ATR applications of minimax entropy models of texture and shape

Song Chun Zhu[a], Alan L. Yuille[b], and Aaron D. Lanterman[c]

[a]Dept. of Computer and Information Sciences, The Ohio State Univ., 2015 Neil Avenue, Columbus, OH 43210

[b]Smith-Kettlewell Eye Research Institute, 2318 Fillmore St., San Francisco, CA 94115

[c]Coordinated Science Laboratory, Univ. of Illinois, 1308 W. Main, Urbana, IL 61801

ABSTRACT

Concepts from information theory have recently found favor in both the mainstream computer vision community and the military automatic target recognition community. In the computer vision literature, the principles of minimax entropy learning theory have been used to generate rich probabilitistic models of texture and shape. In addition, the method of types and large deviation theory has permitted the difficulty of various texture and shape recognition tasks to be characterized by "order parameters" that determine how fundamentally vexing a task is, independent of the particular algorithm used. These information-theoretic techniques have been demonstrated using traditional visual imagery in applications such as simulating cheetah skin textures and such as finding roads in aerial imagery. We discuss their application to problems in the specific application domain of automatic target recognition using infrared imagery. We also review recent theoretical and algorithmic developments which permit learning minimax entropy texture models for infrared textures in reasonable timeframes.

Keywords: information theory, texture modeling, shape modeling, automatic target recognition, performance analysis

1. INTRODUCTION

Automatic target recognition (ATR) systems attempt to detect, locate, and identify targets of interest from collected data.[1] Some systems, such as those which employ radar or sonar, may exploit information which is not readily interpretable by the human visual system. Other systems, such as those employing visual cameras, forward-looking infrared (FLIR) focal plane arrays, or range-imaging laser radars, produce two-dimensional imagery readily understandable by a human observer. Hence, many researchers from the broader field of computer vision have delved into world of ATR.

Although the potential applications of ATR systems are quite broad, the term "automatic target recognition" seems to be most often employed in military contexts. In many machine vision applications, such as automated industrial inspection and medical image analysis, the conditions under which data are collection may be carefully controlled. The object under study can be carefully positioned and clutter kept out of the scene. In other applications, such as autonomous vehicular navagation, less control is available; scenes are often highly cluttered. Military ATR systems may represent the most challenging extreme of this spectrum. Such systems must be able to operate in hostile and highly cluttered environments. The targets sought may employ concealment and camoflague, and may attempt to deliberately "spoof" the ATR system with decoy targets. In the following, we will use the term ATR in reference to military applications.

Another difference between typical computer vision systems and imaging ATR systems is the kind of camera employed. Most computer vision work has been done using standard visual imagery; most ATR systems employ forward-looking infrared* (FLIR) sensors. Infrared is of particular value due to its ability to be used at night and to detect the heat generated by vehicle engines. While the geometric effects of perspective projection and obscuration are the same for visual and infrared sensors, the variability in intensity is quite different. In visual imagery, the prime

E-mail: szhu@cis.ohio-state.edu, yuille@ski.org, lanterma@ifp.uiuc.edu. Web sites: http://www.cis.ohio-state.edu/~szhu, http://www.ski.org/ALYuille_lab, www.ifp.uiuc.edu/~lanterma

*The modifier "forward-looking" is something of a historical accident; it was added to differentiate "staring" systems, such as focal plane arrays, from infrared search-and-track (IRST) systems which scan large areas and detect point targets. A FLIR camera does not necessarily have to be facing "forward."

source of intensity variation is the sources of illumination. In the infrared band, where emissive effects dominate reflectance effects, the varying thermal state of the targets is the prime source of intensity variation. Just as a visual recognition system must account for variability in illumination,[2] an infrared recognition system must be able to handle thermodynamic variability.[3] In infrared, environmental conditions such as ambient temperature also have an important effect.

Designers of ATR systems rarely try to craft generic image understanding algorithms, as is sometimes the goal of researchers in the artificial intelligence and mainstream computer vision community. Instead, the ATR focus is usually on recognizing a specific well-defined and well-characterized set of targets. The targets are usually rigid, having at most a few degrees of freedom (such as a rotating tank turret, or a gun barrel which may be raised or lowered). This contrasts with systems which must recognize highly flexible objects like people and animals.[4]

Throughout the 90s, concepts from information theory originally formulated in communication system contexts became increasingly prevalent in discussions on image processing and imaging understanding problems; see Ref. 5 for an inspiring review of these developments, as well as the special issue on information-theoretic imaging of the IEEE Transactions on Information Theory (August 2000). In particular, Shusterman and colleagues[6] employ rate-distortion theory towards designing target libraries; Cooper and Miller[7,8] consider information measures for characterizing the performance of ATR systems. The results in Ref. 7 draw upon previous research on estimation techniques and error analysis appropriate for rotation groups.[9,10] O'Sullivan considers some other roles of information theory in imaging and ATR in another paper in this proceedings.[11]

This review article focuses on two particular manifestations of information theory which have risen from the computer vision community: 1) *minimax entropy learning theory* for crafting generic models of texture and shape, and 2) *order statistics* for characterizing the fundamental difficulty of certain image understanding problems. To date, experiments with these techniques have focused entirely on visual imagery. Here, we consider their potential application to automatic target recognition systems using infrared data. Our goal is to introduce ATR practitioners to these theoretical developments.

