Robert Pless; Research Summary, 2000-2005

Today, imaging is ubiquitous. In fields from robotics to surveillance to medicine, imaging is the primary sensor for more and more applications. With the ongoing development of better and cheaper sensors, visual processing will take a central role in multi-sensor networks and will find yet more medical applications. But images from real cameras in real environments are complex; to date, success in vision applications has come by developing explicit models for specific applications. This is time consuming and expensive. Instead, my research explores data driven tools to simplify video analysis. I develop algorithms that use thousands of images from a particular sensor, environment or patient, to automatically infer relevant, context specific representations. The highlights of the last several years include: (a) a geometric theory of structure from motion for non-parametric camera models, (b) non-parametric, spatio-temporal statistical models of image variation for extended surveillance applications, and (c) fundamental algorithms for image manifold learning with a specific focus on cardiac MRI imagery, as well as a collection of smaller research efforts in support of collaborative projects.

Generalized Camera Geometry. A wide variety of new imaging systems have been proposed for omni-directional vision; these system designs use fish-eye lenses, cameras viewing spherical or parabolic mirrors, or multi-camera systems. I developed the first unifying geometric theory for understanding moving images from all of these camera types and derived tools to design camera systems that are optimized for particular environments.

Specifically, I derived the structure from motion equations for the generalized imaging model of Nayar and Grossberg [GN05]. This gives a corollary to the standard epipolar constraints [2] and I further derived an analytic form for the sensitivity-to-noise of these constraints [4]. These constraints have been used by others to derive tools for generalized camera calibration [SR04, MP04b, RSL05, CF05], and algorithmic optimizations were proposed to solve these constraints within the context of structure from motion systems [Nis04, MP04a, GA04, Stu05]. These models have been applied to autonomous navigation and model building in the case of a multi-camera system (treated as a single generalized camera) [CC04, LS04, SOÅ05].

Robust, Real-Time Surveillance. In the area of surveillance, I promote a framework I call “Passive Vision”, which builds statistical models of scene appearance over long time periods (minutes to days). Within this model, I have addressed open questions in outdoor surveillance including anomaly detection, tracking, scene annotation, and characterizing patterns of background motion. My algorithms are explicitly designed to be pre-processing modules so that can replace naive background subtraction algorithms to immediately fix problems faced by existing high-level surveillance algorithms in real outdoor environments.

Specifically, I focus on spatio-temporal image derivatives because these measurements of image variation are easy to compute and somewhat invariant to large lighting changes. For fixed viewpoint surveillance cameras, accumulating image derivative distributions gives a simple characterization of complicated background motions. This offers novel tools to find anomalies, segment foreground objects in outdoor scenes [9, 7, 14], and detect global patterns in the local image motions [16]. These local statistics have also turned out to be robust tools for detecting roads in stabilized aerial video [8, 6].

This work is among the first to consider explicitly the extended surveillance problem as an opportunity to characterize the local, spatio-temporal statistical variation of the specific scene in view. It demonstrates that surprisingly simple, local, incrementally updated models are sufficient to detect anomalies in scenes with complicated background dynamics. The work assumes that the background measurements are drawn from a stationary stochastic model, and this strict probabilistic interpretation has continued to be developed by others [MG05, SS05, TID05]. My study of “Passive Vision” is ongoing, and forms one of two major future research directions. Preliminary results indicate that both tracking and video coding algorithms are likely to become more efficient by a factor of 10 when they fully utilize the statistics of motion and image variation captured by a static camera.

Manifold Learning. Manifold Learning is a tool for organizing high-dimensional data that has just a few degrees of freedom. A key example of this is medical imagery of parts of the body that move such as video of a person walking or MR-imagery of the heart or lung, each of which have images which vary in a complicated but recurring patterns. I have specialized Manifold Learning algorithms to be effective for imagery, solved several key technical problems to make them effective for data sets that vary cyclically, and used manifold constraints to make medical image segmentation more accurate and robust.

I was among the earliest popularizers of the Isomap algorithm, using it for pose estimation of unknown objects [11] and video classification [3]. Although recent work has directly applied our approach to estimate illumination direction in graphics applications [WMTG05], classical manifold learning algorithms often fail on noisy, natural images. With my student Richard Souvenir, I have addressed three major shortcomings. First, we used Pattern Theory to define application specific image distance functions [5, 12]. Second, we modified Isomap to find minimal parameterizations for data sets that are sampled from cyclic manifolds (circles, spheres, tori, and their higher dimensional correlates) [10]. This class of images includes nearly all periodic, biological data sets and is recognized as a key limitation of current approaches [SR03, WS04]. Third, we developed an Expectation Minimization algorithm to support clustering and parameterization of data sets that are drawn from multiple intersecting manifolds [13], which is common in motion capture and animation data, and breaks (to some extent) all current Manifold Learning algorithms.

