Challenges in Robust Situation Recognition through Information Fusion for Mission Criticial Multi-agent Systems*

Hiroaki Kitano[1,2,3]

[1] ERATO Kitano Symbiotic Systems Project, Japan Science and Technology Corporation, Suite 6A, 6-31-15 jingumae, Shibuya Tokyo 150-0001 Japan
kitano@symbio.jst.go.jp
[2] The Systems Biology Institute
Suite 6A, 6-31-15 jingumae, Shibuya Tokyo 150-0001 Japan
[3] Sony Computer Science Laboratories, Inc.
3-14-13 Higashi-gotanda, Shinagawa, Tokyo 141-0022 Japan

Abstract. The goal of this paper is to highlights one of emergent scientific issues in RoboCup task domains that has broader applications even outside of the RoboCup task domains. This paper particularly focuses on robust recognition through information fusions issue among numbers of others issues that are equally important. The robust recognition through information fusion is selected because it is one of the most universal issues in AI and robotics, and particularly interesting for domains such as soccer and rescue that has high degree of dynamics and uncertainty, as well as being resource bounded. The author wish to provide a conceptual framework on robust perception from single agent to multi-agent teams.

1 Robustness and Information Fusion

The RoboCup project is an interesting research platform because it has multiple domains that have both difference and similarities in basic domain features. Soccer is a dynamic game with real-time collaboration within teammate, but adversarial against opponents. Major uncertainties are generated by (1) opponent strategies, (2) uncertain and stochastic nature of physics on ball movements and other factors that affect course of each action, and (3) uncertainty of real-time behaviors of each player. On the contrary, rescue domain is a dynamic and mission critical domain with high degree of uncertainty and partial information under hostile terrain that are totally different in each case. Major uncertainties are generated by (1) unknown terrain and victim information, (2) uncertain and stochastic nature of incoming information, success of each operations, numbers of external perturbations, and other social and political factors, and (3) limited information on individual victims and their situations.

* This work was partially supported by the Air Force Office of Scientific Research/ Asian Office of Aerospace Research & Development, under the contract number AOARD-03-4035.

D. Polani et al. (Eds.): RoboCup 2003, LNAI 3020, pp. 35–42, 2004.

It is interesting to note that despite differences in the task domain, there are substantial commonalities in structures of uncertainty. There are uncertainty and limited information at the macroscopic level at the level of entire terrain or theater of operation and at the microscopic level which is the scale of individual players or victims. In addition, there are issues of unknown and unpredictable perturbations throughout the operation.

In order to best accomplish the task, a team (or teams) of agents, either robotics or informational agents, need to be robust in perceptions and actions, as well as their team behaviors. It should be well coordinated to reconstruct a model of the terrain and other agents in the scene against various noise, uncertainty and perturbations, so that effective actions can be taken. A set of actions need to be robust so that failure of one or more of such actions do not leads to catastrophic failure of the overall mission.

Robustness of the system is generally exhibited as capability of the system to (1) adapt to environmental fluctuation, (2) insensitivity against fluctuations in system's internal parameters, and (3) graceful degradation of performance, as opposed to catastrophic failure of the system.

This applies from sensory level to higher multi-agent team level. A brief example at the team level shall make clear what does robustness means. For the soccer team, this means that the team should be able to cope with differences in strategy and physical performance of opponent team, not being seriously affected by changes in fatigue and other factors of players, and removal of one or more players does not results in complete loss of its team performance. For the rescue team, it means that the team can cope with various different disaster scenario and dynamical changes in the situation, ability of cope with unexpected fatigue, damage, and resource shortage, and capability to carry out missions even if some of its rescue teams have to be withdrawn from the scene. For the rescue team that has to cope with extremely hostile environment, robustness is one of the essential features of the system.

Robustness of the system is usually attained by (1) the use of proper control, such as negative feedback, (2) redundancy, or overlapping functions of multiple subsystems, (3) modular design, and (4) structural stability.

