

## Editor's note

Image, speech and text (language) information are closely related to the human visual and auditory perception, and they play important roles in such fields as society, economy, and national security. The amount of such information has retained a rapid growth in this information explosion era. Although such information that can be directly perceived and understood by humans can also be processed by computers, the processing ability of computers is far weaker than that of humans, and their processing efficiency cannot meet the requirement of development in today's society.

In 2008, the National Natural Science Foundation of China (NSFC) started the Major Research Plan "Cognitive Computing of Visual and Auditory Information" (research period: 2008 to 2015). The overall purposes of this Major Research Plan are to study and construct new computational models and methods based on the human visual and auditory cognitive mechanism, and by giving full play to the interdisciplinary advantages of information science, life science and mathematics, to improve the computers' comprehension ability of unstructured visual and auditory perception information and their processing efficiency of massive heterogeneous information so as to overcome the bottleneck in image, speech and text (language) information processing. In 2008 and 2009, this Major Research Plan funded 46 projects as fostering projects or key funding projects from 335 proposals. Furthermore, this plan has entered the integration phase and approximately 30 projects will be funded from more than 130 proposals in 2011.

This special focus presents original and recent advances in ongoing research funded by the Major Research Plan "Cognitive Computing of Visual and Auditory Information", and seven papers have been accepted for publication in this special focus from 19 submissions in accordance with the review procedure of SCIENCE CHINA-Information Sciences. It is a collection of achievements of many researchers over the last three years.

This special focus includes the following papers:

- 1) Salient region detection and segmentation for general object recognition and image understanding (HUANG TieJun, TIAN YongHong, LI Jia, and YU HaoNan)
- 2) Monaural voiced speech segregation based on elaborate harmonic grouping strategies (LIU WenJu, ZHANG XueLiang, JIANG Wei, LI Peng, and XU Bo)
- 3) A flexible framework for HMM based noise robust speech recognition using generalized parametric space polynomial regression (CHENG Ning, LIU XunYing, and WANG Lan)
- 4) Right-and-left visual field stimulation: A frequency and space mixed coding method for SSVEP based brain-computer interface (YAN Zheng, GAO XiaoRong, and GAO ShangKai)
- 5) Spatio-temporal pattern analysis with single-trial EEG signals recorded during visual object recognition (WANG ChangMing, HU XiaoPing, YAO Li, XIONG Shi and ZHANG JiaCai)
- 6) Stable multi-label boosting for image annotation with structural feature selection (ZHUANG YueTing, HAN YaHong, WU Fei, and YANG JiaCheng)
- 7) Part-based on-road vehicle detection using hidden random field (ZHANG XueTao, HE YongJian, and WANG Fei)

Finally the Guest Editors express sincere thanks to numerous reviewers for their valuable contributions and efforts with timely and critical reviews. We are also very grateful to Ms. Song Fei of the SCIENCE CHINA Editorial Office for all her precious advice and active supports through the whole process of this special focus.

ZHENG NanNing, *Guest Editor*  
Xi'an Jiaotong University

CHEN Lin, *Guest Editor*  
Institute of Biophysics,  
Chinese Academy of Sciences

WANG ChengHong, *Guest Editor*  
National Natural Science  
Foundation of China

HU ZhanYi, *Guest Editor*  
Institute of Automation,  
Chinese Academy of Sciences

XIN JingMin, *Guest Editor*  
Xi'an Jiaotong University