PhD Program in Sciences & Technologies For Electronics & Telecommunication

YEAR 2016

Curriculum in

Computer Vision, Pattern Recognition and Machine Learning

Visione Computazionale, Riconoscimento e Apprendimento Automatico

Research Themes

Theme A
Computer vision for behavioral analysis and activity recognition
Tutors: Vittorio Murino

Study and development of techniques and systems for the analysis of behaviours, actions, expressions/emotions, and social signals in general, referred to both single persons and groups. In this context, methods for tracking, recognition, and classification of persons and objects starting from images and/or sequences acquired from cameras distributed in the environment in several sparse locations, and from other types of sensors (e.g., microphones) will be considered. The main goal is to exploit hints and findings coming from social sciences to capture and model human behaviour. Computer vision and machine learning constitute the focus of this research.

Theme B
Computer vision for the prediction of human intentions
Tutors: Vittorio Murino

We want to go further the analysis of human movements and explicit behaviours to design methods and algorithms for the prediction of human intentions. Experiments are already running consisting in performing simple human-object interactions with a goal which is not explicitly stated nor viewed, and the task is to classify the actual intention subsuming that action from 2D and 3D data. We also plan to carry out other visual experiments considering more complex human interactions involving many subjects, while predicting the final outcome of the observed interplay. Computer vision and machine learning constitute the focus of this research.

Theme C
Part-based human body modeling for Socially-Aware Computer Vision
Tutors: Vittorio Murino, Alessio Del Bue

To recognize and interpret human nonverbal behavior it is fundamental to try to identify the subjects involved, especially in the wild, that is in real situations. To this end, part-based human body modeling is a mandatory task aimed at extracting from images the different components of the human body, like head, torso, arms, legs, etc., so as to estimate posture, gesture and gaze, all social cues widely known as useful hints to classify behavior and recognize situations. Further, real time tracking of body parts is equally important to increase such recognition performances, possibly adding prediction functionalities to these algorithms. Computer vision and machine learning methodologies are the main subjects of this study.

Theme D
Crowd behavioral analysis and event recognition
Tutors: Vittorio Murino

Study and development of techniques and systems for the analysis of behaviours, events, social signals in general, referred to a large mass of people (crowd). The analysis and modelling of behaviour of groups and crowd seen as single entities will be considered. There is evidence that large groups of people and crowd are characterised by a collective behaviour which may emerge in different situations and can lead to interesting outcome from the point of view of the surveillance applications, and may help to detect and predict coming events. Machine learning as well as computer vision constitute the focus of this research, starting from early work in human body modelling/tracking to novel social force models able to grasp the complex dynamics of the human flow.

Theme E
Re-identification
Tutors: Vittorio Murino

Study and development of biometric techniques for scene analysis and understanding. The research will mainly focus on person characterization, with possible focus on the usage of soft biometrics cues and in challenging conditions (e.g., crowd). The idea is to recover the identity of persons as viewed in different times and places, also considering face/body attributes, the so-called re-identification problem. Not only optical cameras will be used, but other information derived from different sensors may also be utilized (e.g., range, thermic). Moreover, super resolution techniques could be investigated to increase the resolution of images, particularly for recognition purposes, so as to improve the quality of the images and making them understandable for a human operator or a machine. Other possible options concern the use of a pan-tilt-zoom (PTZ) camera able to identify specific features of a single person or groups, or addressing non-cooperative face recognition at distance. The robustness to environmental (real) conditions and the non-cooperation of the subjects are the main features to which the developed techniques will have to cope with.

Theme F
Time-lapse Computer Vision for long-term learning
Tutor: Alessio Del Bue

Creating a computational model of the visual world requires tools for the understanding of a generic scene using images only. To this end, the aim of this project is to reconstruct the 3D information from a video stream given by a single camera (e.g. outdoor webcams, video surveillance cameras). This problem is apparently very complex to solve using a single viewpoint. However, even with a single view, it is possible to extract 3D cues from the continuous streaming of video (days/weeks/months). In particular, we will leverage photometric cues (gradual change of lighting) and motion patterns of dynamic objects (from people detections and motion flows) in order to infer the dynamic 3D structure of the scene. In this project, we will both use the expertise of PAVIS in 3D Reconstruction, Photometric Stereo and Machine Learning in order to define novel formalisations of this problem. The research will go towards the creation of algorithms working in realistic conditions using the 24h/7d camera network infrastructure of IIT and the high-performance computing infrastructure of PAVIS.

