An Event-driven Context Model in Elderly Health Monitoring

Yuanyuan Cao, Linmi Tao, Guangyou Xu
Tsinghua National Laboratory for Information Science and Technology (TNList), Tsinghua Univ.
Beijing, China
{caoyy1205, tao.linmi, xugy.dcs}@gmail.com

***Abstract*—Elderly health-monitoring systems provide various proactive services according to context. Context plays an essential role in semantic understanding of human activities from sensor data. We present a novel event-driven context model for elderly health-monitoring application, which supports a distributed system with multimodal sensors. The context model is a dynamic hierarchical structure modeling human actions, activities, and environment information in home scenario. Context awareness is achieved by detecting events associated to the context hierarchy. Common knowledge ontology is employed to guide high-level context reasoning. Experiments of this method in elderly health monitoring show the effectiveness of the proposed context model.**

Keywords-elderly monitoring; context model; context aware

I. INTRODUCTION

In the passed decades, the unprecedented scientific progress has continuously pushed the human society into Global Aging Time. This progress has revolutionized how long and how well we live. These advances, in turn, have generated increasing optimism and hope, that most aged people can live independently in their homes and continue participating in community life. Although the scientific discoveries make the majority of elderly people live healthy and active, it is still dangerous for elderly people to live independently when they are getting elder and elder, because of both lonely feeling and physiological diseases. Supporting independent living for elderly people is essential not only for the elderly people themselves, but also for the social and health care sectors. Science offers the best hope to improve the elderly person's quality of life and to help them not only survive but thrive in independently living as long as possible. This will help reduce seniors' health care costs by preventing depression and improving fitness. Furthermore, elderly people can be an important resource for meeting social service needs and overcoming workforce shortages if they still stay active as they age. For this reason, there exists a need for unobtrusive health-monitoring solutions enabling independent living of the elderly adults.

The governments as well as industry show a great interest to develop health-monitoring systems for the seniors. These systems attend to extend the time the seniors can live alone in their home by providing supervision and assistance to help ensure their health, safety, and well-being. Health-monitoring for elderly people in home environments can be accomplished by: 1) ambulatory monitors that utilize wearable sensors to record physiological and motion signals[1]; 2) sensors embedded in the home environment and furnishings to unobtrusively collect behavioral and physiological data[2,3,4]; or 3) a combination of the two [5]. Currently, many research organizations focus on passive non-intrusive monitoring systems aiming to make the elderly people feel comfort in a nature way.

Context plays a significant role in health-monitoring for elderly people due to two main reasons. Firstly, the semantic meaning of human activities in terms of life signals highly depends on the context. Especially in daily living scenarios, activities have to be understood in the context of overall environment. Secondly, signals processed in this domain are often multimodal, including data collected from various sensors, such as cameras, RFIDs, accelerometers, microphones, and pressure sensors. Sensor data processing and fusion demand huge amount of computation. Therefore, selectivity in sensor data processing for semantic analysis is the core of health monitoring systems, which relies on the context [6]. Dey [7] reviews definitions of context, and provides a definition of context as "any information that can be used to characterize situation". This is the sense in which we use the term context. Context specifies the elements that must be observed to model situation. The challenge of context aware is how to capture, process, and exploit concurrent context to provide situated services to users in a human centered system, especially for health monitoring.

Although context aware has been a hot topic in many fields, few related works have been conducted for human centered computing. Zhu [8] presented a Hierarchical Hidden Markov Model based algorithm for the recognition of human hand gestures which considers the context information in the gesture sequence analysis. Hua Si [9] described a context-aware reminder system, called CoReDa, which learns a user's routine of ADLs (Activity in Daily Living), regarded as context, and helps the user complete different activities instead of caregivers. In [10] a distributed system was presented for dynamic selection of the best perspective camera in a multi-camera intelligent space, where context means changes of user's locations and face orientations. Choi [11] developed a system providing knowledge-based assistance according to the contexts, which is recognized by means of continuously monitoring the current position and the time-schedule for their medications. In the works above, context is merely referred to as constraints or conditions for guiding low level processing. Some work has further considered context as a "dynamic structure" [12]. Crowley introduced a conceptual framework to analyze activities within a meeting or lecture [13] and video surveillance [14], where the context model is represented as a network of situations which is defined by a

978-0-7695-3737-5/09 $25.00 © 2009 IEEE
DOI 10.1109/UIC-ATC.2009.47

set of roles and a set of relations between entities playing roles. Our previous work on context model is presented in [15] for meeting scenario, where the users are restricted in fixed location in a single meeting scenario. The system's task is to infer three situations: "break", "discussion", and "presentation". As for the health-monitoring of seniors, the user can be in any place of a house, resulting in changing scenes. Besides, the user engages in many activities and interacts with environment in daily living; a predefined situation network is inadequate. Therefore, it is a challenge to model context for activity analysis.

