

2011 CVPR - Accepted Tutorials and Short Courses

Title	Type	Duration	Proposers	Schedule
Game Theory in Computer Vision and Pattern Recognition	Regular	4h	M. Pelillo and A. Torsello	June 20th, Monday AM
Image and Video Description with Local Binary Pattern Variants	Regular	3h	M. Pietikäinen and J. Heikkilä	June 24th, Friday PM
Tools and Methods for Image Registration	Regular	Full Day	M. Brown, G. Carneiro, A. A. Farag, E. Hancock, A. A. Goshtasby (Organizer), J. Matas, J.M. Morel, N. S. Netanyahu, F. Sur, and G. Yu	June 24th, Friday AM-PM
Frontiers of Human Activity Analysis	Regular	3.5h	J. K. Aggarwal, M. S. Ryoo, and K. Kitani	June 20th, Monday PM
Diffusion Geometry Methods in Shape Analysis	Regular	3h	A. Brontein and M. Bronstein	June 20th, Monday AM
ITK meets OpenCV: A New Open Source Software Resource for CV	Regular	6h (HandsOn)	L. Ibanes, A. Perera, P. Reynolds, M. Leotta	June 20th, Monday PM
Structured Prediction and Learning in Computer Vision	Regular	4-5h	S. Nowozin and C. Lampert	June 25th, Saturday AM
Light Fields in Computational Photography	Advanced	3h	R. Raskar and R. Horstmeyer	June 24th, Friday AM

Details

(1) Game Theory in Computer Vision and Pattern Recognition (M. Pelillo and A. Torsello)

The development of game theory in the early 1940's by John von Neumann was a reaction against the then dominant view that problems in economic theory can be formulated using standard methods from optimization theory. Indeed, most real- world economic problems typically involve conflicting interactions among decision-making agents that cannot be adequately captured by a single (global) objective function, thereby requiring a different, more sophisticated treatment. Accordingly, the main point made by game theorists is to shift the emphasis from optimality criteria to equilibrium conditions. As it provides an abstract theoretically-founded framework to elegantly model complex scenarios, game theory has found a variety of applications not only in economics and, more generally, social sciences but also in different fields of engineering and information technologies. In particular, in the past there have been various attempts aimed at formulating problems in computer vision, pattern recognition and machine learning from a game-theoretic perspective and, with the recent development of *algorithmic game theory*, the interest in these communities around game-theoretic models and algorithms is growing at a fast pace. The goal of this tutorial is to offer an introduction to the basic concepts of game theory and to provide a critical overview of its main applications in computer vision and pattern recognition. We shall assume no pre-existing knowledge of game theory by the audience, thereby making the tutorial self-contained and understandable by a non-expert.

(2) Image and Video Description with Local Binary Pattern Variants (M. Pietikäinen and J. Heikkilä)

This tutorial presents effective image and video descriptors based on recent variants of the highly popular local binary pattern (LBP) texture operator. Part I overviews the basic LBP operators in spatial and spatiotemporal domains. Part II provides an overview of recent LBP variants which improve the discriminative power and robustness of the original LBP. It also describes in more detail the local phase quantization operator (LPQ), which has the property of being highly insensitive to image blurring but also retains excellent discrimination power both with sharp and blurry images. Part III presents examples of using different LBP variants in important computer vision problems, including face and facial expression recognition, biomedical image analysis, motion analysis, recognition of actions and gait, and video texture synthesis. Finally, Part IV concludes the tutorial and presents some directions for future research.

(3) Tools and Methods for Image Registration (M. Brown, G. Carneiro, A. A. Farag, E. Hancock, A. A. Goshtasby (Organizer), J. Matas, J.M. Morel, N. S. Netanyahu, F. Sur, and G. Yu)

This tutorial covers the fundamentals of image registration, including similarity and affine invariant point detectors and descriptors, design of image descriptors for various vision tasks, similarity and affine invariant features, feature learning and optimal features, point correspondence algorithms under various geometric constraints, and various transformation functions for image registration. Also to be covered are registration of temporal, multiview, and multimodality images, as well as applications of image registration in medical imaging and remote sensing.