Theme G
Semantic 3D scene reconstruction and modelling
Tutor: Alessio Del Bue

Classical multi-view geometry problems make use of point correspondences to infer objects dynamic and to estimate the 3D structure of a generic scene. This neglects completely the semantic of the scene that gives important cues about objects dynamics and their spatial configuration. This theme will develop novel tools for the reformulation of multi-view geometry methods in order to include semantic given by object detectors, image segmentation and recognition approaches. To this end, this project will use the output of deep architectures (object positions, region labels) to make hypothesis about 3D geometry. The aim is to make more robust current techniques for 3D reconstruction and motion segmentation when applied to real world problems. This research will study and research of different computational tools such as (Bayesian) factorization, time series analysis, convex and non-linear optimization methods. Possibly, some of the developed techniques might be applied to robotic platforms available in IIT for tasks such as visual navigation and object pose estimation.

Theme H
Sensing humans: enhancing social abilities of the iCub platform
Tutor: Alessio Del Bue, Lorenzo Natale

There is general consensus that robots in the future will work in close interaction with humans. This requires that robots are endowed with the ability to detect humans and interact with them. However, treating humans as simple animated entities is not enough: meaningful human-robot interaction entails the ability to interpret social cues. The aim of this project is to endow the iCub with a fundamental layer of capabilities for detecting humans, their posture and social behaviour. Examples could be the ability to detect if a person is attempting to interact with the robot and to react accordingly. This requires a new set of computational tools based on Computer Vision and Machine Learning to detect people at close distance. On the other hand, this face-to-face interaction requires developing novel algorithms for coping with situations in which large areas of the body are occluded or only partially visible.
Requirements: This PhD project will be carried out within the Visual Geometry and Modelling Lab in the PAVIS department and in collaboration with the iCub Facility. The ideal candidate should have a degree in Computer Science or Engineering (or equivalent) and background in Computer Vision and/or Machine Learning. He should also be highly motivated to work on a robotic platform and have strong computer programming skills.

Theme I
Biomedical imaging
Tutor: Diego Sona, Vittorio Murino

The wide adoption of biomedical sensors (e.g., MRI, TAC, SPECT, MEG, EEG/EMG, Electron and Fluorescent Microscopy, etc.) in various clinical and biological investigations is fostering an increasing interest in advanced tools supporting the expert in the analysis and interpretation of the produced big amount of data. In this perspective, this theme will address all lines of research related to the development of computer aided detection (CAD) systems, ranging from image processing and analysis, from object detection to automatic determination of diseases' biomarkers, etc.. Particular attention will be devoted to structural data (3D) and/or functional data (multivariate time-series and videos). The development of such CAD tools will require the design of novel computer vision and pattern recognition techniques for advanced and automatic analysis of biomedical data.

Theme L
Connectomics
Tutor: Diego Sona, Vittorio Murino

The brain is a complex interconnected system that can be investigated at different granularities, from the macro-scale level (i.e., functional interactions between brain areas) using non-invasive techniques (e.g., MRI, MEG, EEG, etc.) down to the meso-scale level (functional interactions between neurons in a complex network), thanks to the recent evolution of high-density multi-electrode arrays (MEAs). Nevertheless, independently of the granularity, the relationships between the functional dependencies and the structural connectivity remains still unsolved. In this research theme, the topic under investigation is therefore the integration of functional and structural connectivity information with different applications ranging from the characterization of mental diseases' biomarkers (e.g., schizophrenia, autism, multiple sclerosis, etc.) to studies for the analysis of treatments, from the investigation of cognitive functions to their correlation with behavioral patterns. The integration of functional and structural connectivity information will require the development of tools for the analysis of multi-modal data at macro-scale (e.g., functional MRI, diffusion imaging, etc.) and/or meso-scale (MEA electrical activity, fluorescent microscopy imaging). Particular attention will be devoted to machine learning and pattern recognition techniques aiming at integrating/fusing these multimodal source of information.

Theme M
Animal behavior analysis
Tutors: Diego Sona, Vittorio Murino

Behavioural neuroscience is a fundamental research field studying the biological bases of behavior providing insights into the mechanisms of the nervous systems producing anomalous behavior, and experimental subjects mostly involve animals. We are, therefore, involved in a multidisciplinary research activity, which need the development of techniques and systems for the automatic analysis of actions, postures and social behaviours of mice in home cages. In this framework, we aim at designing methods for tracking 24/7 the mice and objects from video recorded from multiple camera with different orientations. Computer vision and machine learning constitute the focus of this research, with particular interest in methodologies exploiting the spatio-temporal information.