Advanced Topics in Computer Vision: Segmentation, Reconstruction and Recognition of Dynamic Scenes (580.464)

Instructor: René Vidal, Phone: 510-516-7306, E-mail: rvidal@cis.jhu.edu

MAT 4:30-6pm, Shaffer 202, Spring 2005

This class will cover state-of-the-art methods in dynamic vision, with an emphasis on segmentation, reconstruction and recognition of static and dynamic scenes. Topics include: reconstruction of static scenes (tracking and correspondences, multiple view geometry, self calibration), reconstruction of dynamic scenes (2-D and 3-D motion segmentation, nonrigid motion analysis), recognition of visual dynamics (dynamic textures, face and hand gestures, human gaits, crowd motion analysis), as well as geometric and statistical methods for clustering and unsupervised learning, such as K-means, Expectation Maximization, and Generalized Principal Component Analysis.

Part I. **Image Formation, Feature Extraction, Matching and Correspondences** (4 lectures)

- Feb 1,2 *Image Formation*: photometry (radiance, irradiance, BRDF), geometry (3-D surfaces, pinhole camera, orthographic, affine, perspective and paracatadioptric projection model), dynamics (rigid-body motions SO(3), SE(3), so(3), se(3), twists and exponential maps).
- Feb 8,9 *Feature Extraction, Matching and Correspondences*: deformation models (2-D translational and 2-D affine), brightness constancy constraint, tracking and optical flow, aperture problem, feature extraction, matching and correspondences (sum of square differences, normalized cross correlation), mosaicing.

Part II. **Reconstruction of Static Scenes [2]** (9 lectures)

- Feb 15,16 *Reconstruction from Two Calibrated Views*: epipolar geometry, epipolar lines, epipoles, essential matrix, essential manifold, pose recovery via eight-point algorithm, planar homographies, structure recovery via triangulation (linear and optimal).
- Feb 22,23 *Reconstruction from Two Uncalibrated Views*: uncalibrated epipolar geometry, the fundamental matrix, transfer properties of the fundamental matrix, camera calibration with a rig, camera self-calibration, Kruppa equations.
- Mar 1 *Reconstruction from Three Views*: the trilinear constraint, the trifocal tensor, transfer properties of the trifocal tensor, the seven-point algorithm, estimation of fundamental and camera matrices from the trifocal tensor. (Read [1] part III)
- Mar 2,8,9 *Reconstruction from Multiple Views*: the multiple view matrix, multilinear constraints, optimal reconstruction from multiview normalized epipolar constraint, minimizing reprojection error, multiframe factorization for affine and perspective cameras.
- Mar 15,16 *SPRING BREAK*
- Mar 22 *MIDTERM I*: all material covered so-far.

Part III. **Clustering and Generalized Principal Component Analysis [3]** (4 lectures)

- Mar 23,29 *Central Clustering*: K-means, Expectation Maximization, Spectral Clustering.
- Mar 30, Apr 5 *Generalized Principal Component Analysis (GPCA)*: line clustering, plane clustering, subspace arrangements and algebraic varieties, Principal Component Analysis (PCA), Probabilistic PCA (PPCA), GPCA, K-subspaces, EM for mixtures of subspaces.

Part IV. **Reconstruction of Dynamic Scenes [3]** (6 lectures)

- Apr 6 *Segmentation of Linear Motion Models from Two Views*: 2-D translational from image intensities, 2-D translational, 2-D similarity, 2-D affine and 3-D translational from point correspondences or optical flow.
- Apr 12,13 *Segmentation of Bilinear Motion Models from Two Views*: multibody brightness constancy constraint, multibody affine matrix, multibody epipolar constraint, multibody fundamental matrix, generalized 8-point algorithm, multibody homography.
- Apr 19 *Motion Segmentation from Three Perspective Views*: multibody trilinear constraint, multibody trifocal tensor, estimation and factorization of the multibody trifocal tensor.
- Apr 20 *Motion Segmentation from Multiple Views*: motion subspaces, multibody factorization algorithms for affine cameras, multibody factorization from optical flow in perspective and central panoramic cameras.
- Apr 26 *Reconstruction of Nonrigid Motions*: shape basis, shape constraints, motion constraints, non-rigid factorization algorithms.
- April 27 *MIDTERM 2*: all material from Midterm 1 till now.

Part V. **Recognition of Visual Dynamics**

- May 3,4 *Modeling, Classification, Segmentation and Recognition of Dynamic Textures, Hand Gestures, Human Gaits and Crow Motion*: linear dynamical models, system identification (N4SID), metrics on the space of dynamical models, nearest neighbor classification and recognition of dynamical models.
- May 11 *PRESENTATION OF FINAL PROJECTS*

References

[1] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge, 2nd edition, 2004.

[2] Y. Ma, S. Soatto, J. Kosecka, and S. Sastry. *An Invitation to 3D Vision: From Images to Geometric Models.* Springer Verlag, 2003.

[3] R. Vidal, Y. Ma, and S. Sastry. *Generalized Principal Component Analysis.* In prep., 2005.

Administrative

Office Hours: 1 hour after each lecture. Office hours will be held in 308B Clark Hall.

Honor system/Grading Policy: Homeworks, midterms and projects will be individual. The strength of the university depends on academic and personal integrity. In this course, you must be honest and truthful. Ethical violations include cheating on exams, plagiarism, reuse of assignments, improper use of the Internet and electronic devices, unauthorized collaboration, alteration of graded assignments, forgery and falsification, lying, facilitating academic dishonesty, and unfair competition. All these will be severely penalized.

(a) *Homeworks*: There will one homework every other week (approximately). Homework problems will include both analytical exercises as well as programming assignments in MATLAB. Homeworks will count toward 30% of the final grade.

(b) *Midterms*: There will be two midterms in mid March and end of April. Midterm problems will include both analytical exercises as well as programming assignments in MATLAB. Each midterm will count toward 20% of the final grade.

(c) *Project*: There will be a final project. Each student will either apply techniques from the course to solve a real problem or solve an open research problem in dynamic vision. Each student will write a report in LaTeX following the authors instructions in the CVPR 2005 website and give a 20 minute presentation during the final examination period. The project will count toward 30% of the final grade.