.

The basic perceptual building blocks represent the perception units of various levels of abstraction and complexity, ranging from a lower level of elementary sensing primitives to a higher level of conceptual recognition units. They are assumed predefined, although individuals may have a different number, type, and quality of these building blocks. In the above example, a sufficient amount of evidences may be quickly assembled from multi-modal sensing cues, including both auditory and visual cues such as calling hand gestures and/or calling facial expressions, generated by an asynchronous and concurrent flow of auditory and visual perception building blocks. Potentially, there exist a large number of possibilities for laying out the flow of building blocks, with individual sensors as origins, and for acquiring distinct evidences from different paths of the chosen asynchronous concurrent flow. However, to explore all the possible flows of evidences may be neither possible due to a limit in computational resources, nor desirable for efficiency in time and energy. The key may be to understand an optimal way of constructing an asynchronous concurrent flow of perceptual building blocks for decision, dynamically to the real-time variation of situations and, perhaps, similarly to the way our brain functions.

The perceptual building blocks are connected to the actions to be taken proactively to collect sensing data of less uncertainty or of additional evidences. Human seldom depends passively on what is sensed only for a decision, if the sensed information is incomplete or uncertain. Rather, human tends to take appropriate actions for gathering a better quality of or a new addition of information. For instance, in the above example, to reduce uncertainty, one may resort to look around to see if he/she can find someone making a calling gesture or to generate a confirmation call, like “Is there anybody calling?”, to get a reply. Human dependability of perception is thus deeply linked to the proactive actions for assisting and guiding perception by incorporating action building blocks into an asynchronous and concurrent flow of perceptual building blocks. The key is how to choose action blocks to be incorporated into an asynchronous and concurrent flow of perceptual building blocks in such a way as to achieve an optimal overall efficiency in reaching the decision. This requires evaluating an action block as an information source against the cost in time and energy to exercise it.

Summarizing the above, human dependability in perception may be conjectured as the result of the following exercises:

1) The spontaneous and self-establishment of ad-hoc perceptual missions in connection to particular sensing that drive the subsequent perceptual processes till satisfied.

2) The choice of particular asynchronous and concurrent flow architecture of perceptual building blocks out of a potentially huge number of possible flow architectures as the basis for deriving evidences to be fused together.

3) The incorporation of action blocks into the chosen asynchronous and concurrent flow architecture of perceptual building blocks as a means of proactively collecting sensing data of less uncertainty and of new evidence, which triggers a dynamic reorganization of the asynchronous and concurrent flow architecture of perceptual building blocks.

4) The optimal process control in terms of the choice of a particular asynchronous and concurrent flow architecture of perceptual building blocks to follow as well as of the choice of particular action blocks to be invoked at each sampling time, where the optimality is defined in terms of the time and energy to be consumed for completing the ad-hoc mission, which is in turn a function of the amount of uncertainty reduction and the time and computational resources required for completing the perceptual and action building blocks to be processed. Note the control strategy may differ by individuals since some heuristics are involved in the strategy, due to the complexity of search space leading no definite optimal solution is feasible. However, it is interesting to see that this heuristics actually represent a personality of an individual or a robot that we can exploit for creating robots with personality.

The environment or toolkit that enables the above asynchronous and concurrent flow of a perceptual process, or, in general, a robotic process, is referred to here as Cognitive Robotic Engine (CRE). In what follows, we present a more details on how to implement the above concept in computer by describing 1) an asynchronous and concurrent architecture for CRE with the definition of perceptual and action building blocks, the representation of search space with the partial order and fusion relation of perceptual building blocks as well as with the exclusion relation and organized actions for action building blocks, 2) a method of connecting perceptual and action building blocks, 3) an optimal control of CRE with self-establishment of ad-hoc missions, of choosing a particular flow architecture with the optimality in terms of speed and time, and, finally, 4) a demonstration of the value of CRE by a caller identification experimentation with a robot .

III. ORGANIZATION OF PERCEPTUAL PROCESSES

A. Perceptual Process

Perceptual processes are defined here as basic building blocks of perception. Fig. 1 and Table I show the input-output relation of a perceptual process and its meaning.

