Statistical Methods and Models for Video-Based Tracking, Modeling, and Recognition

Statistical Methods and Models for Video-Based Tracking, Modeling, and Recognition

Rama Chellappa

UMIACS, University of Maryland
College Park, MD
rama@cfar.umd.edu

Aswin C. Sankaranarayanan

Rice University
Houston, TX
saswin@rice.edu

Ashok Veeraraghavan

Mistubishi Electric Research Laboratory
Cambridge, MA
veerarag@merl.com

Pavan Turaga

UMIACS, University of Maryland
College Park, MD
pturaga@cfar.umd.edu

the essence of knowledge

Boston – Delft

Foundations and Trends® in Signal Processing

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is R. Chellappa, A. C. Sankaranarayanan, A. Veeraraghavan and P. Turaga, Statistical Methods and Models for Video-Based Tracking, Modeling, and Recognition, Foundations and Trends® in Signal Processing, vol 3, nos 1–2, pp 1–151, 2009

ISBN: 978-1-60198-314-5

Foundations and Trends® in Signal Processing
Volume 3 Issues 1–2, 2009

Editorial Scope

Foundations and Trends® in Signal Processing will publish survey and tutorial articles on the foundations, algorithms, methods, and applications of signal processing including the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital and multirate signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations
- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing
- classification and detection
- estimation and regression
- tree-structured methods

Information for Librarians

Foundations and Trends® in Signal Processing, 2009, Volume 3, 4 issues. ISSN paper version 1932-8346. ISSN online version 1932-8354. Also available as a combined paper and online subscription.

Foundations and Trends® in
Signal Processing
Vol. 3, Nos. 1–2 (2009) 1–151

DOI: 10.1561/2000000007

Statistical Methods and Models for Video-Based Tracking, Modeling, and Recognition

Rama Chellappa[1], Aswin C. Sankaranarayanan[2], Ashok Veeraraghavan[3] and Pavan Turaga[4]

[1] *Department of Electrical and Computer Engineering, Center for Automation Research, UMIACS, at University of Maryland, College Park, MD, rama@cfar.umd.edu*
[2] *Department of Electrical and Computer Engineering, Rice University, Houston, TX, saswin@rice.edu*
[3] *Mistubishi Electric Research Laboratory, Cambridge, MA, veerarag@merl.com*
[4] *Department of Electrical and Computer Engineering, Center for Automation Research, UMIACS, at University of Maryland, College Park, MD, pturaga@cfar.umd.edu*

Abstract

Computer vision systems attempt to understand a scene and its components from mostly visual information. The geometry exhibited by the real world, the influence of material properties on scattering of incident light, and the process of imaging introduce constraints and properties that are key to interpreting scenes and recognizing objects, their structure and kinematics. In the presence of noisy observations and other uncertainties, computer vision algorithms make use of statistical methods for robust inference. In this monograph, we highlight

the role of geometric constraints in statistical estimation methods, and how the interplay between geometry and statistics leads to the choice and design of algorithms for video-based tracking, modeling and recognition of objects. In particular, we illustrate the role of imaging, illumination, and motion constraints in classical vision problems such as tracking, structure from motion, metrology, activity analysis and recognition, and present appropriate statistical methods used in each of these problems.

Contents

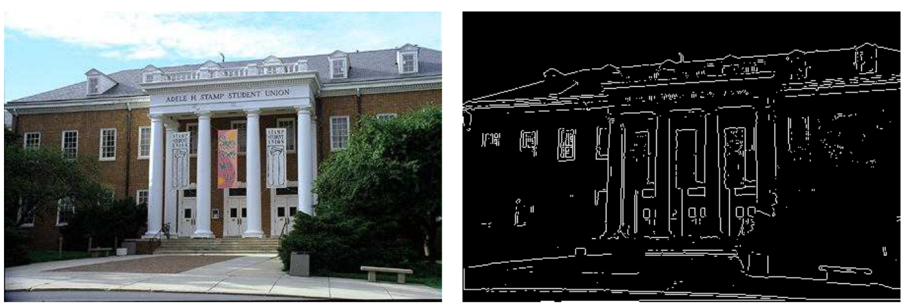
1

Introduction

The goal of computer vision is to enable machines to see and interpret the world. Computer vision algorithms use input from one or more still images or video sequences that are related in a specific manner. The distribution of intensities and their spatial and temporal arrangements in an image or a video sequence contain information about the identity of objects, their reflectance properties, scene structure, and objects in the scene. However, this information is buried in images and video and that makes it a challenging task. One of the fundamental reasons for this difficulty occurs because mapping from the 3D scene to 2D images is generally non-invertible. Most traditional computer vision algorithms make appropriate assumptions about the nature of the 3D world and acquisition of images and videos, so that the problem of inferring scene properties of interest from sensed data becomes recoverable and analytically tractable.

Within this context, reasonably accurate, yet simple geometric models of scene structure (planar scene, etc.), scene illumination (point source), surface properties (Lambertian, Phong, etc.), imaging structure (camera models) serve critical roles in the design of inference algorithms. Moreover, images and video sequences obtained using imaging devices are invariably corrupted by noise. Common noise sources in the imaging system are due to shot noise, thermal noise, etc. Inference in this noisy environment is further complicated by the inherent errors in physical modeling. Real surfaces are never truly Lambertian, real cameras are never truly perspective, illumination in a scene is never a point light source, nevertheless inference algorithms make these assumptions in order to make the problem tractable. In addition, motion of objects in a scene could complicate the recovery of scene and object properties due to blur, occlusion, etc. Therefore, it becomes important that the developed inference algorithms can cope with varying sources of error.