Extensive research has been made on properties of robustness in biological systems, and these traits were shown to be universal using numbers of examples, including bacteria chemotaxis, robustness of gene transcription against point mutations and noises, stable body segment formation, cell cycle and circadian period, etc [4, 5].

Bacteria, for example, swim toward chemoattractants by sensing graduent of concentration. This capability is maintained regardless of concentration level, steepness of the gradient, and keep track of graduent changes consistently. Integral feedback has been identified as a key intrinsic mechanism in which bacterial behaviors are controlled by activation of receptor complex, but deactivated by a negative feedback loop with integral components. This feedback control enables behavior of bacteria dependent on the level of concentration changes took place, but independent of absolute concentration level of chemical in environment

[1, 10]. Similar mechanisms are observed widely amoung different speceies. Feedback control is only one of several mechanisms behind biological robustness.

On the contrary, artificial systems tend to be less robust, and reply on rigid build-in design that may easily fail under fluctuations. How to build robust systems from sensory-level to strategy-level in a consistent manner is one of the major issues in RoboCup research.

In the rest of the paper, possible research directions for robust systems particularly focusing on information fusion aspects are discussed. Information fusion is raised here because it is relatively universal in different domains, and critical for strategic actions to follow. Issues of robust strategic decisions and executions will be the other robustness issues in multi-agent teams, but will not be discussed due to space limitations.

2 Dimensions of Information Fusion for Robust Systems

Information fusion for robust systems has to be considered for multiple aspects:

Abstraction: An interactive processing of different abstraction levels, such as interactive processing of low-level sensory information and high-level recognition and strategy, enhances robustness by providing interlocking feedback of information thereby hypotheses may converge to more plausible one.

Multiple Sensory Channels: Integration of multiple modal perception channels can contribute to improve robust perception by complementing missing information by perception channels with different characteristics.

Perception-Action Loop: Active involvement of probing actions into perception, that is an integration of perception-action loop to enhance recognition by actively manipulating the environment so that ambiguity can be resolved.

Spatio-Temporal Distribution: Integration of spatio-temporally distributed sensory information, as well as absteacted symbolic information is essential to create overall picture of the situation, thereby robust adaptation to the situation can be done with overall re-evaluation of the situation.

Figure 1 illustrates first three dimensions of information fusion for robust perception.

Information fusion in these aspects contribute robust perception of theater of operation through one or more of four mechanisms of robustness.

3 Interaction of Low-Level and High-Level Perception and Actions

Given the multi-scale nature of the domain that requires identification of situation at both macroscopic and microscopic levels, distributed and coordinated information fusion need to be performed that are ultimately combined to create a coherent model. Information fusion at the macroscopic level is called the high level information fusion (HiLIF) and that of the microscopic level is called sensory level information fusion (SLIF).

In the sequential model of AI, a famous Sense Model Plan Act cycle (SMPA cycle) has been used. This paradigm has been criticized as not being able

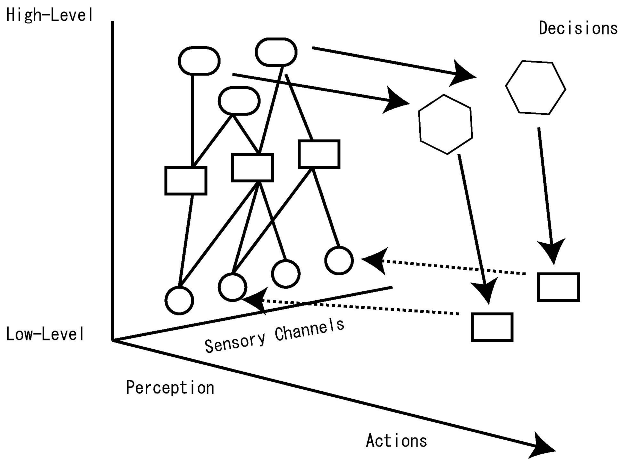

Fig. 1. Integration at Abstraction, Perception channel, and Perception-Action loop

