National Tsing Hua University – Tohoku University Bilateral Workshop

Time: December 28, 2009

Place: EECS Building, National Tsing Hua University, Hsin-Chu

Co-sponsored by

College of Electrical Engineering and Computer Science, National Tsing Tua University, Taiwan

Graduate School of Information Science (GSIS), Tohoku University, Japan

Global COE (Center Of Excellence) Program: Center of Education and Research for Information Electronics Systems

9:30 - 10:00 Opening R-447

Chair: Dean Jyuo-Min Shyu (徐爵民院長)

9:30 - 9:40 Welcome

9:40 - 9:50 Introduction to participants from Tohoku University
Prof. Takao Nishizeki (西關 隆夫教授)

9:50 - 10:00 Introduction to participants from N. Tsing Hua University
Dean Jyuo-Min Shyu (徐爵民院長)

10:00 – 12:00 Group Discussion

Track I Imagine Processing R-447

Chair: Prof. Fuhua Cheng (鄭復華教授)

Attendee: Prof. Kazuyuki Tanaka (田中和之教授), Prof. Takafumi Aoki (青木孝文教授), Prof. Muneki Yasuda (安田宗樹教授), Prof. Fuhua Cheng (鄭復華教授), Prof. Long-Wen Chang (張隆紋教授), Prof. Chaur-Chin Chen (陳朝欽教授), Prof. Shang-Hong Lai (賴尚宏教授), Prof. Hwann-Tzong Chen (陳煥宗教授), Prof. Yung-Chang Chen (陳永昌教授), Prof. Chia-Wen Lin (林嘉文教授)

10:00 – 10:20 Title: Fast Mesh Interpolation and Mesh Expansion with Applications
Professor Fuhua Cheng (鄭復華教授)

10:20 – 10:40 Title: High-Accuracy Machine Vision Technology Using Phase-Only Correlation --- From 3DD Measurement to Biomedical Imaging ---
Professor Takafumi Aoki (青木 孝文教授)

10:40 – 11:00 Title: Landmark-Based Sparse Color Representations for Color Transfer
Prof. Hwann-Tzong Chen (陳煥宗教授)

11:00 – 11:20 Title: Approximate Learning Algorithm in Restricted Boltzmann Machines Using Kullback-Leibler Importance Estimation Procedure
Prof. Muneki Yasuda (安田 宗樹教授)

1**1:20 – 11:40** Title: Recognition, Analysis and Synthesis on Facial Expression Images
Prof. Shang-Hong Lai (賴尚宏教授)

11:40 – 12:00 Discussion

12:00 Adjourn

Track II　　Audio Processing and Logic Design　R-431

Chair: Professor Ting-Ting Hwang (黃婷婷教授)

Attendee：Prof. Takao Nishizeki (西關 隆夫教授), Prof. Yukio Iwaya (岩谷 幸雄教授),
Prof. Kei Uchizawa (內澤 啓教授), Prof. Hsiao-Chuan Wang (王小川教授)
Prof. Jyh-Shing Roger Jang (張智星教授), Prof. Chun-Yao Wang (王俊堯教授)
Prof. Ting-Ting Hwang (黃婷婷教授), Prof. Shih-Chieh Chang (張世杰教授)

10:00 – 10:20 Title: Synthesis of Reversible Sequential Elements
Prof. Chun-Yao Wang (王俊堯教授)

10:20 – 10:40 Title: Size and Energy of Unate Circuits Computing MOD Functions
Prof. Kei Uchizawa (內澤 啓教授)

10:40 – 11:00 Title: Synthesis of Reversible Sequential Elements
Prof. Hsiao-Chuan Wang (王小川教授)

11:00 – 11:20 Title: Three Dimensional Virtual Auditory Display and Its Application
Prof. Yukio Iwaya (岩谷 幸雄教授)

11:20 – 11:40 Title: Introduction to MIR Lab
Prof. Jyh-Shing Roger Jang (張智星教授)

11:40 – 12:00 Discussion

12:00 Adjourn

Abstract

[1] Fast Mesh Interpolation and Mesh Expansion with Applications

Prof. Fuhua (Frank) Cheng (鄭復華教授)

Abstract:
A fast iterative method for constructing a smooth subdivision surface that interpolates the vertices of an arbitrary mesh is presented. The construction is done by iteratively adjusting vertices of the given mesh locally until control mesh of the required interpolating surface is reached. The new interpolation method has not only the simplicity of a local method, but also the capability of a global method in faithfully resembling the shape of a given mesh. The new method does not require solving a linear system, hence it can handle meshes with large number of vertices. Furthermore, the new method is fast and does not require a fairing step in the construction process because the iterative process converges to a unique solution at an exponential rate.
Another important result of this work is, with the new iterative process, each mesh (surface) can be expanded as an infinite series of meshes (surfaces) which carry high and low frequency information of the given model. This mesh expansion scheme provides us with new approaches to some classic applications in computer graphics such as texture mapping and morphing. These new approaches are demonstrated in this paper and test results are included.