2. HIGH-LEVEL INFERENCE ALGORITHMS

Image understanding algorithms generally fall into two broad categories: bottom-up and top-down. Bottom-up algorithms often start with an edge detection algorithm. The edges are then fused together to form lines and regions; these regions are aggregated together and then analyzed to infer scene components. Top-down vision algorithms, on the other hand, tend to avoid pre-processing the data, and try to match object templates directly against the measured data. The deliniation between top-down and bottom-up is not always clear, and there are a spectrum of possibilities between these two extremes.

A variety of image understanding paradigms may be framed as statistical inference problems. One can craft a likelihood function which tells how likely it is that, say, a particular set of objects at different positions (in a top-down algorithm) or a particular configuration of edges (in a bottom-up algorithm) gave rise to a certain image. This likelihood can be fused with a prior distribution to form a Bayesian posterior. In a top-down algorithm, the prior might prefer certain arrangements of objects; for instance, a car is more likely to be on a road than on a cornfield. A bottom-up algorithm might include a prior which encourages edges to connect together to form line segments and which discourages isolated edges.

The resulting statistical inference problems rarely have simple solutions. This has lead many researchers to random sampling techniques, particularly Markov chain Monte Carlo (MCMC) algorithms, for exploring a likelihood or Bayesian posterior surface. For example, Ref. 12 details a jump-diffusion algorithm for recognizing vehicles in infrared imagery. The jumps account for the discrete aspects of the problem, i.e. determining the number of vehicles and their types (for instance, M2, M60, T62, etc.) The diffusions address the continuous aspects of the inference by refining esimates of target position and orientation on the ground plane.

Recently, proposals have been made to integrate the bottom-up and top-down views of object recognition in a coherent fashion, and derive MCMC algorithms which choose their moves based on the data to achieve computationally efficient inference.[13,14]

3. TEXTURE MODELS

3.1. Models for Texture

In minimax entropy texture modeling,[15,16] we apply a set of K linear filters $\{F^{(1)}, \ldots, F^{(K)}\}$ to the image I to form a set of filtered images $\{I^{(1)}, \ldots, I^{(K)}\}$, where $I^{(k)} = F^{(k)} \star I$. For each filtered image $I^{(k)}$, we compute a set of statistics $\{h^{(k,1)}, \ldots, h^{(k,L)}\}$. (Although not explicitly notated, the number of statistics computed, L, may depend on the particular filter k). A variety of statistics have been explored in experiments using this framework, such as conventional moments (means, variances, etc.), rectified functions, and histograms. In studies on *visual* imagery, histograms appeared to offer the most "bang for the buck." Hence, some later works[17,18] seem to focus exclusively on histograms. However, since *infrared* imagery has a different "flavor" than everyday visual imagery, we would like to reserve judgement on how much the extensive experience in the visual domain will transfer to the infrared realm.

Given a set of empirically derived statistics, we want to construct a probability model $p(I)$ which, on average, yields statistics which are the same as the empirical statistics we measured, i.e., $E_p[h(I)] = h_{emp}$. Many different models may fit this criteria; we seek the model which maximizes the entropy $\int p(I) \ln p(I) dI$ subject to the empirically derived constraints, as this makes as few additional assumptions on the distributions as possible. This maximum entropy distribution has the Gibbs form

$$P(I; \lambda) = \frac{1}{Z(\beta)} \exp \left[-\sum_{k=1}^{K} \sum_{\ell=1}^{L(k)} \lambda^{(k,\ell)} h^{(k,\ell)}(I) \right], \tag{1}$$

where $Z(\beta)$ is the normalizing *partition function* and the λ's are Lagrange multipliers which must be estimated. Unfortunately, finding the λ's which satisfy the $E_p[h(I)] = h_{emp}$ is actually quite challenging.

One approach is to gradually improve the λ's by iterating

$$\lambda^{(k)}(\ell) \leftarrow E_{p(I:\lambda)}[h^{(k,\ell)}(I)] - h_{emp}^{(k,\ell)} \tag{2}$$

At each iteration, we approximate the expectation by drawing a random sample image I_{samp} from $p(I : \lambda)$ using a Markov chain Monte Carlo algorithm and taking $E_{p(I:\lambda)}[h^{(k,\ell)}(I)] \approx h^{(k,\ell)}(I_{samp})$. Since the MCMC algorithm itself is iterative, the algorithm for estimating the λ's has an iteration nested inside an iteration. Since λ will not change drastically from iteration to iteration of (2), we can fortunately use the I_{samp} from the previous iteration to initialize the MCMC algorithm for the next iteration. Nontheless, this is a computationally intense procedure.

