Yong-hong Tian, Xi-lin Chen, Hong-kai Xiong, Hong-liang Li, Li-rong Dai, Jing Chen, Jun-liang Xing, Jing Chen, Xi-hong Wu, Wei-min Hu, Yu Hu, Tie-jun Huang, Wen Gao, 2017. Towards human-like and transhuman perception in AI 2.0: a review. *Frontiers of Information Technology & Electronic Engineering*, **18**(1):58-67. http://dx.doi.org/10.1631/FITEE.1601804

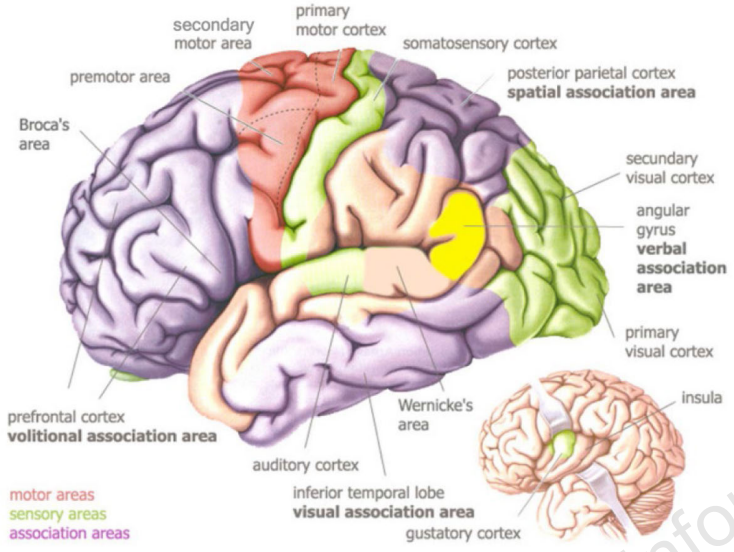
Front Inform Technol Electron Eng

Towards human-like and transhuman perception in AI 2.0: a review

Key words: Intelligent perception; Active vision; Auditory perception; Speech perception; Autonomous learning

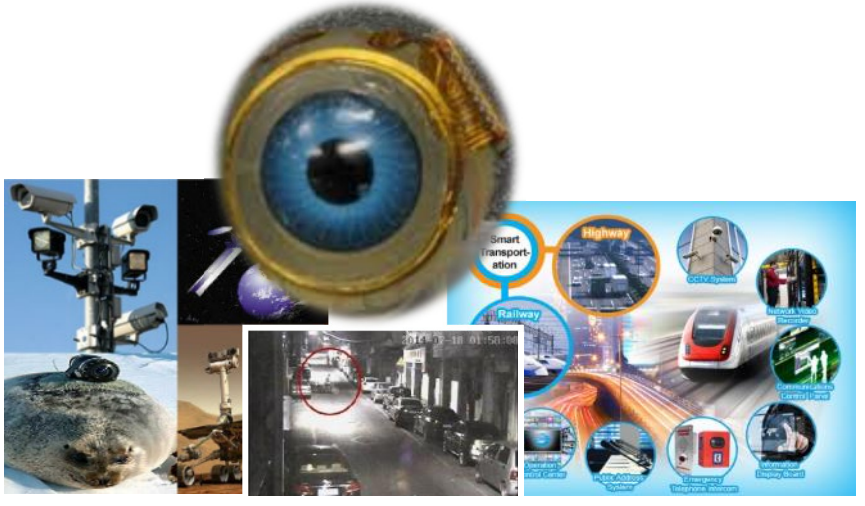
Corresponding author: Yong-hong Tian, Tie-jun Huang
E-mail: yhtian@pku.edu.cn
 ORCID: http://orcid.org/0000-0002-2978-5935

Perception is the source of intelligence



- ✓ **The source of biological intelligence is the perception of external stimuli.**
  - ➢ For example, the human brain perceives the outside world in real time through more than three million nerve fibers (more than one million fibers per eye).

- ✓ **Perception is the interaction interface between an intelligent system and the real world.**
  - ➢ Without sophisticated and flexible perceptual capabilities, it is impossible to create advanced artificial intelligence (AI) systems.