To illustrate these sources of error, let us consider the following simple illustration. Suppose we are interested in designing a robot that can localize and identify the entrances to buildings (see Figure 1.1(a)). To begin, we first define a 'model' of an entrance. For computational tractability, we assume the edges of the entrance form a rectangle. Now, given an image containing the entrance, we might choose to use an edge detector or a corner detector to extract features. Due to image-noise, occlusions, and shadows, the features may not exactly correspond to edge locations. With these noisy feature locations, we proceed to fit two sets of parallel lines, where the lines from different sets are perpendicular to each other. Consider the edge figure in Figure 1.1(b). Finding the set of points corresponding to the entrance and grouping them into a rectangle comprises a combinatorial optimization problem. Suppose



(a) (b)

Fig. 1.1 Fitting a rectangle to an entrance. Various sources of error arise here – feature points are noisy, grouping of the points into a rectangle is a challenge, and a rectangle is not an accurate model for the entrance.

we obtain a solution to this optimization problem, perhaps by using the Hough transform. The final error in fit would have occurred due to noisy measurements, the difficulty in solving the constrained optimization problem, and the error in modeling itself, since the entrance does not appear as a rectangle due to perspective effects. The error would become even worse when the viewing angle moves further from frontal, or if shadows are present or the entrance is partially occluded, etc.

As this example illustrates, computer vision algorithms involve the interplay between geometric constraints that arise from models of the scene. Inference makes assumptions about the imaging devices and about appropriate statistical estimation techniques that can contend with varying sources of error. This tutorial attempts to re-examine and present several computer vision techniques accordingly.

The acceptance of statistical methods in computer vision has been slow and steady. In the early days of the field, the understanding of the geometrical aspects of the problem was given much attention. When uncertainties due to noise and other errors had to be taken into account, and when prior information and massive sensor data became available, the infusion of statistical methods was inevitable. Statistical models and methods entered into computer vision through image models. Non-causal models were first introduced in the analysis of spatial data by Whittle [222]. Subsequently, in the 1960s and 1970s, Markov random fields (MRFs) were discussed in statistical [16, 169] and signal processing literature [223]. In the 1980s, statistical methods were introduced primarily for image representation; thus MRFs [36, 47, 110] and other non-causal representations [37, 110, 111] were suggested for images. This enabled the formulation of problems such as image estimation and restoration [75], and texture analysis (synthesis, classification, and recognition) [43, 50] as maximum a posteriori estimation problems. Appropriate likelihood expressions and prior probability density functions were used to derive the required posterior probability density. Nevertheless, the maximum of the posterior probability density functions did not always yield a closed form expression, requiring techniques such as simulated annealing [75, 118].

The introduction of simulated annealing techniques could be considered a seminal moment as it opened a whole new class of sampling approaches for synthesis and segmentation of textured images [137] and other early vision problems. Simulated annealing techniques were followed by techniques such as mean field annealing [19], iterated conditional mode [18], and maximum posterior marginal [138]. These techniques are now part and parcel of computer vision algorithms. It is worth noting that MRFs and conditional random fields are making a strong resurgence in graphics and machine learning literature.

Applications of Monte Carlo Markov chain techniques for non-linear tracking problems have also been studied [79]. Since the introduction of the CONDENSATION algorithm in 1996 [95], numerous papers have discussed appearance, shape, and behavior-encoded particle filter trackers. Robust estimation methods offer another statistical area that has received attention in the computer vision literature. Many problems such as fitting lines, curves, and motion models relied on least square fitting techniques which are quite sensitive to the presence of outliers. Since the early 1980s, the use of RANSAC [67, 139], M-estimators [93], and least median square estimators [170] has become valuable in all model fitting problems, including fitting moving surfaces and objects to the optical flow generated by them. Discussions of robust estimation with applications in computer vision can be found in Meer et al. [139].

One of the recurring issues in the development of computer vision algorithms is the need to quantify the quality of the estimates. Haralick and his co-workers [108] pioneered this area. In the classical problem of estimating the 3D structure of a scene from motion cues, which is termed as the 'structure from motion' (SfM) problem, one would like to compute the lower bounds on the variances of the motion and structure estimates. Similar needs arise in camera calibration, pose estimation, image alignment, tracking, and recognition problems. A time-tested approach in statistics — the computation of Cramer–Rao bounds [166] and their generalizations — has been adopted for some computer vision problems.

The exploitation of statistical shape theory for object recognition in still images and video sequences has also been studied intensively since the 1980s. In particular, Cooper and collaborators [27, 26] have developed several algorithms based on Bayesian inference techniques for the object recognition problem. Introduction of statistical inference techniques on manifolds which host various representations used in shape, identity, and activity recognition problems is garnering a lot of interest [195].

Kanatani pioneered statistical optimization under the constraints unique to vision problems. His books explore the use of group theoretical methods [106] and statistical optimization [107] in image understanding and computer vision. In particular, Kanatani [107] explores parametric fitting under relationships such as coplanarity, collinearity, and epipolar geometry, with focus on the bounds on the estimate's accuracy. Kanatani also explored the idea of *geometric correction* of data to make them satisfy geometric constraints.

Finally, we will be grossly remiss, if we do not acknowledge Prof. Ulf Grenander, who created the area of probabilistic and statistical approaches to pattern analysis and computer vision problems. His series of books [81, 82, 84], and the recent book with Prof. Mike Miller [83], have laid the foundations for much of what has been accomplished in statistical inference approaches to computer vision. Prof. Julian Besag's contributions to the development of spatial interaction models [16, 18] and Monte Carlo Markov chain techniques [17] are seminal. Many researchers have developed statistical approaches to object detection and recognition in still images. In particular, Professors David Mumford, Don Geman, Stu Geman, Yali Amit, Alan Yuille, Mike Miller, Anuj Srivastava, Song-Chun Zhu and many others have made significant contributions to statistical approaches to still image-based vision problems. As we are focusing on video-based methods and have page constraints, we are unable to provide detailed summaries of outstanding efforts by the distinguished researchers mentioned above.

1.1 Goals

In this monograph, we will examine several interesting video-based detection, modeling, and recognition problems such as object detection and tracking, structure from motion, shape recovery, face recognition, gait-based person identification, and video-based activity recognition. We will explore the fundamental connections between these different problems in terms of the necessary geometric modeling assumptions used to solve them, and we will study statistical techniques that will enable robust solutions to these problems. Of course, a host of other image processing applications exist where statistical estimation techniques have found great use. The goal of some of these applications, such as image denoising, image deblurring, and super-resolution, is to recover an image, not 'understand' the scene captured in the image. We therefore will not delve in detail about these applications in this tutorial. An in-depth discussion of some statistical techniques applied to image processing may be found in [77, 97].