to respond to environment in real-time, and an alternative approach named "behavioral-based approach" has been proposed [3]. While the behavior-based AI demonstrated effectively how simple robotics systems and virtual agents can behave in the real world without creating an internal model, it never scaled to perform complex tasks. In both soccer and disaster rescue, coupling of hierarchy of sensing and actions from low-level sensory units to high-level strategy generators is essential. Reactive control of leg movement to kick a ball must be coordinated with strategic choice of attacking patterns and a pass trajectory to enable such a strategy. Similarly, local actions and decision for the recovery of some critical life-lines has to be carefully coordinated with overall strategy. In this context, low-level perception and action module is not merely behavior-based module. It must be able to recognize local situation that can be aggrigated at a higher level. SLIF should have a certain level of internal model.

While these examples are how overall strategy may constrain local actions, there are cases local actions influence overall strategy. Basically, it is interaction of bottom-up and to-down processes in which the architecture enabling it has been long standing theme in multi-agent systems. While this is well recognized issue, the author would not make further discussions than stating that RoboCup task domains, particularly humanoid league and rescue, are one of ideal platform for seriously tacking this problem. This aspect of integration essentially exploits feedback control of the system for adaptation at certain abstract levels. Low-level perceptions and local decisions are aggrigated to higher-level that feedback differences between desired local situations and actions and actual situation and actions to reduce such descrepancies.

4 Information Fusion of Multiple Sensory Channels

Navigation, localization, and object identification of robotic agents tends to rely on visual information. While vision provides rich information, it is limited in several aspects. First, visual perception can be occluded by obstacles, including dust and smokes. Second, it has to have certain quality of light sources. These features impose serious limitations of visual perception for disaster rescue situation, because disaster scenes are generally highly unstructured, dusty, and may have serious smokes from fire. Lights are often not available in confined environment where victims may be embedded. Auditory perception, on the other hands, has different characteristics. It can transmit over obstacles, do not require light sources. In fact, various noises victims may make is a critically important signals for a victim location identification. Other sensory channels, such as odorant, CO2, and vibrations have different characteristics that complement each other. Information of multiple modalities of perceptions may provide high level of robustness in perceiving the environment. Some of early efforts have been done using integration of auditory and visual perception [6–9]. Vision system often generates false positive identification of objects that is supposed to recognize due to similarity of color and shape in irrelevant objects. When the object is creating certain auditory signals, the use of auditory information to track the sound stream can effectively eliminates false positives. By the same token, objects could be occluded by obstacles so that vision system lost tracking, which can be compensated by keep tracking auditory signals if the object is making some sound streams. Early experiments indicate that well designed integration of multiple modal perception channels based on multiple sensory steam integration is effective for robust perception. Now, the research topic shall be to create basic principle of robust object identification, tracking, and scene understanding by using multiple perception channels, most likely by defining a dynamical defined invariance that corresponds to each object as its signature.This approach is expected to attain robustness by exploring redundancy, or overlapping functions, of perception channels, so that degradation or failure of one perception channel is well compensated by other perception channels with overlapping functions.

5 Integrating Actions to Perception

It is important that the concept of active perception to be integrated into the system, so that ambiguities of information can be resolved by actively directing sensory devices and information gathering agents. Early research of this direction has been proposed as active vision and animated vision [2], but it has to be extended to include not only vision and other perception channels, but also to more high level information collections. Conceptually, this is feedback control to minimize unknown, or ambiguous part of scene to zero.

In the resource constraint situation such as in disaster rescue, the concept of cost of active probing has to be introduced. Any action to obtain information is associated with cost, the use of resources, including time. Decision has to be made on whether actively probe new information or to proceed with ambiguities.