One application of image manifolds is in segmenting and labeling cardiac MRI data, in which images vary due to the patient breathing and heartbeat. In recent work my student Qilong Zhang and I have reformulated both Active Contour and Level Set segmentation algorithms to exploit the additional constraints offered by the manifold parameterization [24, 25]. This reduces by a factor of 30 the amount of time that a Radiology Fellow takes to measure left-ventricle wall thickness in diagnostic studies, and formal verification of our segmentation relative to hand-segmented images is ongoing.

Collaborative Projects. I have been privileged to work with fantastic colleagues at Washington University and have collaborated on a number of interdisciplinary research projects. Significant contributions include (a) tools to allow a multi-sensor robot to recalibrate after vibrations or collisions cause sensors to be misaligned, and (b) algorithms to extend the lifetime of ad-hoc sensor networks.

To support Media and Machines Laboratory efforts in making robust, fault-tolerant robot navigation systems, I worked to develop a framework for ego-motion estimation which accepts motion measurements from arbitrary motion sensors (e.g. cameras, GPS, accelerometer, etc.) and specifies a probabilistic model of data fusion [23]. I provide accurate priors to support classical Bayesian stereo within this framework by experimentally measuring stereo match errors across many surface types [1]. Furthermore, combining data across multiple sensors requires accurate calibration, and in real environments this calibration must continually be corrected or re-calculated. My student Qilong Zhang and I worked to develop the first multi-sensor auto-calibration tools [21], with a special focus on a camera and a laser range finder [22]. This approach has been used by others as the basis for algorithms for auto-calibration of other sensors [DFK05].

My efforts in sensor networks are to make them last longer with limited power; I focus on graph theoretic optimizations to schedule sleep cycles in sensor networks that must ensure sensor coverage and communication connectivity. Our theoretical treatment discovers that when the communication range is at least twice the sensing range, then the sensing problem dominates (making the optimization problem simpler) [15, 20], and greedy distributed communications strategies that are efficient have tightly bounded dilation [17, 18]. We also made one of the first efforts to include more accurate probabilistic models of sensing and communication within the network design problem [19].

Selected publications of mine.

[1] J. Larson and R. Pless. Bayesian stereo: 3d vision designed for sensor fusion. In *Intelligent Robots and Computer Vision XXII: Algorithms, Techniques, and Active Vision*, volume 5608, pages 198–206, 2004.

[2] R. Pless. Two view discrete and differential constraints for generalized imaging systems. In *Proc. of the IEEE Workshop on Omnidirectional Vision*, 2002.

[3] R. Pless. Using isomap to explore video sequences. In *Proc. International Conference on Computer Vision (ICCV)*, pages 1433–1440, 2003.

[4] R. Pless. Using many cameras as one. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 587–593, 2003.

[5] R. Pless. Differential structure in non-linear image embedding functions. In *Proc. the IEEE Workshop on Articulated and non-rigid Motion*, pages 10–17, 2004.

[6] R. Pless. Detecting roads in stabilized video with the spatio-temporal structure tensor. *Geoinformatica*, to appear, 2005.

[7] R. Pless. Spatio-temporal background models for outdoor surveillance. *Journal on Applied Signal Processing*, pages 2281–2291, 2005.

[8] R. Pless and D. Jurgens. Road extraction from motion cues in aerial video. In *Proceedings of the ACM Conference on Geographic Information Systems*, pages 31–38, 2004.

[9] R. Pless, J. Larson, S. Siebers, and B. Westover. Evaluation of local models of dynamic backgrounds. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2003.

[10] R. Pless and I. Simon. Embedding images in non-flat spaces. In *Conference on Imaging Science Systems and Technology*, pages 182–188, 2002.

[11] R. Pless and I. Simon. Using thousands of images of an object. In *Proc. IEEE Conference on Computer Vision, Pattern Recognition and Image Processing (CVPRIP)*, 2002.

[12] R. Souvenir and R. Pless. Isomap and non-parametric models of image deformation. In *Proc. IEEE Workshop on Visual Motion*, pages 195–200, 2005.

[13] R. Souvenir and R. Pless. Manifold clustering. In *Proc. International Conference on Computer Vision (ICCV)*, 2005.

[14] R. Souvenir, J. Wright, and R. Pless. Spatio-temporal detection and isolation: Results on the pets2005 datasets. In *Proc. IEEE Workshop on Performance Evaluation in Tracking and Surveillance*, 2005.

[15] X. Wang, G. Xing, Y. Zhang, C. Lu, R. Pless, and C. Gill. Integrated coverage and connectivity configuration in wireless sensor networks. In *SenSys '03: Proceedings of the 1st international conference on Embedded networked sensor systems*, pages 28–39, 2003.

[16] J. Wright and R. Pless. Analysis of persistent motion patterns using the 3d structure tensor. In *Proc. IEEE Workshop on Visual Motion*, pages 14–19, 2005.

[17] G. Xing, C. Lu, R. Pless, and Q. Huang. On greedy geographic routing algorithms in sensing-covered networks. In *MobiHoc '04: Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, pages 31–42, 2004.

