

The First Tsinghua-Kyoto Symposium on  
Intelligent Technologies and Information Management for Knowledge Society  
August 31<sup>th</sup> ~ September 1<sup>st</sup>, 2009, Tsinghua University

Organized by

State Key Laboratory of Intelligent Technology and Systems, Tsinghua  
University

Informatics Education and Research Center for Knowledge-Circulating  
Society, Kyoto University

Supported by

Department of Computer Science and Technology, Tsinghua University

Graduate School of Informatics, Kyoto University

Participants from Kyoto University:

Professor Toru Ishida (invited talk)

Professor Katsumi Tanaka (invited talk)

Professor Sadao Kurohashi (invited talk)

Professor Toyoaki Nishida (invited talk)

Professor Takashi Matsuyama (invited talk)

Associate Professor Tetsuya Ogata (invited talk)

Participants from Tsinghua University:

Professor Bo Zhang (invited talk)

Professor Shimin Hu (invited talk)

Professor Maosong Sun

Professor Shiqiang Yang

Professor Shaoping Ma

Professor Xiaoyan Zhu

Professor Lianhong Cai

Professor Yuanchun Shi

Professor Juanzi Li

Associate Professor Jianyong Wang, Min Zhang, Lifeng Sun

Assistant Professor Minlie Huang, Jianmin Li, Jie Tang,

Dr. Yu Hao, Jia Jia

August 31th, meeting room: FIT 1-315

9:00-9:30 Opening and Introduction

Professor Xiaoyan Zhu

Professor Katsumi Tanaka

Invited talks

9:40-10:20

**Service-Oriented Collective Intelligence for Intercultural Collaboration**

Professor Toru Ishida

Tea break

10:40-11:20

**Web Search and Information Credibility Analysis**

Professor Katsumi Tanaka

11:20-12:00

**Content Based Image Retrieval and Machine Learning**

Professor Bo Zhang

Lunch

13:00-13:40

**Natural Language Processing based on and for Information Explosion on the Web**

Professor Sadao Kurohashi

13:40-14:20

**Conversational Informatics for Situated Communication**

Professor Toyoaki Nishida

14:20-15:00

**Modeling Dynamics in 3D Video Production and Human Communication**

Professor Takashi Matsuyama

Tea break

15:20-16:00

**Dynamics of Human-Robot Interaction with Robot Audition**

Associate Professor Tetsuya Ogata

16:00-16:40

**PhotoSketch: Internet Image Montage**

Professor Shimin Hu

16:40-17:00 Closing and Future Relations

Professor Maosong Sun

Dinner (with performance)

Sept. 1<sup>st</sup>, meeting room: FIT 3-315 or labs

Lab tour

8:00-9:00

**Information Acquisition**

Prof. Xiaoyan Zhu's group

9:00-10:00

**Information Retrieval**

Prof. Shaoping Ma's group

10:00-11:00

**Natural Language Processing**

Prof. Maosong Sun's group

11:00-12:00

**Video Retrieval**

Prof. Shixiang Yang and Prof. Bo Zhang's group

Lunch (guest house of Tsinghua University)

14:00-15:00

**Core Technologies for Smart Space**

Prof. Yuanchun Shi's group

15:00-16:00

**Human Computer Speech Interaction**

Prof. Lianhong Cai's group

16:00-17:00

**Graph Data Mining and its Applications**

Prof. Jiangyong Wang's group

17:00-18:00

**Semantic Web and Social Network Mining**

Prof. Juanzi Li's group

Dinner (Beijing Duck)

## Web Search and Information Credibility Analysis

Professor Katsumi TANAKA

Graduate School of Informatics (Department of Social Informatics)

Web search, video retrieval, database, multimedia information systems

<http://www.dl.kuis.kyoto-u.ac.jp/~tanaka/>

[tanaka@dl.kuis.kyoto-u.ac.jp](mailto:tanaka@dl.kuis.kyoto-u.ac.jp)

As computers and computer networks become more common, a huge amount of information, such as that found in Web contents including multimedia content (images, videos etc.), has been accumulated and circulated. Such information gives people a framework for organizing their private and professional lives. However, in general, the quality control of Web content is insufficient due to low publishing barriers. In result there is lot of mistaken and unreliable information on the Web that can have detrimental effects on users. This calls for technology that would facilitate judging the trustworthiness of content and the accuracy of the information that users encounter on the Web. Such technology should be able to handle a wide range of tasks: extracting credible information related to a given topic, organizing this information, detecting its provenance, clarifying background, facts, and other related opinions and the distribution of them, and so on. Also, as for Web search, conventional Web search engines still suffer from their low precision/recall ratio especially for searching multimedia contents (images, videos etc.). In this talk, we propose an idea to enhance conventional Web search and to analyze Web information credibility by bridging the gap between Web 1.0 contents and Web 2.0 contents. Based on the idea, we overview our research activities on Web search and their information credibility.