Suppose that a ball that is rolling from right to left is occluded by an opponent player, decision has to be made to use predictions of ball trajectory or actively probe the ball position. While actively proving the ball position may resolve ambiguity of information, it may loose time window to intercept the ball.

By the same token, in the disaster scenario, spreading of fire or cascading collapse of buildings may not be fully monitored by available sensors or information agents. A tactical decision has to be made to dispatch a unit to counter such incidence by predicting possible front of chain reactions or to mobilize information gathering agents to make sure the status of incidents. The cost of information gathering is the use of additional agents and time to wait until the situation to be disambiguated. On the contrary, a quick deployment of counteraction units runs a risk that the prediction is false and deployments are deemed ineffective. Always, there is a trade-off between cost of knowing and risk of not knowing. One of the research topics may be to find out principles of decision on cost of knowing versus risk of not knowing.

6 Spatio-temporal Integration

6.1 Heuristics and Knowledge-Based Estimation

Integration of spatio-temporal information is essential, particularly for disaster rescue operations. Theater of operation is widely distributed, and information is collected only at limited rate from limited locations. Basically it is a problem of making the best estimate of the situation by sparse sampling of complex terrain in 4-D (XYZ+T) space. Certain heuristics, such as continuity of the unfolding events, and a priori on structure of urban infrastructure are expected to be highly useful to constrain possible hypotheses. This applies to both low-level and high-level information fusion, but particularly useful for high-level information fusion where "fog of war" has to be resolved as soon as possible with limited resources. Advantage of this approach is that you can make reasoned estimate of the situation even for the area that cannot be directly measured. The drawback is that it requires substantial knowledge of the urban structures in usable form that are generally not available. At the same time, how to increasing sampling points is the other big issue. One of the best ways is to make sure sensory systems are ubiquitously present in disaster scene. This can be achieve only by creating multi-functional systems that are useful in daily life, but can act as sensory units in emergency. Traffic signals and various monitoring cameras are possible resources that can cover public space. Home security systems, home entertainment robots, and a series of home electronic products are ideal for covering situations in each household. However, there are issues of securing telecommunication with such devices, as well as protection of privacy that are critical, but are outside of AI and robotics issues.

6.2 Adaptive Airborne Spatial Information Fusion

One of possible approach to solve this problem is to develop a small disposable sensory unit and deploy them in large numbers. Figure 2 illustrates one example of such systems which could be deployed airborne. The goal is quickly understand target terrain/city situation for the purpose of disaster rescue.

Phase-I: Large number of small IF (information fusion) devices will be deployed mid-air

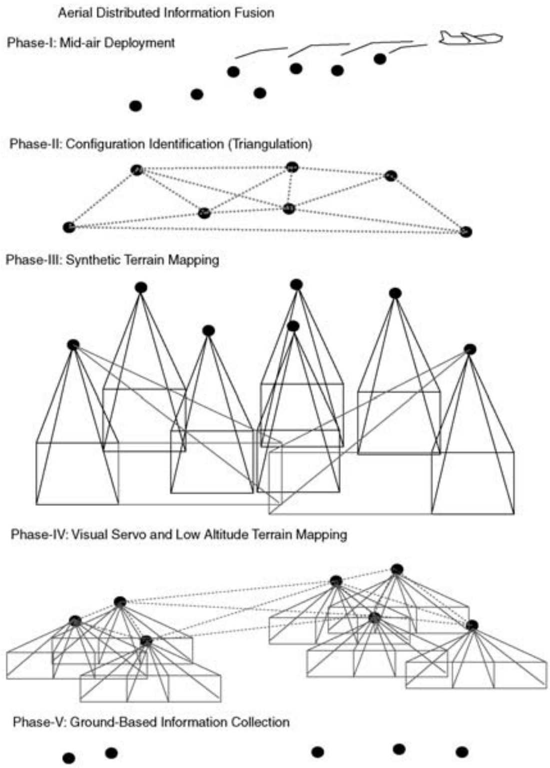

Fig. 2. Aerial Deployment of Small Information Fusion Units for Disaster Area Scanning

Phase-II: Speed break is activated and speed is stabilized. Then each unit has some visual beacon or visual ID so that relation location of units can be determined by triangulation.