A question remains as to how the filters should be chosen. We can adopt a greedy scheme, where we add filters one at a time. At each filter addition, we add the filter whose maximum entropy distribution yields the minimum entropy. (That takes a bit of thinking to sort through; we maximize over the parameters given the filter, but minimize over the filter choices.) Note that this wraps yet another iteration around the two loops discussed in the previous paragraph. One must avoid adding too many filters and thus "overfitting" the data; for this, a model order selection such as the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) may be employed.

The resulting model is called FRAME[16] (Filters, RAndom fields, and Maximum Entropy), since it unifies ideas from filtering theory and Markov random field theory. Some examples of textures simulated from FRAME models are shown in Fig. 1. The models can also be placed in the framework of Gibbs Reaction-Diffusion Equations[19] (GRADE), which offers links to work on anisotropic diffusion.[20] As shown in Fig. 2, the models are particularly powerful for removing clutter from scenes, where separate FRAME models have been learned for the targets and the clutter. The clutter removal problem is framed as one of *maximum a posteriori* (MAP) estimation; the MAP estimate is computed via a Langevin diffusion with annealing.[21]

Note that the minimax criteria described above seeks texture models which are true in appearance to the real textures. If the goal is classification of textures from a particular set of texture classes, it makes sense to derive a different procedure with this goal in mind. Konishi and Yuille[22] use probability distributions of filter responses learned from real imagery in general image segmentation. Instead of the minimax criteria, they use a criteria based on Chernoff information.

3.1.1. Julesz Ensembles to the Rescue!

Wu, Zhu, and Liu define a Julesz ensemble[18] to be a uniform probability distribution on images which is constant over those images whose statistics match (within some tolerance) those of a desired target texture, and zero for those which do not. For large texture images, there is a fortuitious equivalence between the Julesz ensemble and the maximum entropy model described in the previous section. This means that we can simulate textures from a particular Gibbs distribution by just matching texture statistics, and without having to learn the computationally expensive FRAME model.

3.1.2. Further Speedup

We have seen how the equivalence between Julesz ensembles and Gibbs models can speed up model selection. To speed up maximum likelihood estimation of a general Gibbs model, several techniques have been recently proposed,[23] going by names like "maximum partial likelihood estimator," "maximum patch likelihood estimator," and "maximum satellite likelihood estimator." In particular, the maximum satellite likelihood estimator exploits a set of reference models whose parameters are computed off-line. Ref. 23 considers the accuracy of estimating the Lagrange parameters in terms of two sources of error: 1) the fundamental variance of the maximum-likelihood estimator, characterized by the inverse fisher information, and 2) the error associated with approximating the partition function (normalizer of the maximum entropy density) using Monte Carlo integration techniques. The various new proposed techniques allow various tradeoffs between accuracy and speed.

3.2. Related Work

Grenander and Srivastava[24] propose a "transported generator" clutter model which supposes that a cluttered scene consists of a large number of randomly placed identical objects. This simplified model permits a closed-form (although complicated) expression for joint pixel statistics. In particular, they derive a density function for neighboring pixel differences which involves a modified Bessel function. This density has tails which are heavier than those of the Gaussian. Laplacian (double-sided decaying exponential) clutter is a special case of this more general density.

4. SHAPE MODELS

Intriguingly, the same kind of minimax entropy learning theory described in Sec. 3.1 for learning textures models may be used to learn shape models.[25,26] Instead of an image, the subject of study is a parametric description of a two-dimensional shape. Of course, images in the real world are three-dimensional; however, unless multiple cameras are available, we only see them through their projections, so it is reasonable to characterize them in the plane. Such models are useful in image segmentation[27] or object recognition.[4]

In minimax entropy shape learning, filters which operate on contours, instead of the image plane, take the place of the Gabor filters (or other image processing operation) used in the texture study. For instance, first and second derivatives can be used to characterize curvature and circularity. Such local characteristics have been used for many years in deformable contour analysis. However, in traditional formulations, such as snakes,[28,29] a functional form for the energy associated with curvature, etc. is assumed *a priori*. In the minimax entropy framework, we can instead learn the appropriate probability distributions directly from real-world examples.

Simple derivatives of the shape only take into account local characteristics. They do not account for broad, region-based characteristics. The minimax entropy framework permits incorporation of such region-based characteristics, mathematically characterizing ideas from gestalt psychology.[30] In particular, we may exploit symmetry properties of objects and characterize the average thinkness of object components like "limbs." Again, all of these characteristics may be learned from real-world examples. MCMC random sampling algorithms may be used to generate samples from the learned distribution to see how realistic the resulting shapes are. To date, the shapes generated from such models do not seem as realistic as the textures generated by the models described in the previous section. For instance, the bottom row of Fig. 3 shows shapes simulated from a model learned from animal shapes like those shown in the top row. However, the models do seem a sufficient characterization to permit effective inference.