There can be a large number of perceptual processes available as building blocks of a perception system. They are assumed independent as long as they are not under any predefined sequencing constraints.

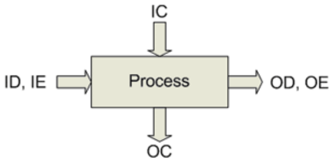

Fig. 1. Input-output relation of a perceptual process

B. Def. Precedence Relation

The perceptual processes are under a precedence relation if the outputs of processes are necessary as an input or inputs of another for processing.

TABLE I
DESCRIPTION OF A PERCEPTUAL PROCESS

Name		Description
Input	IC	Control input for Activation/deactivation of the process
	ID	The inputs from the precedent processes having a precedence relation
	IE	The certainty/uncertainty measures associated with ID
Output	OC	A list of actions that can be invoked to enhance OE
	OD	The output of the process with the associated processing time as an attribute
	OE	The certainty/uncertainty measure associated with OD

C. Def. Fusion Relation

The processes that are inputs of a process A with the OR relations among them are considered in a fusion relation for the process A. Processes under a fusion relation can be used for creating multiple evidences and reducing uncertainty associated with the output for the process A.

Fig. 2 illustrates the precedence and fusion relations defined for a caller identification problem with the assumption of the availability of auditory and visual cues.

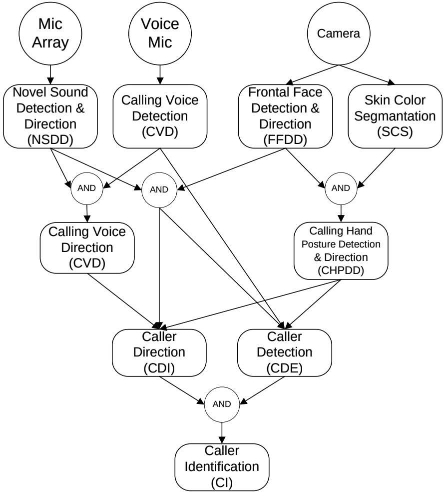

Fig. 2. The Precedence of Perceptual Processes defined for the caller identification problem – All the relations without AND means OR.

D. Asynchronous and Concurrent Architecture for Perceptual Behavior

According to the precedence relations, we can structure the perceptual processes to achieve a given mission. There can potentially be a huge number of possible structures for accomplishing a given mission that satisfy the predefined precedence relations. Here, the key question is how to design an optimal asynchronous and concurrent architecture of available perceptual building blocks under the constraint of time and computational resources, where the optimality criteria may be defined in the sense of the speed and computational resources required for accomplishing a given mission. Since reducing uncertainty plays a crucial role for reaching a decision for the given mission, the architecture shows explicitly not only concurrency for time but also fusion relations for uncertainty management. Refer to Fig. 2. Among the relations, particularly selected flow of processes represent a perceptual behavior.

IV. ORGANIZATION OF ACTION PROCESSES

A. Action Processes

An action process represents a behavior module associated with individual actuators. We organize action processes in terms of whether they are mutually exclusive for execution or not, as well as of whether action processes are under a certain rules for subsumption or coordination, representing collective behavior of a group of action processes. Those action processes under an exclusiveness constraint can not be run concurrently due to resource conflicts. Table II shows a list of action processes and their definitions for caller identification mission.

TABLE II
A List of Action Processes

Behavior	**Sub Behavior**	**Definition**
Head behaviors	Gazing (GA)	Continuous head rotation to the interested object
	Look around (LA)	Head rotation random searching
	Heading (HE)	Head rotation following the direction
Base behaviors	Wandering (WA)	Moving constant velocity with Obstacle avoidance
	Approaching (AP)	Keeping distance after go to goal posture.
	Turning (TU)	Turn to intended direction
Sound behaviors	Speech Action (SA)	Robot's speech that enables to interact with human
Group behaviors	Searching (SE)	Wandering behavior + Look around behavior
	Attentive Approaching (AA)	Gazing + Approaching

V. CONTROL FOR OPTIMAL ACCOMPLISHMENT OF MISSIONS

We implemented the control procedure as a processing module for the experiment. There are three main features in the control module. 1) Selecting the best set of processing modules using the results of each module and the precedence and fusion relationship. 2) To arbitrates among the actions in order to reduce the uncertainty of each module. 3) Each processing modules run and communicate concurrent and asynchronously. The detail description is in the following.