Professor Katsumi Tanaka received the BS, MS and PhD degrees in Information Science from Kyoto University, in 1974, 1976 and 1981, respectively. In 1986, he joined the Department of Instrumentation Engineering, Faculty of Engineering at Kobe University, as an associate professor. In 1994, he became a full professor at the Department of Computer and Systems Engineering Department, Faculty of Engineering, Kobe University. Since 2001, he has been a professor of the Graduate School of Informatics, Kyoto University. His research interests include database theory and systems, Web search, video retrieval, and multimedia information systems.

## **Conversational Informatics for Situated Communication**

Professor Toyoaki NISHIDA

Department of Intelligence Science and Technology

<http://www.ii.ist.i.kyoto-u.ac.jp/~nishida/>

[nishida@i.kyoto-u.ac.jp](mailto:nishida@i.kyoto-u.ac.jp)

Messages need to be situated for communication to make real sense in our life. Conversation allows us to exchange situated messages in simultaneous streams of social interactions at multiple levels. Conversation provides an effective means for circulating knowledge in a community. Conversational Informatics comprises a theoretical foundation of our endeavors of understanding and augmentation of conversations. It consists of four research areas: conversational artifacts, conversational contents, conversational environment design, and conversation measurement, analysis and modeling. As for conversational artifacts, we aim at building embodied conversational agents and conversational robots that can participate in conversational interactions among people. As for conversational contents, we develop technologies that will allow us to capture knowledge arising in a conversational situation and reuse it to enhance later conversations. As for conversational environment design, we design smart environments that can sense conversational behaviors to either help participants become involved in a collaboration, or record conversation for later use. As for conversation measurement, analysis and modeling, we take a data-driven quantitative approach to understanding conversational behaviors. I will overview recent results in conversational informatics and discuss new challenges.

Toyoaki Nishida is a professor of Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University. He received the Doctor of Engineering degree from Kyoto University in 1984. His research centers on artificial intelligence and human computer interaction. In 2001, he founded a series of international workshops on social intelligence design. Then, he broadened the scope of research to include understanding and augmenting conversational communication, and opened up a new field of research called Conversational Informatics. Currently, he leads several projects on social intelligence design and conversational informatics. He is a member of the board of directors of IPS (Information Processing Society) of Japan and JSAI (Japanese Society for Artificial Intelligence). He serves as an editorial board member of several academic journals, including Web Intelligence and Agent Systems, AI & Society, and Journal of JSAI (editor-in-chief).

## **Modeling Dynamics in 3D Video Production and Human Communication**

Professor Takashi MATSUYAMA

Department of Intelligence Science and Technology

<http://vision.kuee.kyoto-u.ac.jp/index.html>

[tm@i.kyoto-u.ac.jp](mailto:tm@i.kyoto-u.ac.jp)

Computer Vision and Visual Pattern Recognition have been analyzing image and video data mainly based on Geometry and Photometry. In this talk, we present our recent research activities on visual information analysis based on Dynamics. The first topic is on 3D Video, which reconstructs complete 3D human actions from multi-view video data. Its applications include digital archiving of traditional dances, sports training, detailed gesture analysis, 3D TV, and so on. The second topic is on the analysis of dynamics of human communication. We will show how communication dynamics can be modeled mathematically and how we can design smooth and natural human-machine communication with several demonstrations.

Takashi Matsuyama received the Doctor of Engineering degrees from Kyoto University Faculty of Engineering in 1980 respectively. He is currently a professor in the Department of Intelligence Science and Technology, Graduate school of Informatics, Kyoto University, Japan. He is also a Director-General of the Institute for Information Management and Communication, also at Kyoto University. He was a recipient of the Marr Prize at the 5th International Conference of Computer Vision in 1995 and won the Funai Best Paper award at IPSJ FIT Conference 2005. He is interested in Image Understanding, Cooperative Distributed Vision and Three Dimensional Video Capture.