Writing this tutorial presented a great challenge. Due to page limitations, we could not include all that we wished. We simply must beg the forgiveness of many of our fellow researchers who have made significant contributions to the problems covered here and whose works could not be discussed.

1.2 Outline

We begin the monograph with an in-depth coverage of the various geometric models that are used in imaging in Section 2. Light from illumination sources interacts with materials, reflects off them, and reaches the imaging system. Therefore, it is important to study the reflectance properties of materials. We describe popular models of reflectance, such as the Lambertian and Phong models, and indicate vision applications where such reflectance models find use. Next, we describe popular models for the imaging sensor — the camera. In particular, we provide a brief description of the perspective projection model and some of its variants. Image sequences obtained from video cameras are related through scene structure, camera motion, and object motion. We also present models for both image motion (optical flow) and object/camera motion and describe how scene structure, motion, and illumination are coupled in a video.

In Section 3, we describe commonly used statistical estimation techniques such as maximum likelihood and maximum a posteriori, estimators. We also describe robust estimators such as M-estimators. We state the problem of Bayesian inference in dynamical systems and describe two algorithms — the Kalman filter and particle filters — that can perform Bayesian inference with applications to object tracking and recognition in video sequences.

In Section 4, we develop models for detection, tracking, and recognition in surveillance applications, highlighting the use of appearance and behavioral models for tracking. Section 5 describes an important fundamental problem in computer vision — structure from motion (SfM). SfM techniques study the relationship between the structure of a scene and its observability given motion. In the section, we highlight various approaches to explore this relationship, then use them to estimate both the structure of the scene and the motion. We also discuss Cramer–Rao bounds for SfM methods based on discrete features and optical flow fields.

Section 6 discusses some applications in vision where the parameters of interest lie on a manifold. In particular, we study three manifolds, the Grassmann manifold, Stiefel manifold, and the shape manifold, and show how several vision applications involve estimating parameters that live on these manifolds. We also describe algorithms to perform statistical inference on these manifolds with applications in shape, identity, and activity recognition. Finally, in Section 7, we conclude the monograph with a discussion on future trends.

References

[1] P.-A. Absil, R. Mahony, and R. Sepulchre, "Riemannian geometry of Grassmann manifolds with a view on algorithmic computation," *Acta Applicandae Mathematicae: An International Survey Journal on Applying Mathematics and Mathematical Applications*, vol. 80, no. 2, pp. 199–220, 2004.

[2] G. Adiv, "Determining three-dimensional motion and structure from optical flow generated by several moving objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, no. 4, pp. 384–401, 1985.

[3] G. Aggarwal, A. Roy-Chowdhury, and R. Chellappa, "A system identification approach for video-based face recognition," in *Proceedings of the 17th International Conference on Pattern Recognition*, vol 4, pp. 175–178, 2004.

[4] G. Aggarwal, A. Veeraraghavan, and R. Chellappa, "3D facial pose tracking in uncalibrated videos," *Lecture Notes in Computer Science*, vol. 3776, pp. 515–520, 2005.

[5] O. Arandjelovic and R. Cipolla, "Face recognition from face motion manifolds using robust kernel resistor-average distance," in *Proceedings of IEEE Workshop on Face Processing in Video*, pp. 88–94, 2004.

[6] E. M. Arkin, L. P. Chew, D. P. Huttenlocher, K. Kedem, and J. S. B. Mitchell, "An efficiently computable metric for comparing polygonal shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 3, pp. 209–216, 1991.

[7] A. Azarbayejani and A. Pentland, "Recursive estimation of motion structure and focal length," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 6, pp. 562–575, 1995.

[8] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 1–8, 2007.

[9] Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association*. San Diego, CA, USA: Academic Press Professional, Inc, 1987.

[10] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[11] E. Begelfor and M. Werman, "Affine invariance revisited," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 2087–2094, 2006.

[12] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projections," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[13] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[14] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509–522, 2002.

[15] R. Berthilsson, "A statistical theory of shape," *Lecture Notes in Computer Science*, vol. 1451, pp. 677–686, 1998.

[16] J. Besag, "Statistical analysis of non-lattice data," *The Statistician*, vol. 24, no. 3, pp. 179–195, 1975.

[17] J. Besag, "Markov chain Monte Carlo for statistical inference," Technical Report, Seattle: Center for Statistics and Social Sciences, University of Washington, 2001.

[18] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 48, no. 3, pp. 259–302, 1986.

[19] G. Bilbro, R. Mann, T. K. Miller, W. E. Snyder, D. E. Van den Bout, and M. White, "Optimization by mean field annealing," in *Proceedings of Advances in Neural Information Processing Systems*, vol. 1, pp. 91–98, 1989.

[20] C. M. Bishop, *Pattern Recognition and Machine Learning*. NJ, USA: Springer-Verlag New York, 2006.

[21] A. Bissacco, A. Chiuso, Y. Ma, and S. Soatto, "Recognition of human gaits," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 52–57, 2001.

[22] S. Biswas, G. Aggarwal, and R. Chellappa, "Efficient indexing for articulation invariant shape matching and retrieval," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2007.

[23] J. Black, T. Ellis, and P. Rosin, "Multi view image surveillance and tracking," in *Proceedings of Workshop on Motion and Video Computing*, pp. 169–174, 2002.

[24] V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1063–1074, 2003.

[25] H. Blum and R. Nagel, "Shape description using weighted symmetric axis features," *Pattern Recognition*, vol. 10, no. 3, pp. 167–180, 1978.

[26] R. M. Bolle and D. B. Cooper, "Bayesian recognition of local 3-D shape by approximating image intensity functions with quadric polynomials," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 4, pp. 418–429, 1984.

[27] R. M. Bolle and D. B. Cooper, "On optimally combining pieces of information, with application to estimating 3-D complex-object position from range data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 5, pp. 619–638, 1986.