Incorporating these higher-level gestalt concepts requires a mechanism for finding the *medial axes* of shapes. The medial axes essentially give the "skeletal" structure of a shape. Although there are several techniques for finding medial axes, a particularly appealing one is a jump-diffusion algorithm,[26] which is similar in spirit to the infrared target recognition algorithm[12] described Sec. 2.

4.1. The Role of Shape Models in ATR

In many ATR applications, the main objects of interest are rigid targets whose three-dimensional geometry is well known. For these objects, it makes sense to manipulate their full three-dimensional CAD representations.[12] The main challenge to detecting, locating, and recognizing these objects of interest is the highly cluttered scenes in which they may be embedded. Some experiments[31] have shown that *structured* clutter tends to be the most troublesome to ATR systems.

Although general texture models may be helpful for modeling some kinds of clutter, an ATR algorithm may find improved performance by directly considering the shape of the elements cluttering the scene. This is where shape models, such as those described in the last section, may be helpful. Of course, the definition of *target* and *clutter* will depend on the scenario; what may be the target in one scenario may be clutter in another, and vice versa! A particularly interesting piece of clutter, although not one of the prime objects of interest (say, tanks), may be of tactical interest to the operator. Indeed, many ATR systems are "automated" instead of "automatic," in that a human operator is kept "in the loop." Marking objects as "interesting," even if it isn't one of the rigid targets in the system's library, allows the human operator to rapidly focus attention on areas of interest without having to constantly scan a large image area.

Structured clutter may be either natural or man-made. For natural objects, such as trees and boulders, the models described in the last section will be appropriate. For man-made objects, we will have to train a different model, perhaps using different extracted features. In particular, natural objects tend to have smooth boundaries, whereas man-made objects such as buildings and roads will have largely straight portions with sharp corners.

5. PERFORMANCE ANALYSIS VIA ORDER PARAMETERS

Contemplate the three segmentation tasks suggested by Fig. 4. In the left panel, the stop sign is readily and quickly seen; indeed, such signs are deliberately created so that they stand out from the background. The lizard in the middle panel, on the other hand, has the benefit of natural camoflage and is trying to blend into the background, and is hence more difficult to see. The ultimate extreme is the rightmost spotted image, which is a longtime favorite of psychophysicists; one could look at it a long time before seeing the dalmation.

Yuille and Coughlan[32] describe how the difficulty of many Bayesian image analysis tasks, such as the segmentation tasks outlined in the previous paragraph, can be characterized by a single *order parameter* K. In particular, a *phase transition* can be observed; if K is below a particular threshold, then the specified task will be impossible for *any* proposed algorithm. Problems are set up by decomposing them into a series of decision problems. In statistical decision theory, two hypotheses are characterized by probability mass functions (or densities) P_A and P_B. A useful characterization of the discrepancy between two distributions, and hence the ease or difficulty of choosing between them given sample data, is the *relative entropy* or *Kullback-Leibler distance*:

$$D(P_A \| P_B) = \sum_x P_A(x) \ln \frac{P_A(x)}{P_B(x)} \tag{3}$$

Quantities of the form, as well as a related discrepancy measure called the *Chernoff information*, are of vital importance in the order parameter framework. In its simplest form, the framework is an application of the theory of *types*, which essentially considers the behavior of empirical histograms via simple counting and bounding arguments. More complicated situations demand an invokation of the much more complex theory of *large deviations*.[†] Jain *et al.*[34] present some work which is similar in spirit. Kullback-Leibler distances are also used to characterize clutter in Ref. 31.

In addition to assessing overall estimator performance from an accuracy standpoint, other kinds of order parameters may be computed to predict the computational and memory complexity of the associated estimation algorithms.[35,36]

The theory has been used to assess the effectiveness of various edge detectors,[37] determine the degree to which high-level knowledge helps in various image understanding tasks,[38] quantify the complexity of road tracking problems,[32] and characterize the difficulty of discriminating textures and shapes[39,40] learned using the minimax entropy theory reviewed in Secs. 3.1 and 4. Order parameters for the road tracking problem are illustrated in Fig. 5

[†]Most tomes on large deviation theory obsess over rather enigmatic mathematical details; this seems to have had the unforuntate effect of scaring many practitioners away. One exception is the refreshingly readable, engineering-oriented account by Bucklew.[33]

6. CONCLUSIONS

We have reviewed minimax entropy models of visual textures and biological shapes. The application of these techniques to infrared images and the characterization of clutter in ATR systems is a wide open area inviting exploration. The use of information theory to characterize the difficulty of various ATR tasks via order parameters may also be fruitful.