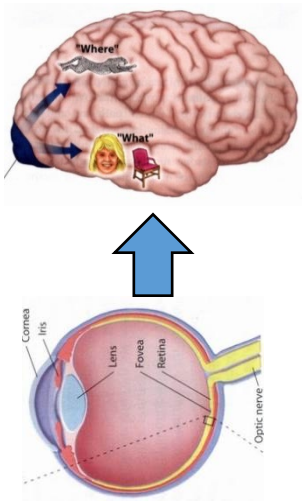
Front Inform Technol Electron Eng

Problems for perception in the current AI



- ✓ **In the AI community, researchers have struggled for several decades with the challenge of designing and implementing intelligent perception systems that can effectively simulate the brain's mechanisms.**
- ✓ **Great success has been achieved with some specific problems and tasks such as face recognition in a constrained environment, especially given the recent advances in deep learning.**
  - ➢ Example: AlphaGo obtains a sense of chess through visual perception

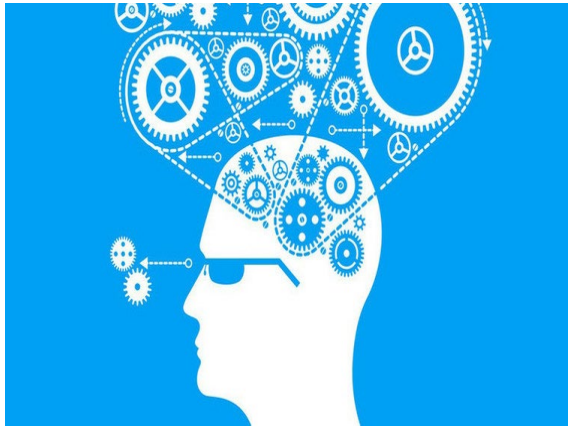
- ✓ **Problems for perception in the current AI**
  - ➢ Be capable of dealing with only specific tasks, as a specialized expert
  - ➢ Need to develop different computational algorithms or tools for different perceptual tasks, while ignoring the correlation or dependency between these tasks





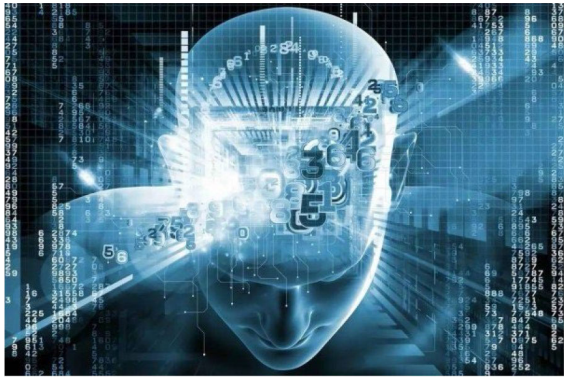
Front Inform Technol Electron Eng

A vision about intelligent perception in AI 2.0



- ✓ **Human-like perception: To simulate the brain's mechanisms**
  - ➢ To achieve intelligent perception methods and technologies that can successfully generate a uniform semantic representation of objects, scenes, behavior, and events in the real world, realize audio analysis and speech recognition in a natural auditory setting, and develop new machine learning algorithms and methods for large-scale perception data.

- ✓ **Transhuman perception: To surpass the human brain in terms of performance.**
  - ➢ To establish transhuman, intelligent perception theories, methods, and technologies. These will include active perception and learning models, human-like auditory perception and understanding technologies in actual auditory settings, and autonomous, self-evolving, and collaborative learning theories and models on intelligent perception.