The solution space S consists of sensors $s(i)$, processing modules $p(j)$, mission modules $m(k)$ and the relationship R.

$$x = \{m(i), p(j), p(k), \ldots\} \tag{1}$$

A solution x is a set of candidates that includes the mission and the modules related with the mission. When a mission $m(i)$ invoked, there should be a set of solution $X(m(i))$.

$$X(m(i)) = \{x(1), x(2), \ldots\} \tag{2}$$

Although $m(i)$ could be achieved by a subset of $Y \subset X$, each solution x of Y gives different results and behaviors, because running time of each module and confidence values provided by the modules are different. To provide efficient solution, the control module selects the best solution x^* in each time t,

$$x^* = \min_i \{C(x(i))\} \tag{3}$$

where $C(x(i))$ is a cost function of the solution $x(i)$. The control module has tables that represent the precedence relationship and the fusion relationship defined in Fig. 2. Because the precedence and fusion graph includes AND-OR relationship, the control have to search based on the relationship. Each module $p(i)$ have a set of child $\{p(j), p(k), \ldots\}$ from the precedence relationship. The cost of $p(i)$ is propagated by the set and determines whether the relation of the set is AND or OR. Cost is divided into values based on confidence values of modules and values based on the processing time and resources.

The control module arbitrates among the actions. Each module has a set of candidate behaviors that could reduce the uncertainty of each module. In each processing step, each module sends an id of the best behavior and requests a permission to act the behavior. It means that every module knows the best action for uncertainty reduction. The control module has a table that includes concurrency relationship among behaviors so that the module could use the table for arbitration.

Up to now, the control for the mission is limited to the certainties which are module certainty, current certainty. Module certainty is evaluated by controller. This certainty value is usually fixed so that it works as weight function. And current certainty is normalized value from each module.

We defined a communication packet in order to communicate to each processing module, that includes a header and data fields, where the header contains an id of the module and the size of the data fields, and the data fields contains the certainty of the results, an id of required action and expected certainty when the action acted. We implemented the module by event driven concept in order to handle concurrent processing. For the purpose, we defined few queues which are including an event, mission, waiting, running, and action queue.

VI. DESIGN OF HARDWARE/SOFTWARE ARCHITECTURE FOR IMPLEMENTATION

The proposed CRE concept is implemented into the real robot system and experimented to investigate accomplishing process of a given mission. Here, we assume that the mission is finding the user calling the robot using his voice and hand motion. This caller identification is one of challenging problems and it is necessary function to intelligent service robot. The conventional approaches to solve caller identification problem are improving component technologies such as voice recognition, face detection, and etc. or multimodal approaches mixing voice/sound and vision based recognition. However, the human response to the calling signal is different from single component based identification or multimodal approaches. He/she not only uses eyes and ears for recognition but also mixes proactive motions such as rotation for finding the caller simultaneously. It means that sometimes motions can reduce the uncertainties effectively rather than only processing the given information. Accordingly, in this experimentation we try to mimic the human approach for dependable caller identification by implementing CRE concept. Fig. 2 describes the precedence relation of component processes for the caller identification. The input/output descriptions are defined as Table II. Detailed description of each perceptual process in Fig. 2 is shown in Appendix.

Fig. 3 shows overall system structure of CRE-Bot system. Asynchronous and concurrent perceptual processes using multi-thread and timer functions are implemented in different computers. After the controller in the server collects the information packets from all the processes, it requests some action behaviors to the arm and base computer.