In this paper, a novel event-driven context model, which support distributed systems with multimodal sensors, is proposed for health-monitoring of seniors. The system analyze the user's activities in context, which is a dynamic hierarchical structure generally consisting of "human context" indicating human status in terms of daily living, and "environment context" indicating the information of the scene where the seniors live. Two kinds of events, namely body events and environmental events, are defined and associated with context hierarchy. Context awareness is achieved through event detection. Common knowledge is structured by ontology for high-level context reasoning and low-level processing guidance. Comprehensible textual descriptions of human activities are designed in the experiment, which shows effectiveness of our framework.

The rest of the paper is organized as follows. In Section 2, the event-driven context model and common knowledge ontology are described. Section 3 presents the context reasoning and demonstrates implementation of the proposed conceptual framework in application of elderly health-monitoring. Conclusions are drawn in Section 4.

II. CONTEXT MODEL AND COMMON KNOWLEDGE ONTOLOGY

In the vision of intelligent systems for health monitoring, multiple devices are diffused into the living environment of elderly people. Relatively, a set of challenging issues is created concerning the interaction between human and system: How to bridge the semantic gap in understanding human physical and mental status via observable body expressions in living environment? How to design implicit interaction for sensor based systems? Context plays a key role in such interaction systems, where the context model deals with these problems by acquiring the user's implicit input from the real world based on building human models, including understanding certain activities. Considering the system tasks, context model defines actions and activities system needs to focus on. Common knowledge ontology which formalizes human knowledge into entities and attributes concerning daily living is also introduced.

A. Event-driven context model

The context model is structured as a hierarchy in two abstract levels; the overview of the model is shown in Fig.1. Two components - "*human context*" and "*environment context*" compose the context model. In low level, the former defines human states by basic body gestures, such as "*walking*", "*standing*", "*sitting*", and "*lying*"; while the latter defines entities' states which are concerned by the system to analyze activities. The high level in the model concerns activities the user is engaged in and emergencies, which are not recognized directly from body gestures but inferred from the human action as well as environment information. This level represents overall human states that the system needs to understand in order to provide situated services, which requires high level semantics. Several basic activities are considered: "*watching TV*", "*drinking water*", "*cooking*", "*sleeping*", and "*falling down*" particularly. Actually, one of the major risks incurred by the fragile population (elderly, illness, etc) is to fall, which is responsible for 70% of accidental deaths in persons ages 75+.

Two kinds of events are defined in compliance with the context hierarchy to form an interweaved context-event hierarchy: "*body event*" and "*environment event*". The former refers to actions causing changes on body gestures. As shown in Fig.1, the body event "*lie down*" causes a shift from "*standing*" to "*lying*". The latter refers to the changes on objects' status in environment, which primarily concerns the results of interaction between user and surrounding environment, such as object use, and household appliances manipulation. Events in our system indicate context switching, which can be regarded as triggers of context transfer.

Context awareness, which bridges the semantic gap between signal processing and understanding of human behavior, is generated from the context model via event detection, which is the observation of the system. Multimodal signal processing modules are integrated to detect human actions and environment events; and then high-level context can be reasoned based on explicit low-level context.