ACKNOWLEDGMENTS

Over the course of the much of the research reviewed here, all of the authors were supported by the U.S. Army Center for Imaging Science under Army Research Office grant DAAH049510494. It was through this multi-university initiative that A.D. Lanterman was first introduced to and began interacting with A.L. Yuille and S.C. Zhu. A.L. Yuille would like to acknowledge funding from the U.S. National Science Foundation, award number IRI-9700446, the Smith-Kettlewell core grant, and the U.S. Air Force Office of Scientific Research, grant F49620-98-1-0197. S.C. Zhu would like to acknowledge funding from ARO grant DAAD19-99-1-0012 and NSF grant IIS-98-77-127. A.D. Lanterman is currently supported by a grant from the U.S. Defense Advanced Research Projects Agency under Contract F49620-98-1-0498, administered by AFOSR.

REFERENCES

1. J. Ratches, C. Walters, R. Buser, and B. Guenther, "Aided and automatic target recognition based upon sensory inputs from image forming systems," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **19**, pp. 1004–1019, September 1997.
2. P. Hallinan, G. Gordon, A. Yuille, P. Giblin, and D. Mumford, *Two- and Three-Dimensional Patterns of the Face*, A.K. Peters, 1999.
3. A. Lanterman, "Bayesian inference of thermodynamic state incorporating Schwarz-Rissanen complexity for infrared target recognition," *Optical Engineering* **39**, pp. 1282–1292, May 2000.
4. S. Zhu and A. Yuille, "FORMS: A flexible object recognition and modeling system," *International Journal of Computer Vision* **20**(3), pp. 187–212, 1996.
5. J. O'Sullivan, R. Blahut, and D. Snyder, "Information-theoretic image formation," *IEEE Trans. on Information Theory* **44**, pp. 2094–2123, October 1998.
6. E. Shusterman, M. Miller, and B. Rimoldi, "Rate-distortion theory applied to automatic object recognition," *IEEE Trans. on Information Theory* **46**(5), pp. 1921–1927, 2000.
7. M. Cooper and M. Miller, "Information measures for object recognition accommodating signature variablility," *IEEE Trans. on Information Theory* **46**(5), pp. 1896–1907, 2000.
8. M. Cooper and M. Miller, "Information measures for object recognition," in *Algorithms for Synthetic Aperture Radar Imagery V*, E. Zelnio, ed., vol. SPIE Proc. 3370, pp. 637–645, (Orlando, FL), April 1998.
9. U. Grenander, M. Miller, and A. Srivastava, "Hilbert-Schmidt lower bounds for estimators on matrix Lie groups," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **20**, pp. 790–801, August 1998.
10. U. Grenander, A. Srivastava, and M. Miller, "Asymptotic performance analysis of Bayesian target recognition," *IEEE Trans. on Information Theory* **46**, pp. 1658–1665, July 2000.
11. J. O'Sullivan, "Some roles of information theory in imaging problems," in *Automatic Target Recognition XI*, F. Sadjadi, ed., vol. SPIE Proc., this volume, (Orlando, FL), April 2001.
12. A. Lanterman, M. Miller, and D. Snyder, "General Metropolis-Hastings jump diffusions for automatic target recognition in infrared scenes," *Optical Engineering* **36**, pp. 1123–1137, April 1997.
13. S. Zhu, R. Zhang, and Z. Tu, "Integrating bottom-up/top-down for object recognition by data driven markov chain monte carlo," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 738–745, 2000.
14. S. Zhu, Z. Tu, and R. Zhang, "Effective Bayesian inference by data-driven Markov chain Monte Carlo," in *Automatic Target Recognition X*, F. Sadjadi, ed., vol. SPIE Proc. 4050, pp. 320–332, (Orlando, FL), April 2000.
15. S. Zhu, Y. N. Wu, and D. Mumford, "Minimax entropy principle and its applications to texture modeling," *Neural Computation* **9**, pp. 1627–1660, November 1997.
16. S. Zhu, Y. Wu, and D. Mumford, "FRAME: Filters, random fields, and maximum entropy - towards a unified theory for texture modeling," *International Journal of Computer Vision* **27**, pp. 1–20, March/April 1998.

