ABSTRACT

Traditionally, how to bridge the gap between low-level visual features and high-level semantic concepts has been a tough task for researchers. In this article, we propose a novel plausible model, namely cellular Bayesian networks (CBNs), to model the process of visual perception. The new model takes advantage of both the low-level visual features, such as colors, textures, and shapes, of target objects and the inter-relationship between the known objects, and integrates them into a Bayesian framework, which possesses both firm theoretical foundation and wide practical applications. The novel model successfully overcomes some weakness of traditional Bayesian Network (BN), which prohibits BN being applied to large-scale cognitive problem. The experimental simulation also demonstrates that the CBNs model outperforms purely Bottom-up strategy 6% or more in the task of shape recognition. Finally, although the CBNs model is designed for visual perception, it has great potential to be applied to other areas as well.

Keywords: Bayesian model; cellular Bayesian networks (CBN)

INTRODUCTION

Cognitive informatics (CI) is a transdisciplinary enquiry of cognitive and information sciences that investigates the internal information processing mechanisms and processes of the brain and natural intelligence (Wang, 2007). It covers a wide range of research fields, including the information-matter-energy (IME) model (Wang, 2003b), the layered reference model of the brain (LRMB) (Wang, Wang, Patel & Patel, 2006), the object-attribute-relation (OAR) model of information representation in the brain (Wang, 2006d; Wang & Wang, 2006), the cognitive informatics model of the brain (Wang, Liu, & Wang, 2003; Wang & Wang, 2006), natural intelligence (NI) (Wang, 2003b), autonomic computing (AC) (Wang, 2004), neural informatics (Nel) (Wang, 2002, 2003b, 2006a), CI laws of software (Wang, 2006b), the mechanisms of human perception processes (Wang, 2005a), the cognitive processes of formal inferences (Wang, 2005b), and the formal knowledge system (Wang, 2006c). Of all these branches,
perception, as an interesting research field of CI, mainly focuses on how human beings perceive external world. Researchers have proposed an excellent model, the motivation/attitude-driven behavioral (MADB) model (Wang & Wang, 2006), to formally and quantitatively describe the relationship between the internal emotion, motivation, attitude, and the embodied external behaviors. In this article, we limit our work to visual perception, and propose a connectivity-based model to formally mimic the perceptual function of human beings.

The primary task of visual perceptual is to organize the visual features of an image into some already known objects. Yet, how to bridge the gap between low-level visual features and high-level semantic concept has long been a tough problem, which puzzles researchers all along. Until now, most of the proposed algorithms just focus on some particular objects, such as human faces, cars, people, and so forth (see e.g., Papageorgiou & Poggio, 2000; Schneiderman & Kanade, 2000; Tamminen & Lampinen, 2003). Researchers utilize various schemes (see e.g., Broek et. al., 2005; Lai, Chang, Chang, Cheng, & Crandell, 2002; Tamminen & Lampinen, 2003) to integrate the low-level visual features, including colors, textures and shapes into the profiles of target objects. Although some people (Murphy, Torralba, & Freeman, 2004) exploit background, or scene, information to improve the recognition, they do not take advantage of interrelationships between objects to help the identification process.

As a matter of fact, interrelationships between objects are of great importance to perceptual organization. Researchers (Geisler & Murray, 2002) have shown that there might exist some visual patterns in human brains, which enable human beings to recognize some simple objects, for example, English letters, as soon as they see them. Besides, they also point out that the speed of processing visual information is very limited in human brains. Consequently, it is highly plausible that human beings utilize interrelationships between objects, or concepts, to facilitate the recognition of complex, unfamiliar objects based on the recognition of some simple, familiar objects. Obviously, the interrelationships will largely reduce the information required for recognition, and improve the effectiveness and efficiency as well. Furthermore, the lateral connections, which widely exist in the cortex of human brains (Choe 2001), also provide the biological support of the usage of interrelationships between objects.

Meanwhile, due to large amount of uncertainty in the process of perceptual organization, Bayesian methods are widely used in the modeling of perceptual organization or object identification (see e.g., (Geisler & Diehl, 2003; Jacobs, 2002; Lee & Mumford, 2003). Typically, uncertainties can rise in a large amount of cases. For example, a target object is too small in some visual scene, or only part of the target object is visible due to some obstructions, or the visual scene is vague in some bad weather. As a result, Bayesian method, which can deal with uncertainties and make use of uncertainties, is obviously a favorite option for many researchers.

In this article, we propose a variation of traditional Bayesian network, a cellular Bayesian network (CBNs), which makes use of both low-level visual features and interrelationships between high-level objects under a Bayesian framework to model visual perception.

The rest of this article is organized as follows: in the second section, we will briefly review the related works, mainly the Bayesian network, and analyze the weakness of traditional Bayesian network in the task of visual perception; next, in the third section, we propose our novel model, the Cellular Bayesian Networks (CBNs), including definitions, learning methods and inference in CBNs; after that, a small simulation is presented in the fourth section, where we demonstrate how to perform visual perception via a CBNs and make an experimental comparison between the CBNs and a purely Bottom-up strategy; then, in the fifth section, we discuss the merits and demerits of CBNs for the task of visual perception; finally, we conclude our article with a short review of