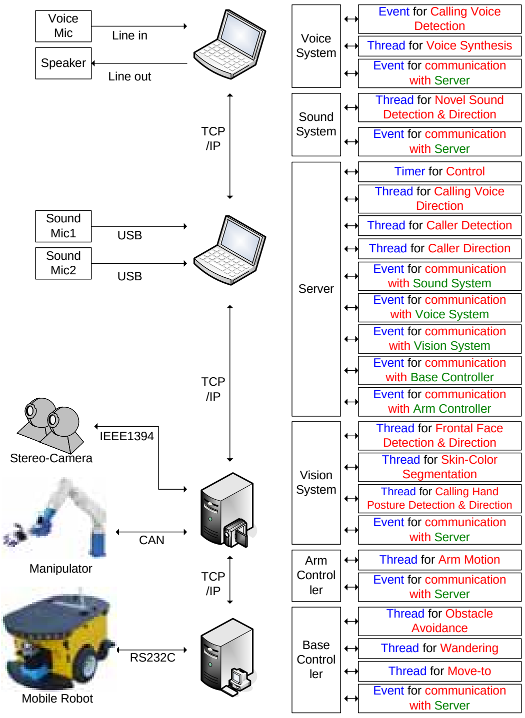

Fig. 3. CRE-Bot system

Fig. 4 also shows the system structure in another robot platform. In this system, a robot and a server computer communicate by what is called Common Robot Interface Framework (CRIF). The robot only gets data from sensing device or performs basic process. The data, transmitted by CRIF, are processed in applications (CRE) of server computer.

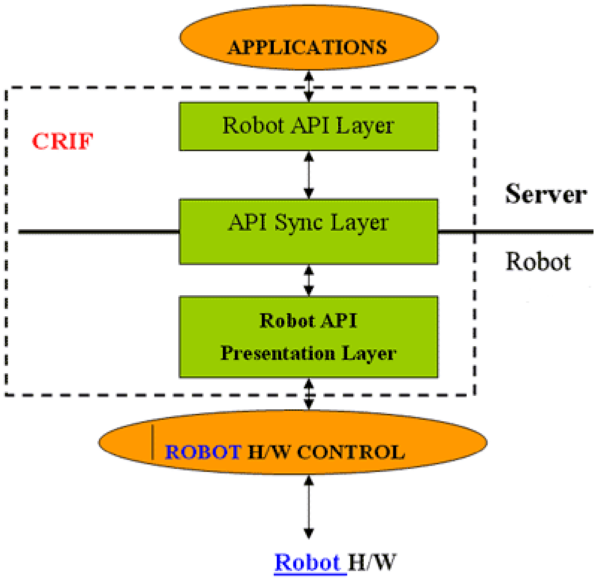

Fig. 4. A concept of Common Robot Interface Framework(CRIF)

For the last, Table III shows the action behaviors from each perception and fusion processes. For instance, in FFDD process, it suggests the controller to invoke Approaching behavior when a face is found with low confidence. All the invoked process suggests one action behavior as they get some information from the environment. And then the controller selects at least one of them.

TABLE III
Description of candidate action behaviors from perception and fusion processes

Process	Candidate action behaviors	
CVD	SA	
NSDD	Low Certainty	LA, SE
	High Certainty	HE, TU
FFDD	Low Certainty	SE, AP
	High Certainty	AA, GA
SCS	Low Certainty	SE
	High Certainty	GA
CHPDD	Low Certainty	LA
	High Certainty	AP, AA
CVD	SA, LA, SE, HE, TU	
CDI	SA, HE, TU, AP, AA, GA	
CDE	LA, SE, AP	
CI	SA, LA, SE, AA, WA	

The implemented control procedure consists of the mission manager and the behavior selector. The mission manager invokes ad-hoc mission based on evidences from each perceptual process and terminate it based on mission certainties. For example, if the output certainty of Calling Voice Detection process is higher than certain threshold value the caller identification mission will be invoked automatically. So the mission manager has the perception-mission relation table that has certainty threshold values to manage ad-hoc missions.

The main role of the behavior selector is selection of proper behavior among proposed behaviors of all perception processes to increase mission certainty. For the selection, the evaluation of perceptual process that considers output certainties and heuristic weights of modules is necessary. Using following criterion, we could find the proper modules and then the suggested candidate action of the module can be selected.

$$x^* = \max_i\{W(i)\cdot CC(i)\} \tag{4}$$

Here, *CC(i)* and *W(i)* are the current certainty of *i*-th module and the weight of *i*-th module, respectively. The weight of each module can be selected considering the relation between the module and the given mission. For example, the weight of CVD module is higher than the weight of NSDD module, because that the output of CVD module is more closely related with the caller identification mission. If one result from perception module has high certainty, the behavior selector allows the module to invoke proposed action behavior so that it can achieve the given mission. The max value varies as the environment changes and then the selected behavior also changed for getting higher certainties of perception modules.