Front Inform Technol Electron Eng

The technological framework of intelligent perception in AI 2.0

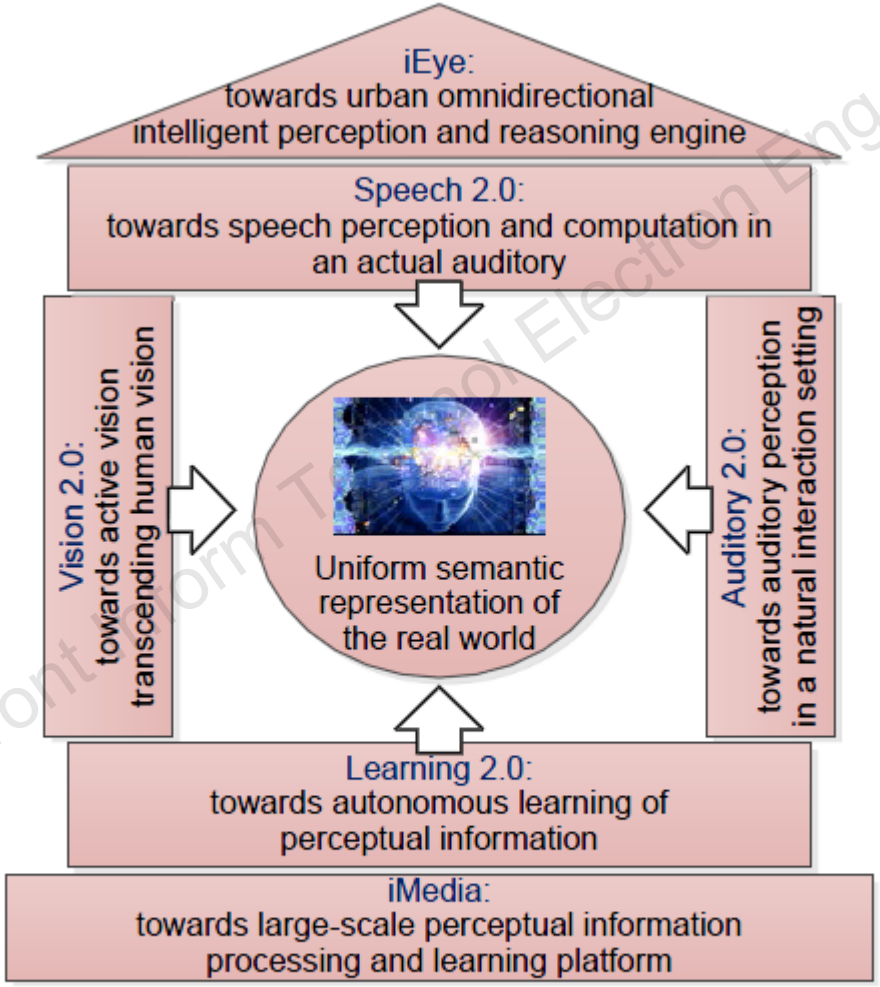


Fig. 2 A vision about the technological framework of intelligent perception in AI 2.0

The technological framework of intelligent perception in AI 2.0

1. **Vision 2.0: towards human-like and transhuman active vision**
   - To study on-site and active learning for vision tasks
   - To develop theory and algorithms beyond human sensing
2. **Auditory 2.0: towards auditory perception and computation in an actual auditory setting**
   - To study the mechanism of auditory binaural processing
   - To study the mechanisms underlying speech perception with interfering sounds
   - To develop novel algorithms and auditory computational models for improving the performance of machines in sound localization and audio understanding in complicated auditory environments
3. **Speech 2.0: towards speech perception and computation in a natural interaction setting**
   - For speech recognition, to explore brain-like models and learning algorithms by integrating the mechanism of speech perception and selective attention, and to develop novel end-to-end speech recognition frameworks with auditory context perception and adaptation
   - For speech synthesis, to develop advanced deep generative models for speech generation, and to explore approaches to extract rich information from texts to boost expressive speech synthesis

Front Inform Technol Electron Eng

The technological framework of intelligent perception in AI 2.0

4. **Learning 2.0: towards autonomous learning of perceptual information**
   - To establish a set of intelligent perceptual information processing and learning frameworks with sufficient theoretical support and autonomous learning capabilities
   - To focus on deciding where to learn by making use of the structured signal processing and the memory management mechanism of the human brain
5. **iMedia: towards large-scale perceptual information processing and learning platforms**
   - To develop the collaborative computing model over multiple datasets for improving the training efficiency of a learning algorithm
   - To develop the processors that support highly parallel floating point arithmetic, such as Chips, distributed parallel computing systems (DPCSs)
6. **iEye: towards urban omnidirectional intelligent perception and reasoning engines**
   - The core technologies include intelligent perception within the scope of a whole city, associative analysis among multiple targets, cross spatial-temporal behavioral understanding, synthesis of heterogeneous information from multiple sources, and urban panorama modeling.

Front Inform Tech El Electron Eng

Conclusions

- ✓ AI systems would greatly benefit from advancements in theory, algorithms, and hardware to enable more robust, reliable, and intelligent perception.
- ✓ Intelligent perception should be one of the most significant features of AI 2.0.
- ✓ The research directions listed in this article should be highlighted in AI 2.0.

Front Inform Technol Electron Eng

Claim: This article is a summary of the specialists' opinions from a subcommittee on intelligent perception technologies, supported by the research project on the National Artificial Intelligence 2.0 Research and Development Strategy from the Chinese Academy of Engineering.