International Journal of Innovative Computing, Information and Control
Volume **3**, Number **6(A)**, December **2007**

ICIC International ©2007 ISSN 1349-4198
pp. **1359–1368**

DYNAMIC RECOGNITION OF HUMAN ACTIONS AND OBJECTS USING DUAL HIERARCHICAL MODELS

Masahiro Saitou[†*], Atsuhiro Kojima[†], Tadahiro Kitahashi[‡] and Kunio Fukunaga[†]

†Osaka Prefecture University
1-1, Gakuen-cho, Nakaku, Sakai, Osaka 599-8531, Japan
saitou@com.cs.osakafu-u.ac.jp; ark@las.osakafu-u.ac.jp; fukunaga@cs.osakafu-u.ac.jp
*Present affiliation: Panasonic Shikoku Electronics Co., Ltd.

‡Kwansei Gakuin University
2-1, Gakuen, Sanda Hyogo 669-1337, Japan
kt@ksc.kwansei.ac.jp

Received December 2006; revised April 2007

Abstract. *In the field of image understanding, it is becoming popular to recognize human actions and objects in a scene. Recently, various methods for scene recognition have been proposed, and most of them focus on recognizing human actions and objects in different ways. We consider, however, actions and objects can be recognized complementarily, because human actions are closely related with objects. In this paper, human actions and related objects are represented in dual hierarchical models and recognized by making use of the relationship between them. First, movement of human head and hands are tracked by stereo vision. Features of actions such as position and direction of the head and hands are extracted and input to Dynamic Bayesian Networks to classify actions roughly. Then the actions and related objects are refined by using conceptual models of human actions and objects. Finally, detailed actions and objects are recognized by referring each model cooperatively.*

Keywords: Cooperative recognition, Human action recognition, Object recognition, Dynamic Bayesian network

1. **Introduction.** The recent trend in the field of computer vision could be recognition of human actions and objects in a scene to realize human-computer interaction. Conventionally, recognition of objects and human actions have been thought to be different. In the field of object recognition, in particular, objects have usually been represented in appearance-based models such as shapes, colors and textures.

Many humans actions, however, are closely related with objects [1]. For example, when a person sits down on a chair, the chair supports him/her upward. Like this, to support human is the intrinsic function of chairs. Humans can infer that there is a chair when he/she sees someone bend the waist on uncertain objects. In the field of cognitive science, it is experimentally confirmed that relationship of action and object plays an important role in the human vision system [2]. Thus, we intend to recognize usage of objects by means of inferring indirectly from observation of human actions interacting with objects. Consequently, unlike the conventional recognition, objects are identified as rather objective of behaviors than categorical labels.