VII. RESULTS

There can be lots of situations which human calls a robot. For the evaluation of implemented CRE, we briefly assumed that there are mainly 2 situations. Maybe a caller exists in front of the robot. The other case is the caller exists besides the robot or exists behind the robot. Fig. 5 shows a caller calls a robot in front of it. Thus the robot can find him/her easily using vision and sound processes. The results of action selection after caller identification mission invocation are shown in Fig. 6 and 7. The second situation and the results of action selection are depicted in Fig. 8, 9 and 10.

We took the robot to other indoor place of 6-square meter size. For each situation, we operated the robot 10 times.

In Fig. 6 and 7, as soon as the robot began the mission it found the caller by vision cues. The robot attentively approached to the caller and completed the mission finally. On the other hand, in Fig. 9 and 10, we can see the robot firstly had to find where the caller is. The robot used to search for caller using SE, WA and LA behaviors. And if vision or sound perception modules detect appropriate direction of caller, the robot look at the caller using HE and TU behaviors.

AA, AP and GA behaviors make the robot close to the caller for achieving higher mission certainty. We tested the robot system a large number of times in various situations. And most of cases, the mission have been achieved in proper time duration as shown above results.

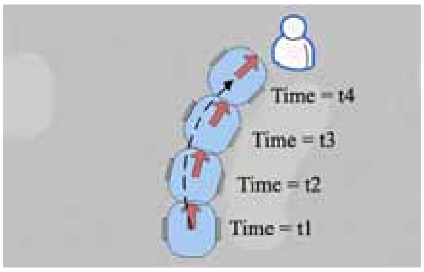

Fig. 5. 1st caller identification situation (User calls a robot in front of it)

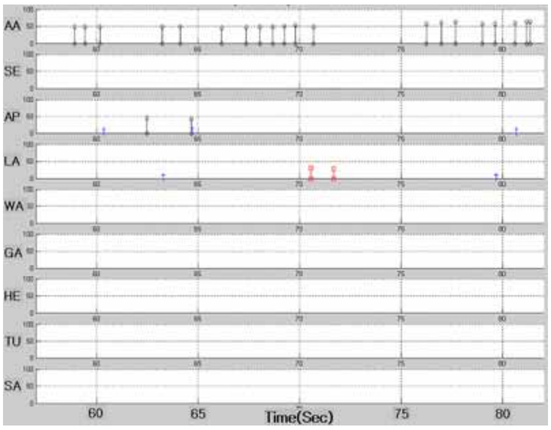

Fig. 6. Triggered action records in the experimentation of Fig. 5

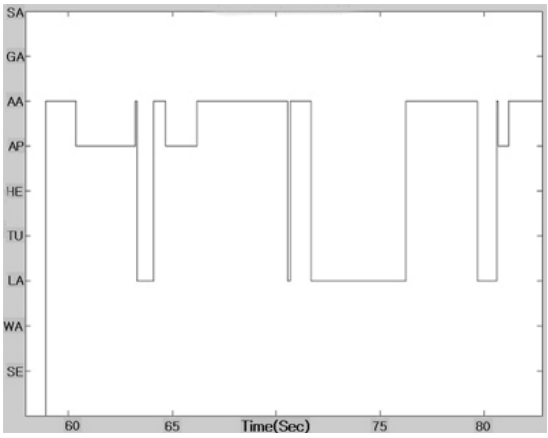

Fig. 7. Action behavior transition graph for the Fig. 6

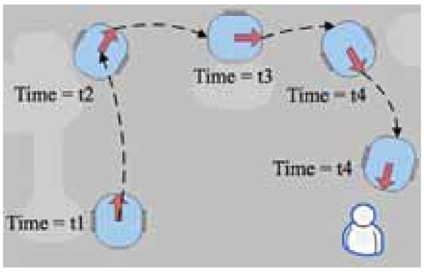

Fig. 8. 2nd caller identification situation (User calls a robot right side of it)