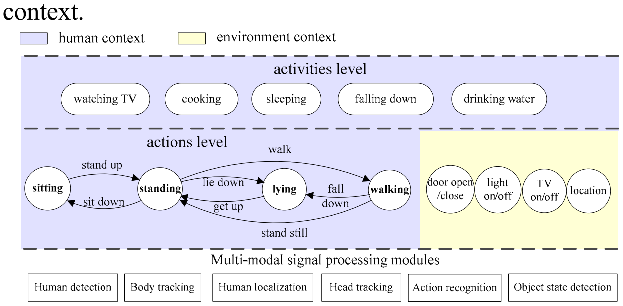

Figure 1. The event-driven context model for elderly health-monitoring

B. Common knowledge ontology

Ontology refers to the subject of existence, which is good at knowledge sharing, logic inference, and knowledge reuse. A specific ontology for health monitoring domain is shown in Fig.2. Several categories of knowledge are considered: objects in the scene and their attributes; the user's personal information, such as medical history, disease and symptoms etc.; camera parameters and planning strategy; signal processor selection guidelines; and proactive services provided by system. The knowledge are formalized and organized as a set of entities, relations, functions, and instances. The upper class includes several categories concerned by the system: "*objects*", "*human*", "*location*", "*activity*", and "*services*". Besides general classes, a number of concrete sub-classes are defined to model specific entities (e.g., the abstract class Objects of home domain are classified into three sub-classes Appliances, Cameras, and Furniture). All classes have a group of attributes of their own.

Common knowledge plays an important role in guiding signal processing, activity reasoning, and comprehensible descriptions for human, which guarantees accuracy and efficiency of the system and strengthens the intelligence of the system to some extent. For example, given the human location, by checking the cameras' attribute "*view field*" the system can choose the best observing camera. The attribute "*service*" to "*activity*" defines what kind of service the system provides for each activity. For instance, activity "*falling down*" will trigger the service "*alarm*". By employing ontology, human knowledge is well formalized and utilized in the system.

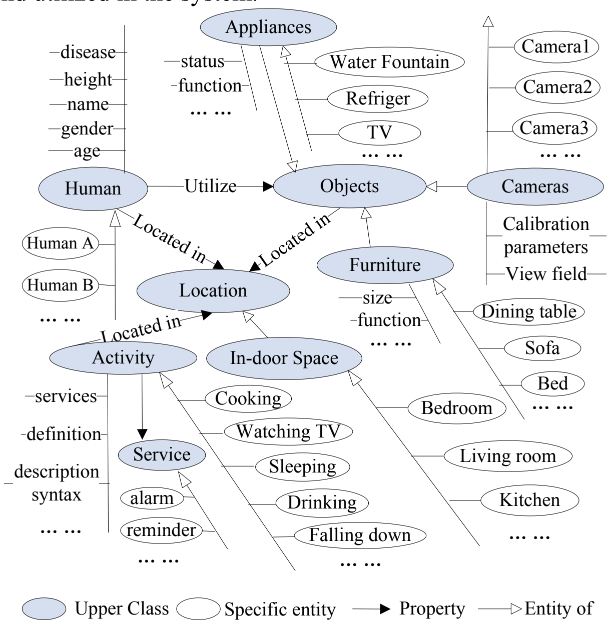

Figure 2. Partial view of common knowledge ontology for elderly health-monitoring

III. CONTEXT REASONING

Context reasoning is to deduce high-level, implicit context from low-level, explicit context based on observation of events. Obviously, high-level context cannot be directly acquired from sensors; it is reasoned from sensor-driven, low-level context such as actions and environmental information. The implementation is presented in this section in detail.

A. *Research Platform*

Our previously developed system for distributed sensor network based multi-streaming data collection and information processing is adopted as the research platform [16]. The architecture is shown in Fig.3. The platform is composed by a set of servers which collaborate with each other to accomplish the tasks like video capture, transmission, buffering and synchronization. The platform is flexible, which is able to support the development of various distributed sensor network based application systems.

Two rooms are set up with furniture and home appliances, which is a simulated home environment with a living room and a kitchen. Five fixed cameras are deployed to monitor the scene from distinct perspectives. Accelerometers are placed on a cup and a TV remote control, which on place on the desk, and the room door respectively. As for the network and software infrastructure, a flexible multi-server platform has been developed to support distributed multimedia information processing.