[18] G. Xing, C. Lu, R. Pless, and Q. Huang. Impact of sensing coverage on greedy geographic routing algorithms. *ACM Transaction on Sensor Networks*, in press, 2005.

[19] G. Xing, C. Lu, R. Pless, and J. A. O'Sullivan. Co-grid: an efficient coverage maintenance protocol for distributed sensor networks. In *IPSN'04: Proceedings of the third international symposium on Information processing in sensor networks*, pages 414–423, 2004.

[20] G. Xing, X. Wang, Y. Zhang, C. Lu, R. Pless, and C. Gill. Integrated coverage and connectivity configurations for energy conservation in sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, in press, 2005.

[21] Q. Zhang and R. Pless. Constraints for heterogeneous sensor auto-calibration. In *IEEE Workshop on Realtime 3D Sensors and Their Use*, pages 38–43, 2004.

[22] Q. Zhang and R. Pless. Extrinsic calibration of a camera and laser range finder. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS)*, pages 2301–2306, 2004.

[23] Q. Zhang and R. Pless. Fusing video and sparse depth data in structure from motion. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pages 3403–3406, 2004.

[24] Q. Zhang and R. Pless. Segmenting cardiopulmonary images using manifold learning with level sets. In *IEEE Workshop on Computer Vision for Biomedical Image Applications*, 2005.

[25] Q. Zhang, R. Souvenir, and R. Pless. Segmentation informed by manifold learning. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMM-CVPR)*, 2005.

Selected Publications Citing My Work

The following list gives recent papers in premier conferences and journals that cite the work described in my research statement.

[CC04] W.Y. Chang and C.S. Chen. Pose estimation for multiple camera systems. In *Proc. International Conference on Pattern Recognition (ICPR)*, pages III: 262–265, 2004.

[CF05] David Claus and Andrew W. Fitzgibbon. A rational function lens distortion model for general cameras. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2005.

[DFK05] R Dupont, P Fuchs, and R Keriven. An improved calibration technique for coupled single-row telemeter and ccd camera. In *3-D Digital Imaging and Modeling (3DIM), 2005*, 2005.

[GA04] Nuno Gonçalves and Helder Araújo. Projection model, 3d reconstruction and rigid motion estimation from non-central catadioptric images. In *Intl. Symposium on 3D Data Processing, Visualization and Transmission (3DPVT 2004)*, pages 325–332, 2004.

[GN05] Michael D. Grossberg and Shree K. Nayar. The raxel imaging model and ray-based calibration. *International Journal of Computer Vision*, 61(2):119–137, 2005.

[LS04] Anat Levin and Richard Szeliski. Visual odometry and map correlation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 611–618, 2004.

[MG05] Joshua Migdal and W. Eric L. Grimson. Background subtraction using markov thresholds. In *Proc. IEEE Workshop on Visual Motion*, pages 58–65, 2005.

[MP04a] M. Menem and T. Pajdla. Constraints on perspective images and circular panoramas. In *Proceedings of the British Machine Vision Conference, (BMVC)*, 2004.

[MP04b] Branislav Micusík and Tomás Pajdla. Autocalibration & 3d reconstruction with non-central catadioptric cameras. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 58–65, 2004.

[Nis04] David Nistér. A minimal solution to the generalised 3-point pose problem. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 560–567, 2004.

[RSL05] Srikumar Ramalingam, Peter Sturm, and Suresh Lodha. Towards complete generic camera calibration. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 1093–1098, Jun 2005.

[SOÅ05] Henrik Stewenius, Magnus Oskarsson, and Kalle Åström. Reconstruction from planar motion image sequences with applications for autonomous vehicles. In *SCIA*, pages 609–618, 2005.

[SR03] Lawrence K. Saul and Sam T. Roweis. Think globally, fit locally: unsupervised learning of low dimensional manifolds. *J. Mach. Learn. Res.*, 4:119–155, 2003.

[SR04] Peter Sturm and Srikumar Ramalingam. A generic concept for camera calibration. In *Proc. European Conference on Computer Vision (ECCV)*, volume 2, pages 1–13. Springer, May 2004.

[SS05] Yaser Sheikh and Mubarak Shah. Bayesian object detection in dynamic scenes. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 74–79, 2005.

[Stu05] Peter Sturm. Multi-view geometry for general camera models. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 206–212, jun 2005.

[TID05] Leonid Taycher, John W. Fisher III, and Trevor Darrell. Incorporating object tracking feedback into background maintenance framework. In *Proc. IEEE Workshop on Visual Motion*, pages 120–125, 2005.

[WMTG05] Holger Winnemöeller, Ankit Mohan, Jack Tumblin, and Bruce Gooch. Light waving: Estimating light positions from photographs alone. *Computer Graphics Forum*, 24(3):to appear, 2005.

[WS04] K. Q. Weinberger and L. K. Saul. Unsupervised learning of image manifolds by semidefinite programming. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume II, pages 988–995, Washington D.C., 2004.