17. S. Zhu, X. Liu, and Y. Wu, "Exploring texture ensembles by efficient Markov chain Monte Carlo - toward a "trichromacy" theory of texture," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **22**, pp. 554–569, June 2000.
18. Y. Wu, S. Zhu, and X. Liu, "Equivalence of Julesz ensembles and FRAME models," *International Journal of Computer Vision* **38**(3), pp. 247–265, 2000.
19. S. Zhu and D. Mumford, "Prior learning and Gibbs reaction-diffusion," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **19**, pp. 1236–1250, November 1997.
20. P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **12**, pp. 629–639, July 1990.
21. S. Geman and C.-R. Hwang, "Diffusions for global optimization," *SIAM J. Control and Optimization* **24**, pp. 1031–1043, 1987.
22. S. Konishi and A. Yuille, "Statistical cues for domain specific image segmentation with performance analysis," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 125–132, 2000.
23. S. Zhu, "Learning in gibbsian fields: How accurate and how fast can it be?," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 2–9, 2000.
24. U. Grenander and A. Srivastava, "Probability models for clutter in natural images," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **23**, April 2001.
25. S. Zhu, "Embedding gestalt laws in Markov random fields," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **21**, pp. 1170–1187, Nov. 1999.
26. S. Zhu, "Stochastic jump-diffusion process for computing medial axes in Markov random fields," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **21**, pp. 1158–1169, Nov. 1999.
27. S. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **18**, pp. 884–900, Sept. 1996.
28. M. Kass, A. Witkin, and D. Terzopolous, "Snakes: Active contour models," *International Journal of Computer Vision* **1**(4), pp. 321–331, 1988.
29. L. Staib and J. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **PAMI-14**(11), pp. 1061–1075, November 1992.
30. K. Koffka, *Principles of Gestalt Psychology*, Harcourt, Brace, and Company, 1935.
31. A. D. Lanterman, J. A. O'Sullivan, and M. I. Miller, "Kullback-Liebler distances for quantifying clutter and models," *Optical Engineering* **38**, pp. 2134–2146, December 1999.
32. A. Yuille and J. Coughlan, "Fundamental limits of bayesian inference: Order parameters and phase transitions for road tracking," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **22**, pp. 160–173, February 2000.
33. J. A. Bucklew, *Large Deviation Techniques in Decision, Simulation, and Estimation*, John Wiley and Sons, 1990.
34. A. Jain, P. Moulin, M. Miller, and K. Ramchandran, "Information-theoretic bounds on target recognition performance," in *Automatic Target Recognition X*, F. Sadjadi, ed., vol. SPIE Proc. 4050, pp. 347–358, (Orlando, FL), April 2000.
35. J. Coughlan and A. Yuille, "Bayesian a^* tree search with expected $o(n)$ convergence rates for road tracking," in *Proc. EMMCVPR*, 1999.
36. A. Yuille and J. Coughlan, "Convergence rates of algorithms for visual search: Detecting visual contours," in *Proc. NIPS*, 1998.
37. S. Konishi, A. Yuille, J. Coughlan, and S. Zhu, "Fundamental bounds on edge detection: An information theoretic evaluation of different edge cues," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 573–379, 1999.
38. A. Yuille and J. Coughlan, "High-level and generic models for visual search: When does high level knowledge help?," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 631–637, 1999.
39. A. Yuille, J. Coughlan, S. Zhu, and Y. Wu, "Order parameters for minimax entropy distributions: When does high level knowledge help?," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 558–565, 2000.
40. A. Yuille, "A unified framework for performance analysis of bayesian inference," in *Automatic Target Recognition X*, F. Sadjadi, ed., vol. SPIE Proc. 4050, pp. 333–346, (Orlando, FL), April 2000.