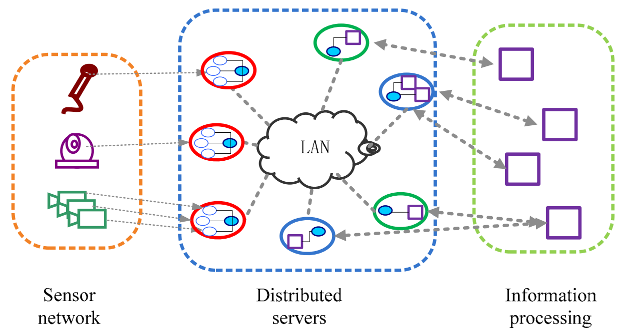

Figure 3. The architecture of the system

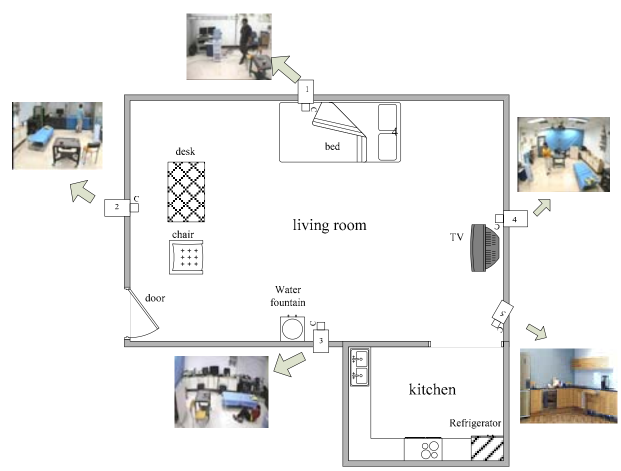

Figure 4. Environment settings of our system

B. *Implementation*

A flowchart of the proposed method is shown in Fig. 5. Context reasoning procedure consists of two major steps: (1) primary processing of multimodal signals, such as foreground extraction, human localization etc.; (2) human action recognition and environmental events detection; (3) high-level context reasoning and description. The detailed processing algorithms are presented as follows.

(1) Signals coming from distributed sensors are processed in the first step, which mainly include foreground extraction and human tracking. Hu's method is employed to locate human [17]. Location is one of the most fundamental contexts for capturing information for high-level context

reasoning. For example, location decides observing cameras and give clues of which object may be used. That means the system does not process all the signals from multimodal sensors, but choose some of that.

(2) Action recognition is challenging because the user moves in an arbitrary manner in the home resulting in changeable location and viewpoint. A viewpoint insensitive action recognition method proposed in our previous work [18] is employed to recognize basic actions, including "*walking*", "*sitting*", "*lying*", and "*standing*". This method requires two cameras with optical axes being orthogonal approximately, which can be any paired neighboring cameras in the scene. To recognize a sequence of frames as an action, body shape is represented by "envelope shape" in every time step, based on which features are extracted from "envelope shape" and form a feature vector; The sequence of feature vectors is then input into a bank of HMMs for action recognition. The action recognition algorithm is demonstrated in Fig. 6. Therefore human context in action level is specified.

