ΠΑΡΟΥΣΙΑΣΗ
ΔΙΔΑΚΤΟΡΙΚΗΣ ΔΙΑΤΡΙΒΗΣ

ΗΜΕΡΟΜΗΝΙΑ: Πέμπτη, 25 Φεβρουαρίου 2016

ΩΡΑ: 09:00

ΑΙΘΟΥΣΑ: Αίθουσα Σεμιναρίων
Κτήριο Τμήματος Μηχανικών Η/Υ & Πληροφορικής

ΟΜΙΛΗΤΗΣ: **Μιχαήλ Βρίγκας**

Θ έ μ α

«Human Activity Recognition Using Conditional Random Fields and Privileged Information»
ή
«Αναγνώριση Ανθρώπινης Δραστηριότητας με υπό Συνθήκη Τυχαία Πεδία και Προνομιακή Πληροφορία»

Επταμελής Εξεταστική Επιτροπή:

1. **Χριστόφορος Νίκου**, Αναπληρωτής Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων (*επιβλέπων*).
2. **Ioannis Kakadiaris**, Professor, Department of Computer Science, University of Houston.
3. **Λυσίμαχος-Παύλος Κόντης,** Αναπληρωτής Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων.
4. **Αντώνιος Αργυρός,** Καθηγητής του Τμήματος Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης.
5. **Αριστείδης Λύκας,** Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων.
6. **George Bebis,** Professor, Department of Computer Science and Engineering, University of Nevada, Reno, USA.
7. **Κωνσταντίνος Μπλέκας,** Αναπληρωτής Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων.

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

Τ.Θ. 1186, ΙΩΑΝΝΙΝΑ 45110
T: 265100 8817 - 8813 - 7196
http://www.cse.uoi.gr/

DEPT. OF COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF IOANNINA

P.O. BOX 1186, IOANNINA
GR - 45110 GREECE
T: +30 265100 8817 - 8813 - 7196
http://www.cse.uoi.gr/

Π ε ρ ί λ η ψ η

Recognizing human activities from video sequences or still images is a challenging task due to problems such as background clutter, partial occlusion, changes in scale, viewpoint, lighting, and appearance. Many applications, including video surveillance systems, human-computer interaction, and robotics for human behavior characterization, require a multiple activity recognition system.

In the first part of this thesis, after a review of the state-of-the-art methods, a learning based framework for action representation and recognition relying on time series of optical flow motion features is presented. In the learning step, the motion curves representing each action are clustered using Gaussian mixture modeling (GMM). In the recognition step, the optical flow curves of a probe sequence are also clustered using a GMM, then each probe sequence is projected onto the training space and the probe curves are matched to the learned curves using a non-metric similarity function based on the longest common subsequence, which is robust to noise and provides an intuitive notion of similarity between curves.

Next, a human behavior recognition method with an application to political speech videos is presented. The behavior of a subject is modeled using a conditional random field (CRF). To evaluate the performance of the model, a novel behavior dataset is introduced, which includes low resolution video sequences depicting different people speaking in the Greek parliament. The subjects of the Parliament dataset are labeled as friendly, aggressive or neutral depending on the intensity of their political speech.

An extension of the aforementioned human behavior recognition method using multimodal features is also presented. Individual and social behaviors of a subject are modeled using a hidden conditional random field (HCRF). Each video is represented by a vector of spatio-temporal visual features along with audio features. To remove irrelevant features a feature pruning method based on the spatio-temporal neighborhood of each feature in a video sequence is presented. The proposed framework assumes that human movements are highly correlated with sound emissions and canonical correlation analysis is employed to find relationship between the audio and video features prior to fusion.

Besides the classical learning frameworks, a novel method based on the learning using privileged information (LUPI) paradigm for recognizing complex human activities is proposed that handles missing information during testing. A supervised probabilistic approach that integrates LUPI into an HCRF model is presented. The proposed model employs a self-training

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

Τ.Θ. 1186, ΙΩΑΝΝΙΝΑ 45110
Τ: 265100 8817 - 8813 - 7196
http:/www.cse.uoi.gr/

DEPT. OF COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF IOANNINA

P.O. BOX 1186, IOANNINA
GR - 45110 GREECE
T: +30 265100 8817 - 8813 - 7196
http://www.cse.uoi.gr/

technique for automatic estimation of the regularization parameters of the objective function. Moreover, the method provides robustness to outliers by modeling the conditional distribution of the privileged information by a Student's t-density function. Different forms of additional information were investigated.

In many human activity recognition systems the size of the unlabeled training data may be significantly large due to expensive human effort required for data annotation. Moreover, the insufficient data collection process from heterogenous sources may cause dissimilarities between training and testing data. To address these limitations, a novel probabilistic approach that combines LUPI and active learning is proposed. A pool-based privileged active learning approach is presented for semi-supervising learning of human activities from multimodal labeled and unlabeled data.

In the last part of this dissertation, the LUPI paradigm is also investigated for solving biometric applications such as facial expression recognition. As facial image sequences may contain information for heterogeneous sources, facial data may be asymmetrically distributed between training and testing, as it may be difficult to maintain the same quality and quantity of information. To this end, a novel probabilistic classification method that combined the LUPI framework and conditional random fields is proposed to indirectly propagate knowledge from privileged to regular feature space. Each feature space owns specific parameter settings, which are combined together through a Gaussian prior, to train the proposed t-CRF+ model and allow the different tasks to share parameters and improve classification performance.