[28] F. L. Bookstein, "Size and shape spaces for landmark data in two dimensions," *Statistical Science*, vol. 1, no. 2, pp. 181–222, 1986.

[29] T. Broida and R. Chellappa, "Performance bounds for estimating three-dimensional motion parameters from a sequence of noisy images," *Journal of Optical Society of America A*, vol. 6, no. 6, pp. 879–889, 1989.

[30] T. J. Broida, S. Chandrashekhar, and R. Chellappa, "Recursive 3-D motion estimation from a monocular image sequence," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 4, pp. 639–656, 1990.

[31] M. J. Brooks, *Shape from Shading (Ed. B. K. P. Horn).* Cambridge, MA, USA: MIT Press, 1989.

[32] E. Candes, "Compressive sampling," in *Proceedings of the International Congress of Mathematicians*, vol. 3, pp. 1433–1452, 2006.

[33] E. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007.

[34] G. Casella, R. L. Berger, and R. L. Berger, *Statistical Inference.* CA: Duxbury Pacific Grove, 2002.

[35] V. Cevher, A. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa, "Compressive sensing for background subtraction," in *Proceedings of European Conference on Computer Vision*, pp. 12–18, 2008.

[36] R. Chellappa, "Two-dimensional discrete Gaussian Markov random field models for image processing," in *Progress in Pattern Recognition 2*, (L. N. Kanal and A. Rosenfeld, eds.), pp. 79–112, New York: Elsevier, 1985.

[37] R. Chellappa and R. Kashyap, "Texture synthesis using 2-D noncausal autoregressive models," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 33, no. 1, pp. 194–203, 1985.

[38] C. C. Chen, "Improved moment invariants for shape discrimination," *Pattern Recognition*, vol. 26, no. 5, pp. 683–686, 1993.

[39] D. Chen, J. Zhang, S. Tang, and J. Wang, "Freeway traffic stream modeling based on principal curves and its analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 246–258, 2004.

[40] Y. Chikuse, *Statistics on Special Manifolds.* Springer Verlag, 2003.

[41] A. K. R. Chowdhury and R. Chellappa, "Face reconstruction from monocular video using uncertainty analysis and a generic model," *Computer Vision and Image Understanding*, vol. 91, no. 1–2, pp. 188–213, 2003.

[42] K. D. Cock and B. De Moor, "Subspace angles and distances between ARMA models," in *Proceedings of the International Symposium of Mathematical Theory of Networks and Systems*, vol. 1, 2000.

[43] F. S. Cohen and D. B. Cooper, "Simple parallel hierarchical and relaxation algorithms for segmenting noncausal Markovian random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 2, pp. 195–219, 1987.

[44] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 142–149, 2000.

[45] A. Criminisi, *Accurate Visual Metrology from Single and Multiple Uncalibrated Images*. Springer Verlag, 2001.

[46] A. Criminisi, I. Reid, and A. Zisserman, "Single view metrology," *International Journal of Computer Vision*, vol. 40, no. 2, pp. 123–148, 2000.

[47] G. R. Cross and A. K. Jain, "Markov random field texture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, no. 1, pp. 25–39, 1983.

[48] M. H. DeGroot, *Optimal Statistical Decisions*. Wiley-Interscience, 2004.

[49] Q. Delamarre and O. Faugeras, "3D articulated models and multi-view tracking with silhouettes," in *Proceedings of IEEE International Conference on Computer Vision*, vol. 2, pp. 716–721, 1999.

[50] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 39–55, 1987.

[51] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[52] D. L. Donoho and C. Grimes, "Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data," in *Proceedings of the National Academy of Sciences*, pp. 5591–5596, 2003.

[53] A. Doucet, "On sequential simulation-based methods for Bayesian filtering," Technical Report, Department of Engineering, University of Cambridge, 1998.

[54] A. Doucet, N. D. Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag, 2001.

[55] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[56] I. Dryden, "Statistical shape analysis in high-level vision," *Mathematical Methods in Computer Vision*, p. 37, 2003.

[57] I. L. Dryden and K. V. Mardia, *Statistical Shape Analysis*. Wiley New York, 1998.

[58] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM Journal on Matrix Analysis and Application*, vol. 20, no. 2, pp. 303–353, 1999.

[59] G. J. Edwards, T. F. Cootes, and C. J. Taylor, "Face recognition using active appearance models," in *Proceedings of the European Conference on Computer Vision*, vol. 2, pp. 581–695, 1998.

[60] A. Elgammal, R. Duraiswami, D. Harwood, and L. S. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1151–1163, 2002.

[61] A. Elgammal and C. Lee, "Separating style and content on a nonlinear manifold," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 1, 2004.

[62] A. Elgammal and C. S. Lee, "Inferring 3D body pose from silhouettes using activity manifold learning," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 681–682, 2004.

[63] A. Erdelyi, *Asymptotic Expansions*. Dover, 1956.

[64] O. Faugeras, Q. T. Luong, and T. Papadopoulo, *The Geometry of Multiple Images*. Massachusetts: MIT Press Cambridge, 2001.

[65] P. F. Felzenszwalb and D. P. Huttenlocher, "Pictorial structures for object recognition," *International Journal of Computer Vision*, vol. 61, no. 1, pp. 55–79, 2005.

[66] H. Fillbrandt and K. H. Kraiss, "Tracking people on the ground plane of a cluttered scene with a single camera," *WSEAS Transactions on Information Science and Applications*, vol. 2, pp. 1302–1311, 2005.

[67] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[68] F. Fleuret, J. Berclaz, and R. Lengagne, "Multi-camera people tracking with a probabilistic occupancy map," Technical Report, EPFL/CVLAB2006.07, July 2006.

[69] A. Forner-Cordero, H. Koopman, and F. van der Helm, "Describing gait as a sequence of states," *Journal of Biomechanics*, vol. 39, no. 5, pp. 948–957, 2006.

[70] G. D. Forney Jr, "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, 1973.