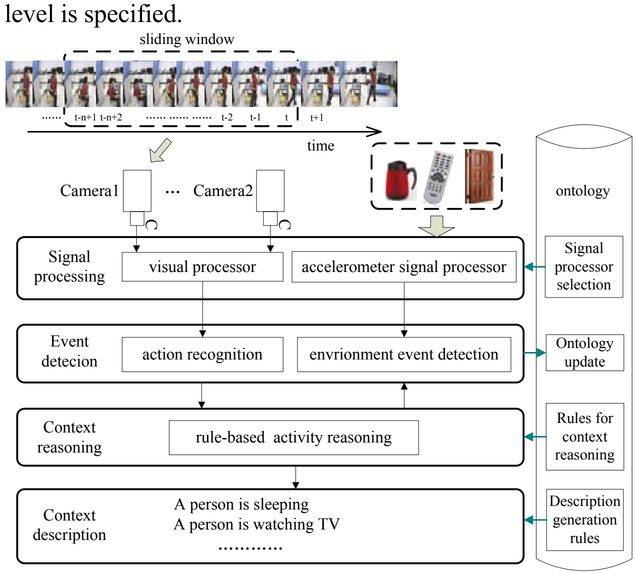

Figure 5. Flowchart of the proposed context reasoning

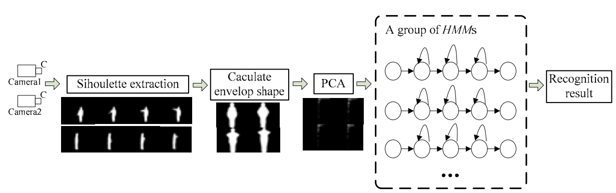

Figure 6. Action recognition flow Diagram

Accelerometers on the door, the cup, and the TV remote control send signals when they are moved. Not all of them are processed at every time step, instead, only when the object is in the operational region of the user, the signals from this object is processed. This selectivity mechanism increase efficiency of the system. In signal processing, we do not analyze the object's moving trajectory and velocity but only detect whether the object is moved or not. Accelerometers help avoid precise and detailed hand movement analysis, which is a big challenge in visual processing, and help detect interaction between human and environment. Detected environmental events include: "*TV turned on*", "*TV turned off*", "*door open*", "*TV remote control used*", and "*cup used*". The TV screen is monitored by one of the cameras and a visual processor detects whether the TV is on by judging the lightness of screen area. The other three events are detected based on signals from the accelerometers.

(3) We define reasoning rules within the entailment of first-order predicate, a wide range of higher-level, conceptual context such as "what the user is doing" can be deduced from relevant low-level context. Table 1 shows the defined context reasoning rules that are employed to derive the user's activities in the home life scenario.

The procedure of context reasoning from low-level to high-level specifies explicit context and acquires overall context, based on which the system automatically generates textual descriptions for every frame. Textual descriptions act as video annotation, which provides possibility to video retrieval and user activity pattern study. Descriptions are generated as follows:

[SUB: subject, PRED: action, LOC: location, …]

In the system, since it is assumed that there is only one user in the current stage, so that the identity of user is default. Thus agentive case always indicates a constant expression "the user". As for instantaneous actions, an action itself and the resultant state are distinguished. For example, "sit down" is an action and "be sitting" is the resultant state after that. We explicitly specify a resultant state in description. As for the locus, text translation adds a preposition before locus, which is selected from a vocabulary according to the spatial relationship between subject and location. For example, "in front of the desk", "in the chair", and "in the kitchen". In most cases, we assume that high-level context may be suitable for expression; when an activity is reasoned, the activity is used in the textual description: e.g. "a person is cooking".

TABLE I. DEFINED CONTEXT REASONING RULES

Context	Reasoning rules
Watching TV	((action SIT) ∨ (object used TV REMOTE CONTROL))∧(locatedIn CHAIR LIVING ROOM)∧ (TV status ON) =>(WATCHING TV)
Sleeping	(action LIE) ∧ (locatedIn BED LIVING ROOM) ∧ (living room light OFF) =>(SLEEPING)
Falling down	(action LIE)∧((locatedIn FLOOR LIVING ROOM)∨ (locatedIn FLOOR KITCHEN)) =>(FALLING DOWN)
Drinking water	(object used CUP)∧(locatedIn DESK LIVING ROOM) =>(DRINKING WATER)
Cooking	(action STAND)∧(locatedIn CUPBOARD KITCHEN) =>(COOKING)

C. *Results and discussion*

We performed experiments on 4 different scenarios acted by 3 different actors. In each scene, typically, a person sleeps, falls down, walks, watches TV, and drinks in unrestricted order. Fig.7 shows a sequence of video frames with textual descriptions. Further, descriptions for human understanding still need well defined syntax for accurate and natural expression.

frame 0087	the user is walking in the living room
frame 0161	the user is sitting in the chair
frame 0195	the user is watching TV
frame 1624	the user is drinking some water
frame 2064	the user is standing in front of the desk
frame 2120	the user is walking in living room
frame 2151	the user is falling down
frame 2602	the user is sleeping

Figure 7. Input images and the generated text.

IV. CONCLUSIONS

Context plays a significant role in human centered computing. However, context modeling and awareness still remains a challenging problem, especially in daily living scenarios, due to the dynamic nature and complexity of context. Current context cannot be determined by simple combination of sensor data, on the contrary it has to be inferred by a multi-level processing model. In this paper, we address the problem by analyzing human activities in home health monitoring system. A novel event-driven hierarchical context model is proposed to model human activities in daily living. Awareness of the context is based on the detection of events associated to the context hierarchy, which serves as the conceptual framework for context-aware computing. Event detection is further presented to enable context aware computing, based on which context reasoning is achieved. It is characterized as a great advantage of our approach. Experiments show the effectiveness of the proposed model.

Future work includes extension of multimodal sensors and improvement of inference. Proactive services are also to be integrated in the system.

ACKNOWLEDGMENT

The work described in this paper was supported by National Science Foundation of China (No. 60673189) and National Science Foundation of China (No. 60873266).