Phase-III: First aerial photo/video will be taken using multiple camera unit, but mostly using cameras that are facing terrains below. and send back to airborne server or other servers. Focus of attention is determined and rough model will be created. If processing power is sufficient, this can be done in each unit. Infra-red sensors or other sensors can be used in addition to CMOS imager to identify specific signature on terrain.

Phase-IV: Each unit deploys small and simple flaps to control trajectory, so that visual servo will be made effective to regroup IF units, so that areas that are more significant will be assigned with dense IF units. Photo/video analysis and reconstruction continues. In low altitude, all 360 angle camera as well as microphone and other sensors are activated to capture maximum information on the ground. If possible, frame rate will be increase dramatically.

Phase-V: For those units that safely reached the ground and survived impact, 360 degree vision and microphone systems, as well as other sensors will be activated to identify objects and terrain structure.

This approach integrates self-organization of agents and sensory feedback on-the-fly, and can be enormously useful for rapid deployment at disaster site, as well as being an excellent test of modern AI techniques.

7 Conclusion

This paper addressed several issues in robust systems for RoboCup domains. Several aspects of robustness have been identified, and four aspects of information fusion have been discussed briefly. Both soccer and rescue provides an excellent platform for such research, and several research issues have been raised. Integration of three dimensions (abstraction levels, perception modality, and perception-action loops) of information fusion poses particularly interesting problems that community shall tackle. Spatio-temporal integration applies to both soccer and rescue, but more seriously to rescue scenario. Resolving this issue requires not only improvement of each robotic and AI agents, but also how such systems are deployed before and after the onset of the disaster. A systematic approach for various levels of information fusion is necessary for robust perception of multi-agent teams.

References

1. Alon, U., Surette, M., Barkai, N., and Leibler, S., (1999) "Robustness in bacterial chemotaxis," *Nature,* 397(6715):168-71
2. Ballard, D., (1989) "Reference Frames for Animate Vision," *Proc. of International Joint Conference on Artificial Intelligence* Detroit.
3. Brooks, R., (1986) "A robust layered control system for a mobile robot," *IEEE J. robotics and Automation,* RA-2, 14-23.
4. Kitano, H.,(2002) "Systems biology: a brief overview, *Science,* 295(5560):1662-1664
5. Kitano, H., (2002) "Computational systems biology, *Nature,* 420(6912):206-210
6. Nakagawa, Y., Okuno, H.G., and Kitano, H. (1999) "Using Vision to Improve Sound Source Separation," *Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI-1999),* pp.768-775, AAAI, Orlando, Jul. 1999.
7. Nakadai, K., Lourens, T., Okuno, H.G., and Kitano, H. (2000) "Active Audition for Humanoid," *Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI-2000),* 832-839, AAAI, Austin, Aug. 2000.
8. Nakadai, K., Hidai, K., Mizoguchi, H., Okuno, H.G., and Kitano, H. (2001) "Real-Time Auditory and Visual Multiple-Object Tracking for Robots," *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-01),* IJCAI and AAAI, Vol.2, 1425-1432, Seattle, Aug. 2001.
9. Nakadai, K., Hidai, K., Okuno, H.G., and Kitano, H. (2002) "Real-Time Speaker Localization and Speech Separation by Audio-Visual Integration," *Proceedings of IEEE/RSJ International Conference on Robots and Automation (ICRA-2002),* IEEE, pp.1043-1049, Washington DC., May 2002.
10. Yi, T., Huang, Y., Simon, M., and Doyle, J., (2000) "Robust perfect adaptation in bacterial chemotaxis through integral feedback control," *Proc. Natl Acad Sci USA,* 97(9):4649-4653