[71] W. Forstner, "Uncertainty and projective geometry," *Handbook of Geometric Computing: Applications in Pattern Recognition, Computer Vision, Neuralcomputing and Robotics*, pp. 493–534, 2005.

[72] H. Freeman, "On the encoding of arbitrary geometric configurations," *IRE Transactions on Electronic Computers*, vol. 10, no. 2, pp. 260–268, 1961.

[73] D. M. Gavrila and L. S. Davis, "3-D model-based tracking of humans in action: A multi-view approach," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 73–80, 1996.

[74] D. Geiger, T. Liu, and R. V. Kohn, "Representation and self-similarity of shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 1, no. 1, pp. 86–99, 2003.

[75] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.

[76] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, no. 6, pp. 643–660, 2001.

[77] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Prentice Hall, 2007.

[78] C. R. Goodall and K. V. Mardia, "Projective shape analysis," *Journal of Computational and Graphical Statistics*, vol. 8, no. 2, pp. 143–168, 1999.

[79] N. Gordon, D. Salmon, and A. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings on Radar and Signal Processing*, vol. 140, pp. 107–113, 1993.

[80] A. Goshtasby, "Description and discrimination of planar shapes using shape matrices," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, no. 6, pp. 738–743, 1985.

[81] U. Grenander, *Abstract Inference*. John Wiley & Sons, 1981.

[82] U. Grenander, *General Pattern Theory: A Mathematical Study of Regular Structures*. USA: Oxford University Press, 1993.

[83] U. Grenander and M. I. Miller, *Pattern Theory: From Representation to Inference*. USA: Oxford University Press, 2007.

[84] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*. Chelsea Publishing Company, Incorporated, 1984.

[85] G. D. Hager and P. N. Belhumeur, "Efficient region tracking with parametric models of geometry and illumination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 10, pp. 1025–1039, 1998.

[86] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of the Alvey Vision Conference*, vol. 15, pp. 147–151, 1988.

[87] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.

[88] R. I. Hartley and P. Sturm, "Triangulation," *Computer Vision and Image Understanding*, vol. 68, no. 2, pp. 146–157, 1997.

[89] R. Horaud, F. Dornaika, and B. Lamiroy, "Object pose: The link between weak perspective, paraperspective, and full perspective," *International Journal of Computer Vision*, vol. 22, no. 2, pp. 173–189, 1997.

[90] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1–3, pp. 185–203, 1981.

[91] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, 1962.

[92] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 34, no. 3, pp. 334–352, 2004.

[93] P. J. Huber, *Robust Statistics*. Wiley New York, 1981.

[94] A. Hyvrinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. John Wiley and Sons, 2001.

[95] M. Isard and A. Blake, "Condensation — conditional density propagation for visual tracking," *International Journal of Computer Vision*, vol. 29, no. 1, pp. 5–28, 1998.

[96] M. Isard and A. Blake, "ICONDENSATION: Unifying low-level and high-level tracking in a stochastic framework," in *Proceedings of European Conference on Computer Vision*, vol. 1, pp. 767–781, 1998.

[97] A. K. Jain, *Fundamentals of Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1989.

[98] T. S. Jebara and A. Pentland, "Parametrized structure from motion for 3D adaptive feedback tracking of faces," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 144–150, 1997.

[99] A. D. Jepson, D. J. Fleet, and T. F. El-Maraghi, "Robust online appearance models for visual tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1296–1311, 2003.

[100] S. Joo and Q. Zheng, "A temporal variance-based moving target detector," in *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS)*, 2005.

[101] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, "An introduction to variational methods for graphical models," *Learning in graphical models*, pp. 105–161, 1999.

[102] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–481, 2000.

[103] S. Julier and J. K. Uhlmann, "A general method for approximating nonlinear transformations of probability distributions," Technical Report, Department of Engineering Science, University of Oxford, 1996.

[104] A. Kale, A. N. Rajagopalan, A. Sundaresan, N. Cuntoor, A. Roy Cowdhury, V. Krueger, and R. Chellappa, "Identification of Humans using gait," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1163–1173, 2004.

[105] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME Journal of Basic Engineering*, vol. 82D, no. 1, pp. 34–45, 1960.

[106] K. Kanatani, *Group Theoretical Methods in Image Understanding*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1990.

[107] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*. New York, NY, USA: Elsevier Science Inc., 1996.

[108] T. Kanungo, M. Y. Jaisimha, J. Palmer, and R. M. Haralick, "A methodology for quantitative performance evaluation of detection algorithms," *IEEE Transactions on Image Processing*, vol. 4, no. 12, pp. 1667–1674, 1995.

[109] R. Kashyap and R. Chellappa, "Stochastic models for closed boundary analysis: Representation and reconstruction," *IEEE Transactions on Information Theory*, vol. 27, no. 5, pp. 627–637, 1981.

[110] R. Kashyap and R. Chellappa, "Estimation and choice of neighbors in spatial-interaction models of images," *IEEE Transactions on Information Theory*, vol. 29, no. 1, pp. 60–72, 1983.

[111] R. L. Kashyap, "Analysis and synthesis of image patterns by spatial interaction models," *Progress in Pattern Recognition*, vol. 1, pp. 149–186, 1981.

[112] Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 506–513, 2004.

[113] D. G. Kendall, “Shape manifolds, procrustean metrics, and complex projective spaces,” *Bulletin of the London Mathematical Society*, vol. 16, no. 2, pp. 81–121, 1984.

[114] S. M. Khan and M. Shah, “A multiview approach to tracking people in crowded scenes using a planar homography constraint,” in *Proceedings of European Conference on Computer Vision*, vol. 4, pp. 133–146, 2006.

[115] Z. Khan, T. Balch, and F. Dellaert, “An MCMC-based particle filter for tracking multiple interacting targets,” in *Proceedings of European Conference on Computer Vision*, vol. 4, pp. 279–290, 2004.

[116] A. Khotanzad and Y. H. Hong, “Invariant image recognition by Zernike moments,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 489–497, 1990.

