Short Term Mobility Report
Ercan Engin Kuruoglu
Submitted to CNR and Shanghai Jiao Tong University

My visit to Shanghai Jiao Tong University, Institute of Image Communication and Information Processing took place between dates 9-30 October 2008. During my visit I had the opportunity to establish strong research connections with Shanghai Jiao Tong University. The trip has been very educative for me into video coding state of the art methodology and about Chinese research structures. My main research activity was carried out in collaboration with Prof Jia Wang and Prof Xiaokang Yang deputy director of Institute of Image Communication and Information Processing Institute but I also had contacts with Prof Peilin Liu, director of multimedia research center and Prof Tong Zhou, director of Georgia Tech-Shanghai.

During my visit I gave one academic talk with title "Image Estimation and Component Separation Using Numerical Bayesian Techniques" which was followed by a crowd of 40 researchers and students and excited a number of questions leading to research ideas.

My research activity in Shanghai Jiao Tong University in this period can be summarised in five main topics which are detailed below.

1. Video event tracking

I continued supervising the Master student Zhaowen Wang in collaboration with Prof Xiaokang Yang and Dr Xi Yu on the subject of human motion tracking in video sequences. The project has several applications:

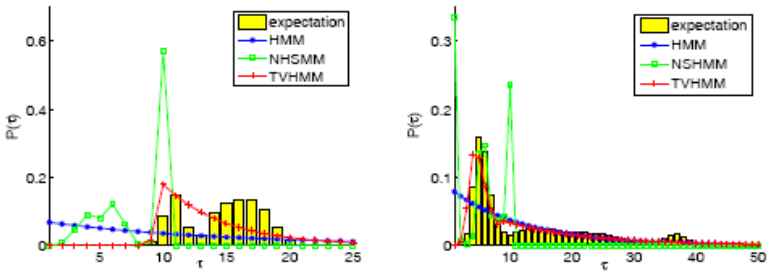
- Detecting unusual behaviours from video sequences such as drunken driving or safety in a hospital environment
- Modelling human behaviour in certain environments such as a bank service queue

Standard Hidden Markov Model (HMM) and the more general Dynamic Bayesian Network (DBN) models assume stationarity of state transition distribution. However, this assumption does not hold for many real life events of interest. In this work, we propose a new time sequence model that extends HMM to time varying scenario. The time varying property is realized in our model by explicitly allowing the change of state transition density as the time spent in a particular state passes by. Rather than keeping these transition densities at different time spots independent of each other, we exploit their temporal correlation by applying a hierarchical Dirichlet prior. This leads to a more robust time varying model, especially when training data are scarce. We also employ Monte Carlo Markov Chain (MCMC) sampling in learning the MAP estimate of time varying parameters, with a transition kernel incorporating linear optimization. The proposed model is applied to recognize real video events, and outperforms the existing HMM-based methods. We emphasize that the contribution we made on Dynamical Bayesian Network estimation is applicable in various applications.

Below in Table 1 and Figure 1, an evaluation of the success of the technique is given:

Table 1. Average recognition rate for human motion.

model	running	walking	jumping
HMM	0.994	0.551	0.414
NHSMM	0.996	0.791	0.858
TVHMM	0.999	0.856	0.984



(a) state 3, running event (b) state 6, walking event

Figure 1. State duration distribution (TVHMM: time varying Hidden Markov Model suggested in this work).

The output of this research is already a paper submitted to ICASSP, the main signal processing conference in the world [1]. In the continuation of the work, we will study a continuous distribution as the prior. Industrial applications will also be considered.

2. Source coding with impulsive distributions

As in various other fields, in the field of source coding, the Gaussian assumption is very common. However various real world scenes in image and video applications involve data which is rather impulsive and distinctively non-Gaussian. The rate distortion theory is developed for the Gaussian data and the results are not valid for non-Gaussian data. In this work, we aim to extend the rate distortion theory results to impulsive distributions.

We will use a simple algorithm designed by Blahut [2] for the calculation of the rate-distortion curves for the alpha-stable distribution and look at various real life image and video signals compression which we believe are more realistically described with the rate distortion curves for alpha-stable distributions which are very good models for data with impulsive and non-symmetric characteristics. I am collaborating with Prof Jia Wang and his student Tao Zhou in this problem.

3. Calculation of channel capacity for impulsive channels

Various recent communication technologies such as mobile communications, satellite communications and xDSL, defy the Gaussian channel assumption since they operate under impulsive noise channels. These channels are characterised by sudden bursts of noise which have high amplitudes frequently surpassing signal amplitude. Traditional communication systems are designed under Gaussian channel assumption and the channel capacities developed with this assumption are not valid for impulsive channels. In this work, we aim to develop channel capacity for alpha stable distribution which is a very popular model for impulsive noise. We will employ Blahut's algorithm [2] for this objective. This theoretical result will help us design communication systems with more realistic conditions and therefore we expect important applicative results from this work.

4. Reflection cancellation in videos

Reflections from windows and other shiny surfaces create an important problem in video sequences. For video for leisure applications, user would like to avoid any reflection from windows while they are registering a video. In security applications, reflections prevent correct recognition of subjects.

A recent paper by Kayabol and Kuruoglu [3] propose a method which utilizes colour information for reflection cancellation from still colour images. Our work aims at extending this work to video sequences and utilize also motion information. We foresee important industrial applications of this problem. In this work I am collaborating with Prof Xiaokang Yang and his student Guo Kai.

5. Encoder structure optimization for data compression systems

Kuruoglu and Ayanoglu have demonstrated in [4] that in a data compression system, optimisation of the encoder structure (next-state map) provides very significant gains. However, in the literature and in many standards, the encoder structure is assumed fixed and only the decoder is optimised. We would like extend the work of Kuruoglu and Ayanoglu and see its success in commonly used coding schemes such as Trellis Coded Quantisation (TCQ).
In this project I am collaborating with Prof Jia Wang and his student.

Final issues

This trip has given me an opportunity in developing important applicative research ideas. It has been agreed to look for other types of funding outside CNR and SJTU that can lead to a collaboration on a bigger scale and possibly bring benefits to CNR and SJTU. A possible venue has been identified to be 7^{th} framework program. Inclusion of Trinity College Dublin and University of Freiburg have been discussed. Industry funding ways are also being considered.

I have been impressed once again by the Chinese way of hosting. All is very well thought of and considered from beginning. I was offered a comfortable environment in which I could think of only research. I appreciate very much SJTU's hospitality and would like to express my gratitude. In Italy, we should seriously consider how we can equal their level of organisation for hosting visitors.

References:

[1] Z. Wang, E. E. Kuruoglu, X. Yang, Y. Xu, and S. Yu, "Event recognition with time varying hidden Markov model" submitted to ICASSP 2009.

[2] R. E. Blahut, ``Computation of Channel Capacity and Rate-Distortion Function," *IEEE Transactions on Information Theory*, Vol. 18, No. 4, pp. 460-473, July 1972.

[3] K. Kayabol, E. E. Kuruoglu and B. Sankur, "Image Source Separation using Color Channel Dependencies" submitted to ICA 2009.

[4] E. E. Kuruoglu and E. Ayanoglu, "The Design of Finite-State Machines for Quantization Using Simulated Annealing," in *Coding and Quantization*, pp. 175-184, Eds. R. Calderbank, G.D. Forney and N. Moayeri, American Mathematical Society, Series in Discrete Mathematics and Theoretical Computer Science, 1993.