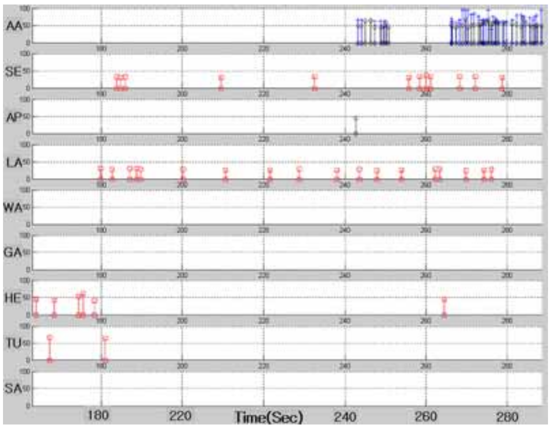

Fig. 9. Triggered action records in the experimentation of Fig. 8

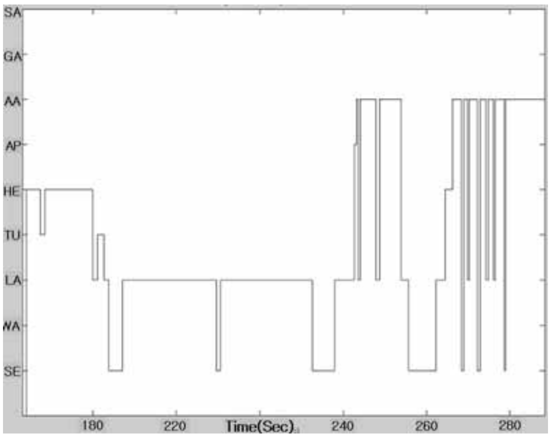

Fig. 10. Action behavior transition graph for the Fig. 9

VIII. Discussion And Conclusion

In this paper we introduced the Cognitive Robotic Engine that takes the best out of the human approach to reliable and dependable perception. To implement this dependable perception, we used asynchronous and concurrent perceptual processes and proactive action behaviors for reducing uncertainty and collecting additional evidences.

As shown in results, proactive action behaviors make the robot to combine outputs of perceptual processes and to achieve the caller identification mission finally. For example, if the robot couldn't detect any face or something that implies human by visual perception, it actively began to find caller due to the sound perception processed. Note that reactions of the robot can be varied if some parameters such as heuristic weights of perceptual processes are modified. It means that the proposed system is deferent from traditional rule based action selection systems or behavior based reaction systems.

There are still lots of research issues such as how to find optimal data flow path in the search space considering resource limitation and how to manage multiple missions.

APPENDIX

DESCRIPTION OF PERCEPTUAL PROCESSES IMPLEMENTED FOR CALLER IDENTIFICATION

NS DD	**Definition**	When the sound volume exceeds the threshold, estimates the direction of source
	Source	Mic Array
	Input	Raw data of sound
	Output	Direction of novel sound, Certainty (CC) Candidate of action behavior and certainty improvement expected (BE) Processing Time (PT)
	Destination	CVD, CDI, CDE
CVD	**Definition**	Recognize voice commands which this module knows
	Source	(Voice) Mic Array
	Input	Raw data of voice sound
	Output	CC, BE, PT
	Destination	CVD, CDE
FF DD	**Definition**	Finds face region by image feature
	Source	Camera
	Input	Raw image from Camera
	Output	Coordinate, and size of detected face CC, BE, PT
	Destination	CHPDD, CDI
SCS	**Definition**	Distinguishes skin region by RGB condition and makes others black in image
	Source	Camera
	Input	Raw image from Camera
	Output	Skin color segmented image. Most probable direction that callers exist in. CC, BE, PT
	Destination	CHPDD
CHP DD	**Definition**	Estimates calling hand by skin color in face adjacent area
	Source	FFDD, SCS
	Input	Coordinate and size of detected face Skin segmented image
	Output	Direction, and distance of caller CC, BE, PT
	Destination	CDI, CDE
CVD	**Definition**	Finds voice direction with voice and noble sound direction
	Source	NSDD, CVD
	Input	Calling voice and Novel sound direction
	Output	Direction of callers CC, BE, PT
	Destination	CDI
CDI	**Definition**	Finds caller direction by putting together sound and image direction
	Source	NSDD, CVD, FFDD, CHPDD
	Input	Direction information and certainties from calling voice and novel sound. Face location and certainty
	Output	Location of callers CC, BE, PT
	Destination	Caller Identification
CDE	**Definition**	Judges the certainty that callers exist
	Source	NSDD, CVD, FFDD, CHPDD
	Input	Certainty of calling voice Certainty of calling hand posture
	Output	CC, BE, PT
	Destination	Caller Identification
CI	**Definition**	Integrates certainties from pre-step modules and identifies callers existence and location
	Source	CDI / CDE
	Input	Certainties of CDI and CDE
	Output	CC, BE, PT
	Destination	N/A