[117] K. Kim and L. S. Davis, “Multi-camera tracking and segmentation of occluded people on ground plane using search-guided particle filtering,” in *Proceedings of European Conference on Computer Vision*, vol. 3, pp. 98–109, 2006.

[118] S. Kirkpatrick, “Optimization by simulated annealing: Quantitative studies,” *Journal of Statistical Physics*, vol. 34, no. 5, pp. 975–986, 1984.

[119] O. Kouropteva, O. Okun, and M. Pietikainen, “Classification of handwritten digits using supervised locally linear embedding algorithm and support vector machine,” in *Proceedings of 11th European Symposium Artificial Neural Networks*, pp. 229–234, 2003.

[120] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.

[121] M. La Cascia, S. Sclaroff, and V. Athitsos, “Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3D models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 4, pp. 322–336, 2000.

[122] J. H. Lambert, *Photometria Sive de Mensure de Gratibus Luminis, Colorum Umbrae.* Eberhard Klett, 1760.

[123] S. L. Lauritzen, *Graphical Models.* Oxford University Press, 1996.

[124] K. C. Lee, J. Ho, M. H. Yang, and D. Kriegman, “Video-based face recognition using probabilistic appearance manifolds,” in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 313–320, 2003.

[125] F. F. Li and P. Perona, “A bayesian hierarchical model for learning natural scene categories,” in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 524–531, 2005.

[126] J. Li and R. Chellappa, “Structure from planar motion,” *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3466–3477, 2006.

[127] H. Ling and D. W. Jacobs, “Deformation invariant image matching,” in *Proceedings of IEEE International Conference on Computer Vision*, vol. 2, pp. 1466–1473, 2005.

[128] H. Ling and D. W. Jacobs, “Shape classification using the inner-distance,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, pp. 286–299, 2007.

[129] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *Journal of American Statistician Association*, vol. 93, no. 443, pp. 1032–1044, 1998.

[130] J. S. Liu, R. Chen, and T. Logvinenko, "A theoretical framework for sequential importance sampling with resampling," in *Sequential Monte Carlo Methods in Practice*, (A. Doucet, N. de Freitas, and N. Gordon, eds.), New York: Springer-Verlag, 2001.

[131] Z. Liu and S. Sarkar, "Improved gait recognition by gait dynamics normalization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 6, pp. 863–876, 2006.

[132] H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, vol. 293, no. 1, pp. 133–135, 1981.

[133] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[134] F. Lv, T. Zhao, and R. Nevatia, "Camera calibration from video of a walking human," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1513–1518, 2006.

[135] T. Lv, B. Ozer, and W. Wolf, "A real-time background subtraction method with camera motion compensation," in *Proceedings of IEEE International Conference on Multimedia and Expo*, vol. 1, pp. 331–334, 2004.

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

[137] B. S. Manjunath, T. Simchony, and R. Chellappa, "Stochastic and deterministic networks for texture segmentation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 6, pp. 1039–1049, 1990.

[138] J. Marrowuin, S. Mitter, and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision," *Journal of the American Statistical Association*, vol. 82, no. 397, pp. 76–89, 1987.

[139] P. Meer, D. Mintz, A. Rosenfeld, and D. Y. Kim, "Robust regression methods for computer vision: A review," *International Journal of Computer Vision*, vol. 6, no. 1, pp. 59–70, 1991.

[140] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equations of state calculations by fast computing machines," *Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1091, 1953.

[141] W. Mio, A. Srivastava, and S. H. Joshi, "On Shape of Plane Elastic Curves," *International Journal of Computer Vision*, vol. 73, no. 3, pp. 307–324, 2007.

[142] A. Mittal and L. S. Davis, "M2 Tracker: A multi-view approach to segmenting and tracking people in a cluttered scene," *International Journal of Computer Vision*, vol. 51, no. 3, pp. 189–203, 2003.

[143] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 696–710, 1997.

[144] F. Mura and N. Franceschini, "Visual control of altitude and speed in a flying agent," in *From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior*, pp. 91–99, MIT Press, 1994.

[145] T. R. Neumann and H. H. Bulthoff, "Insect inspired visual control of translatory flight," in *Proceedings of the 6th European Conference on Advances in Artificial Life*, pp. 627–636, 2001.
[146] D. Nister, "An efficient solution to the five-point relative pose problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 756–770, 2004.
[147] B. Ochoa and S. Belongie, "Covariance propagation for guided matching," in *Proceedings of Workshop on Statistical Methods in Multi-Image and Video Processing (SMVP)*, 2006.
[148] J. Oliensis, "A critique of structure from motion algorithms," Technical Report, NEC Research Institute, 2000.
[149] S. K. Parui, S. E. Sarma, and D. D. Majumder, "How to discriminate shapes using shape vector," *Pattern Recognition Letters*, vol. 4, no. 3, pp. 201–204, 1986.
[150] V. Patrangenaru and K. V. Mardia, "Affine shape analysis and image analysis," in *22nd Leeds Annual Statistics Research Workshop*, 2003.
[151] T. Pavlidis, "A review of algorithms for shape analysis," *Document Image Analysis*, pp. 145–160, 1995.
[152] P. Peers, D. Mahajan, B. Lamond, A. Ghosh, W. Matusik, R. Ramamoorthi, and P. Debevec, "Compressive light transport sensing," *ACM Transactions on Graphics*, vol. 28, no. 1, pp. 1–3, 2009.
[153] E. Persoon and K. Fu, "Shape discrimination using Fourier descriptors," *IEEE Transactions on Man Machine and Cybernetics*, vol. 7, no. 3, pp. 170–179, 1977.
[154] J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. Bowyer, "The gait identification challenge problem: Data sets and baseline algorithm," in *Proceedings of IEEE International Conference on Pattern Recognition*, August 2002.
[155] B. T. Phong, "Illumination for computer generated pictures," *Communications of the ACM*, vol. 18, no. 6, pp. 311–317, 1975.
[156] C. J. Poelman and T. Kanade, "A paraperspective factorization method for shape and motion recovery," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 3, pp. 206–218, 1997.
[157] F. Porikli, O. Tuzel, and P. Meer, "Covariance tracking using model update based on lie algebra," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 728–735, 2006.
[158] M. J. Prentice and K. V. Mardia, "Shape changes in the plane for landmark data," *The Annals of Statistics*, vol. 23, no. 6, pp. 1960–1974, 1995.
[159] G. Qian and R. Chellappa, "Structure from motion using sequential Monte Carlo methods," *International Journal of Computer Vision*, vol. 59, no. 1, pp. 5–31, 2004.
[160] G. Qian and R. Chellappa, "Structure from motion using sequential Monte Carlo methods," *International Journal of Computer Vision*, vol. 59, no. 1, pp. 5–31, 2004.
[161] G. Qian, R. Chellappa, and Q. Zheng, "Spatial self-calibration of distributed cameras," in *Proceedings of Collaborative Technology Alliances Conference — Sensors*, 2003.