REFERENCES

[1] Zhu, C., Sun, W., Sheng, W.: Wearable Sensors based Human Intention Recognition in Smart Assisted Living Systems. In: Proc. of the 2008 IEEE Intl. Conf. on Information and Automation, pp. 954-959. IEEE Press, Zhangjiajie, China (2008)

[2] Virone, G., Alwan, M., Dalal, S., Steven, W., et al.: Behavioral Patterns of Older Adults in Assisted Living. IEEE Transaction on Information Technology in Biomedicine. Vol. 12, No. 3, pp. 387-398(2008)

[3] Litz, L., Gross, M.: Covering Assisted Living Key Areas based on Home Automation Sensors. In: Proc. of the 2007 IEEE Intl. Conf. on Networking, Sensing and Control, pp. 15-17. IEEE Press, London, UK(2007)

[4] Fleck, S., StraQer, W.: Smart Camera Based Monitoring System and Its Application to Assisted Living. In: Proc. of the IEEE, Vol. 96, No. 10, pp. 1698-1714 (2008)

[5] Korhonen, I., Parkka, J., Van Gils, M.: Health Monitoring in the Home of the Future. IEEE Engineering in Medicine and Biology Magazine, Vol. 22, Issue 3, pp.66-73(2003)

[6] Pantic, M., Pentland, A., Nijholt, A., Huang, T.: Human Computing and Machine Understanding of Human Behavior: A Survey. In Proc. of Int. Conf. Multimodal Interfaces, pp. 239-248(2006)

[7] Dey, A.: Understanding and Using Context. Personal and Ubiquitous Computing, Vol. 5, No. 1, pp 4-7(2001)

[8] Zhu, C., Cheng, Q., Sheng, W.: Human Intention Recognition in Smart Assisted Living Systems Using A Hierarchical Hidden Markov Model. In: Proc. of 4th IEEE Conf. on Automation Science and Engineering, pp.23-26. IEEE Press, Washington DC.(2008)

[9] Hua, S., Seung, J., Kawanishi, N., Morikawa, H., Kawanishi, N., Morikawa, H.: A Context-aware Reminding System for Daily Activities of Dementia Patients. In: Proc. of 27th Intl. Conf. on Distributed Computing Systems Workshops(2007)

[10] Trivedi, M., Huang, K., Mikic, I.: Dynamic Context Capture and Distributed Video Arrays for Intelligent Spaces. IEEE Trans. Syst., Man, Cybern. A, Syst. Humans, vol. 35, no. 1, pp. 145-163, (2005)

[11] Choi, J., Lim, M., Kim, D., Park, S.: Proactive Medication Assistances based on Spatiotemporal Context Awareness of Aged Persons. 30th Annual International IEEE EMBS Conference. Vancouver, British Columbia, Canada(2008)

[12] Greenberg, S.: Context as a Dynamic Construct. Human-Computer Interaction, vol. 16, pp. 257-268(2001)

[13] Crowley, J.: Situated Observation of Human Activity. Computer Vision for Interactive and Intelligent Environments, Vol. 2005, pp.97-108(2005)

[14] Brdiczka, O., Yuen, P., Zaidenberg, S., Reignier, P., Crowley, J.: Automatic Acquisition of Context Models and Its Application to Video Surveillance. In: Proc. Int. Conf. Pattern Recognition, pp.1175-1178(2006)

[15] Dai, P., Di, H., Dong, L., Tao, L., Xu, G.: Audio-visual Fused Online Context Analysis toward Smart Meeting Room. In: Proc. 4th Intl. Conf. on Ubiquitous Intelligence and Computing (2007)

[16] Wang, Y., Tao, L., Liu, Q., Zhao, Y., Xu, G.: A Flexible Multi-Server Platform for Distributed Video Information Processing. In: Proc. 5th Intl. Conf. on Computer Vision Systems (2007)

[17] Hu, W., Zhou, X.: Principal Axis-based Correspondence Between Multiple Cameras for People Tracking. IEEE Trans on Pattern Analysis and Machine Intelligence, 28(4), pp.663-671(2006)

[18] Huang F., Xu G.: Viewpoint Insensitive Action Recognition Using Envelop Shape. In: Proc. 8th Asian Conference on Computer Vision, pp.477-486, Tokyo, Japan (2007)