## **Service-Oriented Collective Intelligence for Intercultural Collaboration**

Professor Toru ISHIDA

Department of Social Informatics, Graduate School of Informatics

<http://www.ai.soc.i.kyoto-u.ac.jp/~ishida/index.htm>

[ishida@i.kyoto-u.ac.jp](mailto:ishida@i.kyoto-u.ac.jp)

To increase the accessibility and usability of online language services, this talk explains the Language Grid, which facilitates the creation of composite language services for various intercultural collaboration activities. The Language Grid is an initiative to build an infrastructure that allows end users to create new language services for their intercultural / multilingual activities. To this end, language resources (including data and programs) are wrapped as web services so that users can easily combine these services to create workflows that suit their own activities. Thus, the Language Grid can be seen as collective intelligence based on language services. The Language Grid is called "horizontal," when the grid connects standard languages, or "vertical," when the grid combines the language services generated by communities. There are four types of stakeholders in the Language Grid: Language Resource Provider, Computation Resource Provider, Language Service User, and Language Grid Operator (who coordinates the other stakeholders). Though there can be various operation models for the Language Grid, we propose a non-profit operation model in this talk. This model limits the usage of language services solely to non-profit operations, tries to match the incentives of stakeholders, and manages various issues associated with intellectual property rights, user privacy, and operation costs.

Toru Ishida is a professor of Department of Social Informatics, Kyoto University and a leader of the NICT Language Grid Project. He is an IEEE fellow from 2002. He has been working on autonomous agents and multiagent systems for twenty years. He also studies social informatics and is running research projects related to digital cities and intercultural collaboration. His professional services include an associate editor of IEEE Transactions on Pattern Analysis and Machine Intelligence, an associate editor of Journal of Autonomous Agents and Multi-Agent Systems, a co-editor in chief of Elsevier Journal on Web Semantics, a program co-chair of International Conference on Multiagent Systems (ICMAS'96), and a general co-chair of the first international conference on Autonomous Agents and Multi-Agent Systems (AAMAS'02).

## **Natural Language Processing based on and for Information Explosion on the Web**

Professor Sadao KUROHASHI

Graduate School of Informatics (Department of Intelligence Science and Technology)

<http://nlp.kuee.kyoto-u.ac.jp/~kuro/index.html>

[kuro@i.kyoto-u.ac.jp](mailto:kuro@i.kyoto-u.ac.jp)

Information explosion on the web has brought NLP research community two big impacts. One is that it provides us very huge knowledge sources from which we can extract linguistic knowledge and extra-linguistic knowledge (common sense knowledge). This situation makes it possible to solve the deadlock between language understanding and knowledge acquisition in a bootstrap way. Another impact is that NLP-based intelligent support to exploit the web information becomes a killer application for NLP, since information on the web becomes more and more important, giving judgment criteria for people's daily life and starting to have a strong influence on governmental policy and business management. This talk introduces our several on-going projects concerning NLP based on and for information explosion on the web.

Sadao Kurohashi received the B.S., M.S., and PhD in Electrical Engineering from Kyoto University in 1989, 1991 and 1994, respectively. He has been a visiting researcher of IRCS, University of Pennsylvania in 1994. He is currently a professor of the Graduate School of Informatics at Kyoto University. His research interests include natural language processing, knowledge acquisition/representation, and information retrieval.

## **Dynamics of Human-Robot Interaction with Robot Audition**

Associate Professor Tetsuya OGATA

Graduate School of Informatics (Department of Intelligence Science and Technology)

<http://winnie.kuis.kyoto-u.ac.jp/~ogata/hp>

[ogata@i.kyoto-u.ac.jp](mailto:ogata@i.kyoto-u.ac.jp)

In human robot interaction, robot should understand an arbitrary sound mixture including non-speech sounds and music as well as voiced speech, obtained by its' ears. It also should ground the sounds to "symbols" which dynamically changes with interaction context. This talk includes two topics. One topic is robot audition techniques for computational auditory scene analysis, that is, sound source localization, separation, and recognition of separated sounds for a mixture of speech signals as well as polyphonic music signals. In this talk, we present our recent results on missing feature theory based integration of sound source separation and automatic speech recognition, and those in music information processing. The other is a synthetic approach named "System informatics of interaction emergence" for designing human robot interactions as dynamical systems. This approach couples multiple dynamics of the recurrent neural networks, the robot system, and the environment in order to describe symbols using dynamical systems language. In this talk, we present our recent studies on motion imitation, multi-modal translation and motion-language translation.