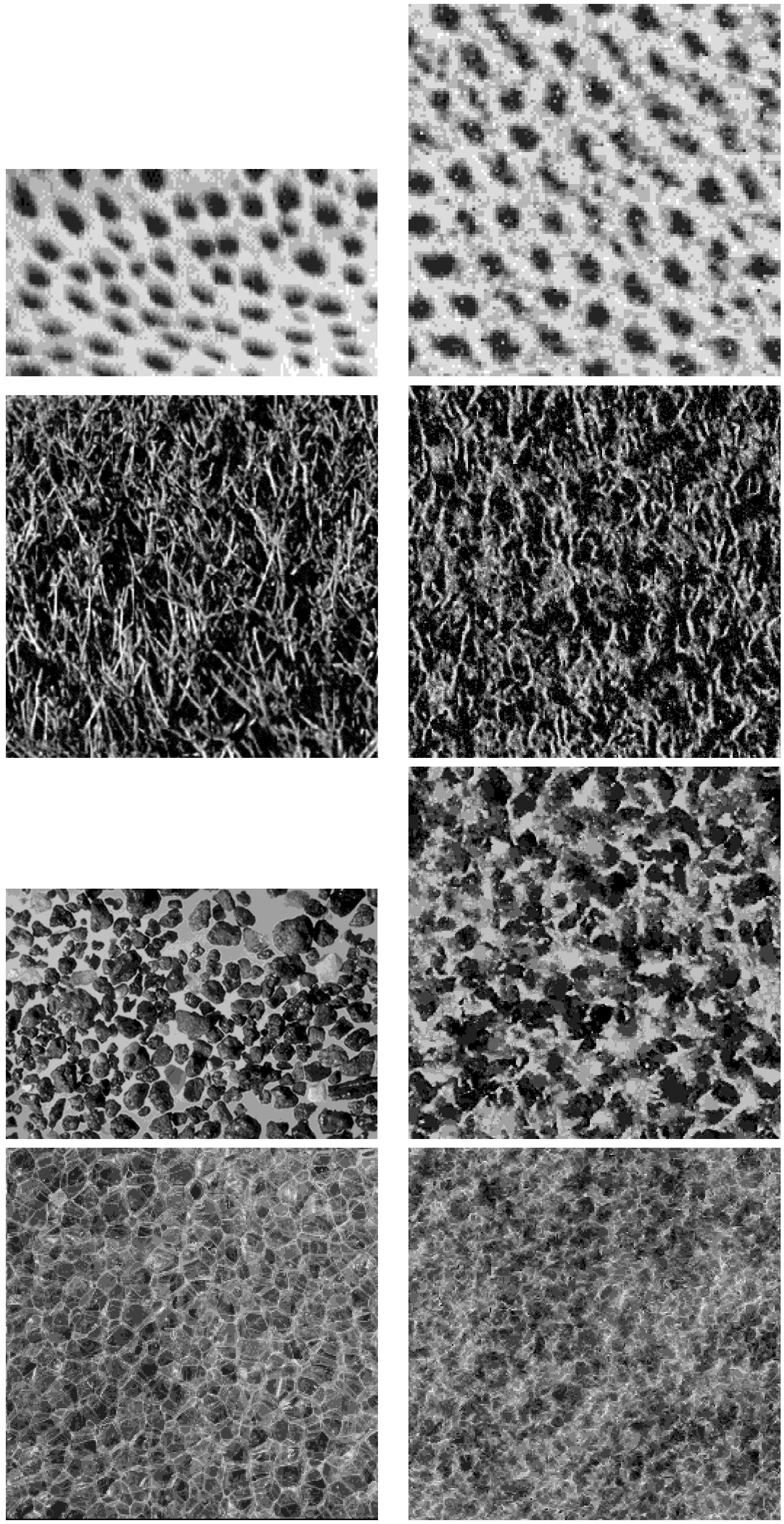

Figure 1. Left column: original texture images. Right column: texture images simulated from the learned FRAME models. Visit Y.N. Wu's webpage at http://www.stat.ucla.edu/~ywu for many more examples.