(4) Frontiers of Human Activity Analysis (J. K. Aggarwal, M. S. Ryoo, and K. Kitani)

Human activity recognition is an important area of computer vision research. The goal of the activity recognition is an automated analysis (or interpretation) of ongoing events and their context from video data. Its applications include surveillance systems, patient monitoring systems, and a variety of systems that involve interactions between persons and electronic devices such as human-computer interfaces. Most of these applications require recognition of high-level activities, often composed of multiple simple (or atomic) actions of persons. This tutorial provides a detailed overview of various state-of-the-art research papers on human activity recognition. We discuss both the methodologies developed for simple individual-level activities and those for high-level multi-person (and object) interactions. An approach-based taxonomy is chosen, comparing the advantages and limitations of each approach. We first review early history of human activity recognition briefly, and discuss methodologies designed for recognition of activities of individual persons. Approaches utilizing space-time volumes and/or sequential models are covered. Next, hierarchical recognition methodologies for high-level activities are presented and compared. We categorize human activities into human actions, human-human interactions, human-object interactions, and group activities, discussing approaches designed for their recognition. Hierarchical state-based approaches and syntactic approaches that interpret videos in terms of stochastic strings are covered. Finally, we discuss description-based approaches that analyze videos by maintaining their knowledge on activities' temporal, spatial, and logical structures. Recent video datasets designed to encourage human activity recognition research will be discussed as well. This tutorial will provide the impetus for future research in more productive areas.

(5) Diffusion Geometry Methods in Shape Analysis (A. Brontein and M. Bronstein)

Over the last decade, 3D shape analysis has become a topic of increasing interest in the computer vision community. Nevertheless, when attempting to apply current image analysis methods to 3D shapes (feature-based description, registration, recognition, indexing, etc.) one has to face fundamental differences between images and geometric objects. Shape analysis poses new challenges that are non-existent in image analysis. The purpose of this tutorial is to overview the foundations of shape analysis and to formulate state-of-the-art theoretical and computational methods for shape description based on their intrinsic geometric properties. The emerging field of diffusion geometry provides a generic framework for many methods in the analysis of geometric shapes objects. The tutorial will present in a new light the problems of shape analysis based on diffusion geometric constructions such as manifold embeddings using the Laplace-Beltrami and heat operator, heat kernel local descriptors, diffusion and commute-time metrics.

(6) ITK meets OpenCV: A New Open Source Software Resource for CV (L. Ibanes, A. Perera, P. Reynolds, M. Leotta)

This tutorial is a hands-on introduction to the combined use of the Insight Toolkit (ITK) and the OpenCV library. ITK is an image segmentation and registration toolkit developed by a consortium of partners from industry and academia, and funded mainly by the US National Library of Medicine. Although ITK has been traditionally used in medical imaging applications, the ITK toolkit provides hundreds of image processing filters, as well as modules for performing segmentation and registration of N-Dimensional images that are generic and usable for non-medical applications. Distributed under an Apache 2.0 License, ITK is an open source resource supported by a vibrant community. During the past year, ITK has been revamped to produce version 4.0 that includes improved support for video applications. As part of that effort, a software bridge has been developed to interconnect ITK with the OpenCV library, making it very easy for application developers and researchers to take advantage of the hundreds of algorithms that are available in ITK, without having to migrate away from their existing OpenCV-based applications. This tutorial will introduce participants to the wealth of algorithms available in the ITK toolkit, and will provide hands-on exercises on how to use ITK from OpenCV applications. Attendees will receive a virtual appliance (VirtualBox) with the full ITK+OpenCV build environment pre-configured, and will execute programming exercises ranging from the introductory level to the intermediate level. The material, including the virtual appliance, will be distributed in pre-configured USB memory sticks.

(7) Structured Prediction and Learning in Computer Vision (S. Nowozin and C. Lampert)

Powerful statistical models that can be learned efficiently from large amounts of data are currently revolutionizing computer vision. These models possess rich internal structure reflecting task-specific relations and constraints. This tutorial introduces the reader to the most popular classes of structured prediction models in computer vision. This includes discrete graphical models which we cover in detail together with a description of algorithms for both probabilistic inference and maximum a posteriori (MAP) inference. We discuss separately recently successful techniques for prediction in general structured models. In the second part of this tutorial we describe methods for parameter learning. We distinguish the classic maximum likelihood based methods, such as conditional random fields, from the more recent prediction-based parameter learning methods, such as structured output support vector machines. We highlight recent developments to enrich models such as kernelization and latent variable models. Throughout the tutorial we provide examples of successful application of the methods in the computer vision.

(8) Light Fields in Computational Photography (R. Raskar and R. Horstmeyer)

Computational photography involves optical processing as well as digital image processing. The concepts are often represented via higher dimensional data structures. The ray-based 4D lightfield representation, based on simple 3D geometric principles, has led to a range of new algorithms and applications in Computer Vision and Graphics. They include digital refocusing, depth estimation, synthetic aperture, image stabilization and glare reduction within a camera or using an array of cameras. The lightfield representation is, however, inadequate to describe interactions with diffractive or phase-sensitive optical elements. Fourier optics principles are used to represent wavefronts with additional phase information. This course reviews the current and future directions in exploiting higher dimensional representation of light transport. We hope the course will inspire researchers in computer vision comfortable with ray-based analysis to develop new tools and algorithms based on joint exploration of geometric and wave optics concepts. The notes are aimed at readers familiar with the basics of computer vision and computational photography.