ACKNOWLEDGMENT

This work is supported by a project "Ubiquitous Robotic Companion (URC)," and was funded from the Ministry of Information and Communications (MIC) Republic of Korea.

REFERENCES

[1] K. Dautenhahn and A. Billard, “Bringing up Robots or The Psychology of Socially Intelligent Robots: From Theory to Implementation,” In Proc. of the Autonomous Agent, 1999.

[2] J. Fritsch, M. Kleinehagenbrock, A. Haasch, S. Wrede, and G. Sagerer, “A flexible infrastructure for the development of a robot companion with extensible HRI-capabilities,” IEEE International Conference on Robotics and Automation, Barcelona, Spain, April 2005.

[3] S. Li, M. Kleinehagenbrock, J. Fritsch, B. Wrede, and G. Sagerer, “BIRON, let me show you something: evaluating the interaction with a robot companion,” IEEE International Conference on Systems, Man and Cybernetics, vol.3, pp. 2827-2834, Oct. 2004.

[4] C. Breazeal, D. Buchsbaum, J. Gray, D. Gatenby, and B. Blumgerg, “Learning From and About Others: Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots,” Artificial Life, 2004

[5] H. Huttenrauch and K. Serverinson Eklundh. Fetch-and-carry with CERO: Observations from a long-term user study with a service robot. In Proc. IEEE Int. Workshop on Robot-Human Interactive Communication(ROMAN), pages 158-163. IEEE Press, 2002

[6] H. G. Okuno, K. Nakadai, and H. Kitano, “Social interaction of humanoid robot based on audio-visual tracking,” In Proc. Int. Conf. on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, Cairns, Australia, 2002. Lecture Notes in Artificial Intelligence, Springer.

[7] R. Bischoff and V. Graefe. Demonstrating the Humanoid Robot HERMES at an Exhibition:. A Long-Term Dependability Test. In Proc. IEEE/RSJ Int. Conf. on Intelligence Robots and Systems; Workshop on Robot at Exhibitions, Lausanne, Switzerland, 2002

[8] Marcus Kleinehagenbrock, Jannik Fritsch, Gerhard Sagerer. Supporting Advanced Interaction Capabilities on a Mobile Robot with a Flexible Control System. in Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, vol. 3, Sendai, Japan, 2004, pp. 3649-3655

[9] Designing Sociable Robots, MIT Press, 2005

[10] Kerstin Dautenhahn, “Robots We Like to Live with? A Developmental Perspective on a Personalized, Life-long Robot Companion”

[11] S. Thrun, M. Beetz, M. Bennewiz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hahnel, C. Rosenberg, N. Roy, J. Schulte, D. Schulz. Probabilistic algorithms and the interactive museum tour guide robot Minerva. Intern. Journal of Robotics Research, Vol. 19, NO. 11, pp.972-999

[12] I. Nourbakhsh, J. Bobenage, S. Grange, R. Lutz, R. Mayer, A. Soto. An Affective Mobile Educator with a Full-time Job. Artificial Intelligence, 114(1-2), pp. 95-124

[13] R. Simmons, J. Fernandez, R. Goodwin, S. Koenig, J. O’Sullivan. Xavier: An autonomous mobile robot on the web. Robotics and Automation Magazine 1999.

[14] R. C Arkin, “Behavior-Based Robotics”, MIT Press, Cambridge, MA, 1998