[162] L. Rabiner and B. Juang, "An Introduction to Hidden Markov Models," *IEEE ASSP Magazine*, vol. 3, no. 1, pp. 4–16, 1986.

[163] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[164] L. R. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*. Prentice Hall, 1993.

[165] A. Rangarajan, H. Chui, and F. L. Bookstein, "The softassign Procrustes matching algorithm," in *Proceedings of the International Conference on Information Processing in Medical Imaging*, pp. 29–42, 1997.

[166] C. R. Rao, *Linear Statistical Inference and its Applications*. John Wiley & Sons, 1973.

[167] H. Robbins and S. Monro, "A stochastic approximation method," *Annals of Mathematical Statistics*, vol. 22, no. 3, pp. 400–407, 1951.

[168] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York: Springer-Verlag, 1999.

[169] Y. A. Rosanov, "On Gaussian Fields with given conditional distributions," *Theory of Probability and its Applications*, vol. 12, no. 3, pp. 381–391, 1967.

[170] P. J. Rousseeuw, "Least median of squares regression," *Journal of American Statistical Association*, vol. 79, no. 388, pp. 871–880, 1984.

[171] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[172] A. K. Roy-Chowdhury and R. Chellappa, "Statistical bias in 3-D reconstruction from a monocular video," *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1057–1062, 2005.

[173] H. Rubin, "Robust Bayesian estimation," in *Proceedings of symposium on Statistical Decision Theory and Related Topics II*, pp. 351–356, New York: Academic Press, 1977.

[174] S. Saha, C. C. Shen, C. J. Hsu, G. Aggarwal, A. Veeraraghavan, A. Sussman, and S. S. Bhattacharyya, "Model-Based OpenMP implementation of a 3D facial pose tracking system," *International Workshop on Parallel Processing*, pp. 66–73, 2006.

[175] P. Saisan, G. Doretto, Y. N. Wu, and S. Soatto, "Dynamic texture recognition," in *Proceedings of IEEE International Conference on Computer Vision*, vol. 2, pp. 58–63, 2001.

[176] A. Sankaranarayanan, A. Veeraraghavan, and R. Chellappa, "Hybrid subspace sparse signals and their application to reflectance measurement," Manuscript under preparation.

[177] A. C. Sankaranarayanan and R. Chellappa, "Optimal multi-view fusion of object locations," in *Proceedings of IEEE Workshop on Motion and Video Computing (WMVC)*, pp. 1–8, 2008.

[178] A. C. Sankaranarayanan, J. Li, and R. Chellappa, "Fingerprinting vehicles for tracking across non-overlapping views," in *Army Science Conference*, 2006.

[179] A. C. Sankaranarayanan, A. Veeraraghavan, and R. Chellappa, "Object detection, tracking and recognition for multiple smart cameras," *Proceedings of the IEEE*, vol. 96, no. 10, pp. 1606–1624, 2008.

[180] S. Sarkar, P. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, "The humanID gait challenge problem: Data sets, performance, and analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162–177, 2005.

[181] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "Multiplexing for optimal lighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 8, pp. 1339–1354, 2007.

[182] J. Shao, S. K. Zhou, and R. Chellappa, "Video mensuration using stationary cameras," *IEEE Transactions on Image Processing* (under review).

[183] J. Shi and C. Tomasi, "Good features to track," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 593–600, 1994.

[184] J. Sivic, B. C. Russell, A. Efros, A. Zisserman, and W. T. Freeman, "Discovering object categories in image collections," in *Proceedings of IEEE International Conference on Computer Vision*, p. 65, 2005.

[185] L. I. Smith, *A Tutorial on Principal Components Analysis*. Cornell University, USA, 2002.

[186] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," *Autonomous Robot Vehicles*, pp. 167–193, 1990.

[187] R. C. Smith and P. Cheeseman, "On the representation and estimation of spatial uncertainty," *The International Journal of Robotics Research*, vol. 5, no. 4, pp. 56–68, 1986.

[188] S. Soatto, G. Doretto, and Y. Wu, "Dynamic textures," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 439–446, 2001.

[189] H. Sorenson, *Parameter Estimation: Principles and Problems*. M. Dekker, 1980.

[190] G. Sparr, "Depth computations from polyhedral images," *Image and Vision Computing*, vol. 10, no. 10, pp. 683–688, 1992.

[191] M. Srinivasan, S. Zhang, M. Lehrer, and T. Collett, "Honeybee navigation en route to the goal: Visual flight control and odometry," *Journal of Experimental Biology*, vol. 199, no. 1, pp. 237–244, 1996.

[192] S. Srinivasan, "Extracting structure from optical flow using the fast error search technique," *International Journal of Computer Vision*, vol. 37, no. 3, pp. 203–230, 2000.

[193] A. Srivastava, S. H. Joshi, W. Mio, and X. Liu, "Statistical shape analysis: Clustering, learning, and testing," *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 27, no. 4, pp. 590–602, 2005.

[194] A. Srivastava and X. Liu, "Tools for application-driven linear dimension reduction," *Neurocomputing*, vol. 67, pp. 136–160, 2005.

[195] A. Srivastava, W. Mio, E. Klassen, and S. Joshi, "Geometric analysis of continuous, planar shapes," in *Proceedings of 4th International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 341–356, 2003.