Tetsuya Ogata received the BS, MS and DE degrees in Mechanical Engineering in 1993, 1995, and 2000, respectively, from Waseda University. From 1999 to 2001, he was a Research Associate in Waseda University. From 2001 to 2003, he was a Research Scientist in the Brain Science Institute, RIKEN. Since 2003, he has been a Faculty Member in the Graduate School of Informatics, Kyoto University, where he is currently an Associate Professor. Since 2005, he has been a Visiting Associate Professor of the Humanoid Robotics Institute of Waseda University. His research interests include human-robot vocal-sound interaction, dynamics of human-robot mutual adaptation and active sensing with robot systems.

## Content Based Image Retrieval and Machine Learning

Professor Bo Zhang

Department of Computer Science and Technology

[http://www.csai.tsinghua.edu.cn/personal\\_homepage/zhang\\_bo/index.html](http://www.csai.tsinghua.edu.cn/personal_homepage/zhang_bo/index.html)

[dcszb@tsinghua.edu.cn](mailto:dcszb@tsinghua.edu.cn)

Content based image retrieval is to search for a set of images with the same semantic meaning from a given image-base under a given query (image). This is an instance of content based information processing and it contains several key problems of the latter. So it is one of the hot topics in information processing. The main method used in image retrieval recently is the statistical learning approach or data driven approach. We also address this popular methodology.

First, the representation problem: An image (object) can be represented at different granularities in computers. We will discuss how the grain-size affects the searching performance such as accuracy and how to choose a proper grain-size. We will show that multi-granular representation can improve the searching performance as well and present a mathematical model for the multi-granular analysis.

Second, the structural analysis: The structure of an image contains plenty of semantic information. We will show how to detect the structural knowledge and how the knowledge affects the searching performance.

Third, we will discuss the essence of recent data driven approach and its limitation. In order to overcome the limitation, we should learn some things from human being, that is, the human visual mechanism.

Finally, we will show the future direction of the content based image retrieval.

In 1958 Bo Zhang graduated from Automatic Control Department of Tsinghua University, and became a faculty member since then. From 1980/02 to 1982/02 he visited University of Illinois at Urbana-Champaign, USA as a scholar. He is now a professor of Computer Science and Technology Department of Tsinghua University, the fellow of Chinese Academy of Sciences. He is engaged in the research on artificial intelligence, artificial neural networks, machine learning and so on. And he also is engaged in the research applying technology that applies the theories mentioned above into pattern recognition, knowledge engineering, and robotics. He took part in building up the State Key Lab of Intelligent Technology and Systems and served as director of the lab from 1991 to 1996. From 1987 to 1994 he served as a member of specialist group of Intelligent Robots theme of National '863' High-Tech Program.

## **PhotoSketch: Internet Image Montage**

Professor Shimin Hu

Department of Computer Science and Technology

<http://cg.cs.tsinghua.edu.cn/~shimin/>

[shimin@tsinghua.edu.cn](mailto:shimin@tsinghua.edu.cn)

We present a system that composes a realistic picture from a user provided sketch with text labels. The composed picture is generated by seamlessly stitching several photographs automatically searched from internet according to the sketch and its text labels. While on line image search generates noisy results, our system can automatically select suitable photographs to generate a high quality composition. To achieve this, we first design a filtering scheme to exclude undesirable images from searched results. Then we propose a novel image blending algorithm for seamless image composition. Our blending algorithm returns a numeric score for each blending, which is used to optimize the combination of searched images. Several vivid results are generated in the experiments. We also perform a user study to demonstrate the advantages of our system.

Dr Shi-Min Hu is currently a chair professor in the Department of Computer Science and Technology, Tsinghua University, Beijing. He received the PhD degree from Zhejiang University in 1996. His research interests include digital geometry processing, video processing, rendering, computer animation, and computer-aided geometric design. He is on the editorial board of Computer Aided Design (Elsevier). He is member of the CCF, IEEE and ACM.