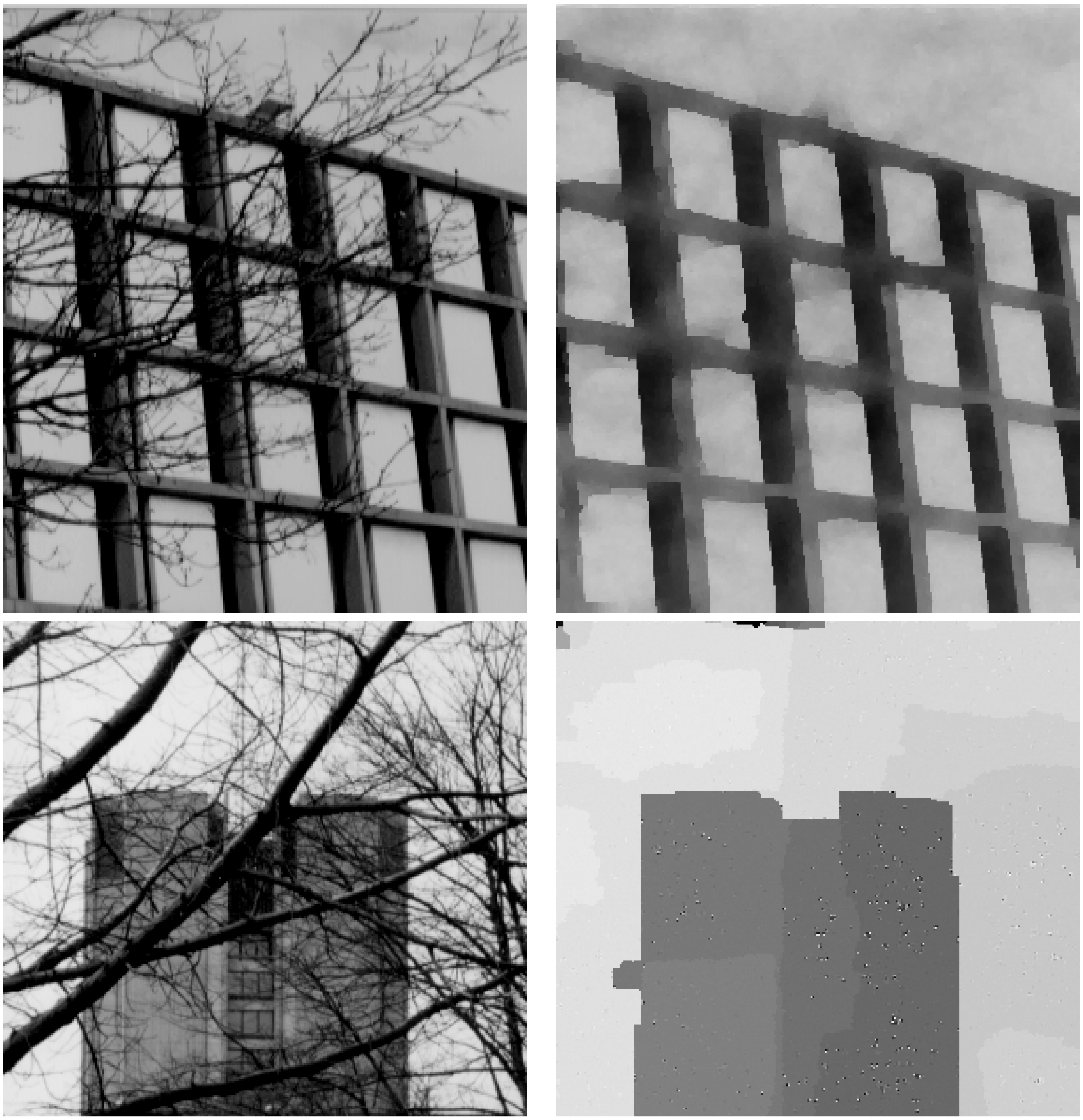

Figure 2. Examples of clutter removal using learned FRAME models for target (buildings) and clutter (trees). Left column shows the original image; right column shows result of clutter removal.

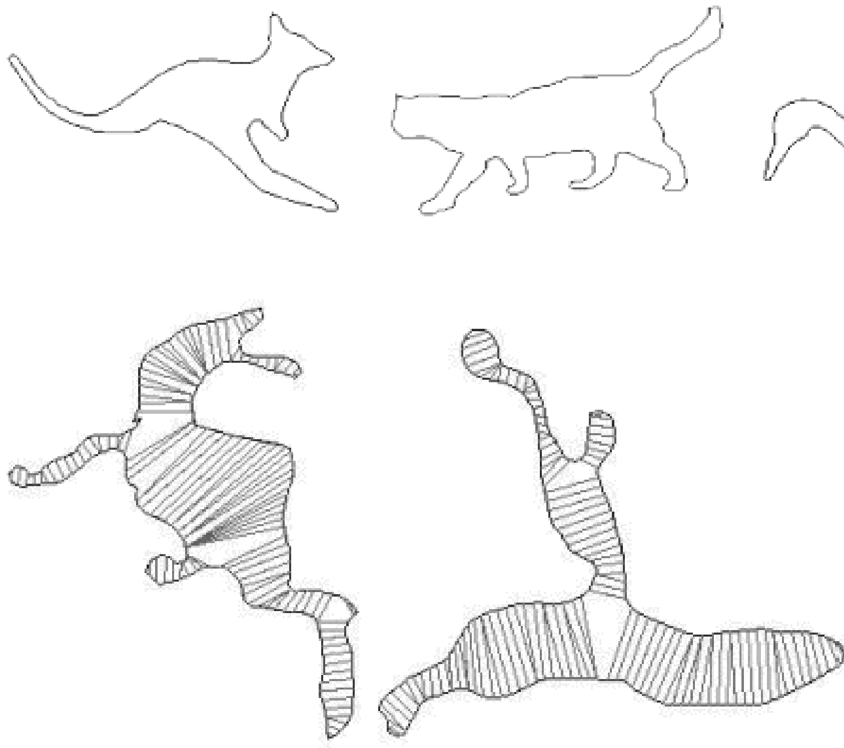

Figure 3. Top row: example animal shapes. Bottom row: shapes drawn from the minimax entropy distribution learned from shapes like those in the top row.

Figure 4. A series of increasingly difficult segmentation tasks.

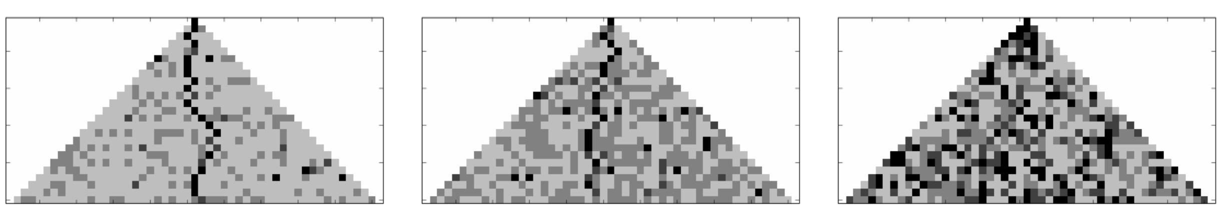

Figure 5. Three examples of increasingly difficult problems of tracking a road in clutter. From left to right, the order parameter $K = 0.8647, 0.2105$, and -7.2727. The negative K found for the right image suggests that *no* algorithm would be able to successfully track the road. The high K of the left image denotes that it would require less computation to track the road than in the middle panel.