[196] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, 2000.

[197] J. Stuelpnagel, "On the parametrization of the three-dimensional rotation group," *SIAM Review*, vol. 6, no. 4, pp. 422–430, 1964.

[198] Z. Sun, A. M. Tekalp, and V. Ramesh, "Error characterization of the factorization method," *Computer Vision and Image Understanding*, vol. 82, no. 2, pp. 110–137, 2001.

[199] A. Sundaresan and R. Chellappa, "Model driven segmentation of articulating humans in Laplacian Eigenspace," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1771–1785, 2008.

[200] D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, "A New Compressive Imaging Camera Architecture using Optical-Domain Compression," in *Proceedings of the SPIE Computational Imaging IV*, vol. 6065, pp. 43–52, 2006.

[201] J. B. Tenenbaum, V. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[202] C. Tomasi and T. Kanade, "Detection and tracking of point features," Technical Report, CMU-CS-91–132, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1991.

[203] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," *International Journal of Computer Vision*, vol. 9, no. 2, pp. 137–154, November 1992.

[204] B. Triggs, "Factorization methods for projective structure and motion," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 845–851, 1996.

[205] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment — A modern synthesis," in *Proceedings of the International Workshop on Vision Algorithms*, pp. 298–372, 2000.

[206] P. Turaga and R. Chellappa, "Locally time-invariant models of human activities using trajectories on the Grassmannian," in *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 2435–2441, 2009.

[207] P. Turaga, A. Veeraraghavan, and R. Chellappa, "Statistical analysis on Stiefel and Grassmann manifolds with applications in computer vision," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2008.

[208] P. K. Turaga, A. Veeraraghavan, and R. Chellappa, "From videos to verbs: Mining videos for events using a cascade of dynamical systems," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, 2007.

[209] P. K. Turaga, A. Veeraraghavan, and R. Chellappa, "Unsupervised view and rate invariant clustering of video sequences," *Computer Vision and Image Understanding*, vol. 113, no. 3, pp. 353–371, 2009.

[210] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[211] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 586–591, 1991.

[212] O. Tuzel, F. Porikli, and P. Meer, "Region covariance: A fast descriptor for detection and classification," in *European Conference on Computer Vision*, pp. 589–600, 2006.

[213] O. Tuzel, F. Porikli, and P. Meer, "Pedestrian detection via classification on Riemannian manifolds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1713–1727, 2008.

[214] A. Veeraraghavan, R. Chellappa, and A. K. Roy-Chowdhury, "The function space of an activity," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 959–968, 2006.

[215] A. Veeraraghavan, R. Chellappa, and M. Srinivasan, "Shape-and-behavior encoded tracking of bee dances," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 3, pp. 463–476, 2008.

[216] A. Veeraraghavan, D. Reddy, and R. Raskar, "Coded strobing camera for high speed periodic events," *IEEE Transactions on Pattern Analysis and Machine Intelligence* (under revision), 2009.

[217] A. Veeraraghavan, A. K. Roy-Chowdhury, and R. Chellappa, "Matching shape sequences in video with applications in human movement analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 12, pp. 1896–1909, 2005.

[218] A. Veeraraghavan, A. Srivastava, A. K. Roy-Chowdhury, and R. Chellappa, "Rate-invariant recognition of humans and their activities," *IEEE Transactions on Image Processing*, vol. 18, no. 6, pp. 1326–1339, 2009.

[219] K. von Frisch, *The Dance Language and Orientation of Bees*. Cambridge, Massachusetts: Belknap Press, 1967.

[220] Y. Wang, H. Jiang, M. S. Drew, Z. N. Li, and G. Mori, "Unsupervised discovery of action classes," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 1654–1661, 2006.

[221] D. Weinland, R. Ronfard, and E. Boyer, "Free viewpoint action recognition using motion history volumes," *Computer Vision and Image Understanding*, vol. 104, no. 2, pp. 249–257, 2006.

[222] P. Whittle, "On stationary processes in the plane," *Biometrika*, vol. 41, no. 3–4, pp. 434–449, 1954.

[223] J. Woods, "Two-dimensional discrete Markovian fields," *IEEE Transactions on Information Theory*, vol. 18, no. 2, pp. 232–240, 1972.

[224] C. R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Realtime tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780–785, 1997.

[225] M. Xu, J. Orwell, and G. Jones, "Tracking football players with multiple cameras," in *Proceedings of IEEE International Conference on Image Processing*, pp. 2909–2912, 2004.

[226] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Computing Surveys (CSUR)*, vol. 38, no. 4, pp. 1–45, 2006.

[227] G.-S. J. Young and R. Chellappa, "Statistical analysis of inherent ambiguities in recovering 3-D motion from a noisy flow field," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 995–1013, 1992.

[228] J. Zhang, S. Z. Li, and J. Wang, "Nearest manifold approach for face recognition," in *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 17–19, 2004.

[229] Z. Zhang, "Parameter estimation techniques: A tutorial with application to conic fitting," *Image and Vision Computing*, vol. 15, no. 1, pp. 59–76, 1997.

[230] Z. Zhang and H. Zha, "Principal manifolds and nonlinear dimensionality reduction via tangent space alignment," *SIAM Journal on Scientific Computing*, vol. 26, no. 1, pp. 313–338, 2005.

[231] T. Zhao and R. Nevatia, "Tracking multiple humans in complex situations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1208–1221, 2004.

[232] T. Zhao and R. Nevatia, "Tracking multiple humans in crowded environment," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 406–413, 2004.

[233] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.

[234] S. Zhou, V. Krueger, and R. Chellappa, "Probabilistic recognition of human faces from video," *Computer Vision and Image Understanding*, vol. 91, no. 1–2, pp. 214–245, 2003.

[235] S. K. Zhou, R. Chellappa, and B. Moghaddam, "Visual tracking and recognition using appearance-adaptive models in particle filters," *IEEE Transactions on Image Processing*, vol. 13, no. 11, pp. 1491–1506, 2004.