[2] High-Accuracy Machine Vision Technology Using Phase-Only Correlation ---From 3D Measurement to Biomedical Imaging---

Prof. Takafumi Aoki (青木 孝文教授)

[3] Landmark-Based Sparse Color Representations for Color Transfer

Prof. Hwann-Tzong Chen (陳煥宗教授)

Abstract:
We present a novel image representation that characterizes a color image by an intensity image and a small number of color pixels. Our idea is based on solving an inverse problem of colorization: Given a color image, we seek to obtain an intensity image and a small subset of color pixels, which are called landmark pixels, so that the input color image can be recovered faithfully using the intensity image and the color cues provided by the selected landmark pixels. We develop an algorithm to derive the landmark-based sparse color representations from color images, and use the representations in the applications of color transfer and color correction. The computational cost for these applications is low owing to the sparsity of the proposed representation. The landmark-based representation is also preferable to statistics-based representations, e.g., color histograms and Gaussian mixture models, when we need to reconstruct the color image from a given representation.

[4] Approximate Learning Algorithm in Restricted Boltzmann Machines Using Kullback-Leibler Importance Estimation Procedure

Prof. Muneki Yasuda (安田 宗樹教授)

Abstract:
A restricted Boltzmann machine (RBM) is one of the learning models which is represented by the Markov network defined on the bipartite graph with two layers. The one is the visible layer and the other is the hidden layer. The RBM is a component of Hinton's deep belief networks [1] which is paid attention by many researchers recently. Studying a learning algorithm in the RBM, then, is important for the machine learning researching field and other scientific fields. However, in general, the learning problem in the RBM belongs to the NP-hard class which requires an effective approximate method.
Roux and Bengio (RB) proposed a new learning algorithm for the RBM which can optimize the variational bound of the deep belief network [2]. In their algorithm, one finds suitable parameters matching an empirical distribution with the distribution of the visible layer reconstructed through the hidden layer. The RB learning algorithm has a double-sum problem. In this learning algorithm, one needs to take costly double-sum over empirical data points at each update step. Thus, the RB learning algorithm is not implementable without any effective approximations. In our presentation, using the Kullback-Leibler importance estimation procedure (KLIEP) proposed by Sugiyama et al. [3], we will propose the effective approximate learning algorithm based on the RB learning algorithm. The present works are collaborations with Mr. Tetsuharu Sakurai and Prof. Kazuyuki Tanaka.
References
[1] G. E. Hinton, S. Osindero, and Y. W. Teh: A fast learning algorithm for deep belief nets, Neural Computation, Vol.18, No.7, pp.1527-1554, 2006.
[2] N. L. Roux and Y. Bengio: Representational Power of Restricted Boltzmann Machines and Deep Belief Networks, Neural Computation, Vol.20, No.6, pp.1631-1649, 2008.
[3] M. Sugiyama, S. Nakajima, H. Kashima, P. von Bunau, and M. Kawanabe:
Direct Importance Estimation with Model Selection and Its Application to Covariate Shift Adaptation, In J. C. Platt, D. Koller, Y. Singer, and S. Roweis (Eds.), Advances in Neural Information Processing Systems 20, pp.1433-1440, 2008.

[5] Recognition, Analysis and Synthesis on Facial Expression Images

Prof. Shang-Hong Lai (賴尚宏教授)

Abstract:
Face image processing has attracted considerable attention in the computer vision and image processing fields in the past two decades. In this talk, I will present some of my recent researches related to facial expression image processing. They can be divided into three parts. The first part is about recognition from facial expression images. I will talk about our recent work on face recognition and facial expression recognition based on optical flow estimation. Secondly, I will talk about our research on the analysis of facial expression images. This is mainly the construction of a low-dimensional nonlinear manifold for representing facial expressions, which has been applied for the 3D face model

reconstruction from a single face image. Finally, I will talk about our work on the synthesis of facial expression images which is also based on this expression manifold.

[6] Synthesis of Reversible Sequential Elements

Prof. Chun-Yao Wang (王俊堯教授)

Abstract:
To construct a reversible sequential circuit, reversible sequential elements are required. This work presents novel designs of reversible sequential elements such as D latch, JK latch, and T latch. Based on these reversible latches, we also construct the designs of the corresponding flip-flops. Comparing with previous work, the implementation cost of our new designs, including the number of gates and the number of garbage outputs is considerably reduced.

[7] Size and Energy of Unate Circuits Computing MOD Functions

Prof. Kei Uchizawa (內澤　啓教授)

Abstract:
Let C be a unate logic circuit computing a Boolean function $\mathrm{MOD}_m:\{0,\ 1\}^n\rightarrow\{0,\ 1\}$, where $n \geq 1$ and $m \geq 2$. Then C outputs "0" if the number of "1"s in an input $x \in \{0, 1\}^n$ to C is a multiple of m and, otherwise, C outputs "1." The function MOD_2 is the so-called PARITY function, and MOD_{n+1} is the OR function. Let s be the size of the circuit C, that is, C consists of s unate gates, and let e be the energy complexity of C, that is, at most e gates in C output "1" for any input $x \in \{0, 1\}^n$. In the paper, we prove that a very simple inequality $n/(m-1) \leq s^e$ holds for every circuit C computing MOD_m. The inequality implies that there is a tradeoff between the size s and energy complexity e of circuit C computing MOD_m, and yields a lower bound $e = \Omega((\log n - \log m)/\log\log n)$ on e if $s = O(\mathrm{polylog}(n))$. We actually obtain a general result on the so-called generalized mod function, from which the result on the ordinary mod function MOD_m immediately follows.

[8] Synthesis of Reversible Sequential Elements

Prof. Hsiao-Chuan Wang (王小川教授)

Abstract:
The acoustic models of corpus-based automatic speech recognition (ASR) system may not be able to describe the detail of articulation procedure. When the speech quality is degraded because of background noise, channel distortion, or speech variation, the performance of ASR system becomes unsatisfactory. A knowledge-based plus data-driven paradigm should be an alternative approach to the new ASR system. Speech landmark detection is a way to reach this goal. A set of various types of landmark detectors will be the front-end processor of ASR system where each detector is designed for a specific articulation mode. Burst onset is one of most crucial speech landmarks in speech signal. It would be helpful to the

recognition of stop/affricate phones. This study proposes the use of two-dimensional cepstral (TDC) feature and random forest (RF) technique to generate the burst onset detector. Only limited training data are needed for generating a reliable RF-based burst onset detector. By appending the burst presence probability obtained from RF-detector to MFCC feature, the performance of stop/affricate recognition can be greatly improved.

[9] Three dimensional virtual auditory display and its application

Prof. Yukio Iwaya (岩谷 幸雄教授)

Abstract:
We can perceive sound position as sound image with sound localization ability. In sound localization, we use head-related transfer functions (HRTFs) empirically. Using this principle reversely, we can control a sound position to arbitrary position based on convolution of HRTFs to a sound. Such system is called as three-dimensional virtual auditory display (VAD).
There are some points to establish high accuracy of VAD system; Individualization of HRTFs to listener is essentially needed. Because localization error is often large, when HRTFs used in VAD system is not individualized to a listener.
Another important point in sound localization is head movement. When VAD system is realized by headphones without head-movement-sensing device, the absolute position of sound image moves. This does not occur in real world. Therefore, VAD system should become responsive to listener's head movement. In the first half of this talk, investigations of sound image perception in virtual world are discussed. We also have put a lot of efforts into developing a high performance VAD software engine as middleware on personal computers. This VAD works with a native CPU of a personal computer and outputs sounds for headphones installed with a three- dimensional position sensor and with individualized HRTFs. In this presentation, the overview and applications of the VAD middleware are introduced.

[10] Speech and Music-related Research and Application at MIR Lab

Prof. Jyh-Shing Roger Jang (張智星教授)

Abstract:
In this talk, I will give a brief introduction to the research foci and related applications at MIR lab. In particular, I shall describe our recent research goals for speech assessment and music retrieval. Moreover, I will also demonstrate some of the prototype systems based on our current and previous research results.