# Improving Edge Detectors on Compressed Images—a Trainable Markov Random Field Approach

Davin Milun and David Sher
Computer Science Department
SUNY at Buffalo
226 Bell Hall
Buffalo NY 14260

milun@cs.buffalo.edu sher@cs.buffalo.edu

September 30, 1992

**Abstract**

We use Markov random fields to improve the output of the thinned Sobel edge detector, applied to images compressed using the JPEG technique. JPEG compression saves a lot of file space, however it introduces correlated errors into the images. This is exactly a circumstance for which our recently developed double neighborhood MRFs are suited (Milun and Sher, 1992a). Double neighborhood MRFs are constructed by sampling from pairs of original images together with noisy imagery. Thus we create a probability density function for pairs of neighborhoods across both images. This models the noise within the MRF probability density function without having to make assumptions about its form. This provides an easy way to generate Markov random fields for annealing or other relaxation methods.

We train the double neighborhood MRF on true edge-maps and edge-maps generated as output of a Sobel edge detector (Duda and Hart, 1973) on compressed images. Our method improves the generated edge-maps as verified by visual inspection, counting the number of misclassified pixels, and Pratt's figure of merit for edge detectors (Pratt, 1978).

## Key words and phrases

Markov random field; training; image reconstruction; image compression; JPEG; edge detection.

$$
\begin{array}{llllll}
\boxed{\begin{matrix} I & I & I \\ I & I & I \\ I & I & I \end{matrix}} & : 0.21667 &
\boxed{\begin{matrix} I & I & B \\ I & I & B \\ I & I & B \end{matrix}} & : 0.11250 &
\boxed{\begin{matrix} B & I & I \\ B & I & I \\ B & I & I \end{matrix}} & : 0.11806 \\
\boxed{\begin{matrix} B & B & B \\ I & I & I \\ I & I & I \end{matrix}} & : 0.09722 &
\boxed{\begin{matrix} I & I & I \\ B & B & B \\ I & I & I \end{matrix}} & : 0.07777 &
\boxed{\begin{matrix} B & B & B \\ B & B & B \\ B & B & B \end{matrix}} & : 0.00000 \\
\boxed{\begin{matrix} I & I & I \\ I & I & B \\ I & I & B \end{matrix}} & : 0.00277 &
\boxed{\begin{matrix} I & B & B \\ I & B & I \\ B & B & I \end{matrix}} & : 0.00139 &
\boxed{\begin{matrix} I & I & B \\ B & B & B \\ I & I & I \end{matrix}} & : 0.01111
\end{array}
$$

Figure 1: Examples of neighborhoods with probabilities. (These are probabilities from an image of overlapping, upright rectangles.)

# 1 Introduction

We are studying segmentation of imagery with correlated, difficult to characterize, noise. We chose lossy compressed imagery as a domain because the noise that is added by the compression algorithm is correlated. Segmenting such imagery is also important because JPEG is a technique that is becoming more popular, and some way is needed to remove any artifacts that the compression introduces into the segmentation.

We chose the JPEG technique for lossy compression because it is a standardized image compression mechanism (Wallace, 1991). JPEG is an acronym for Joint Photographic Experts Group, the name of the committee that developed the standard. The degree of image compression is usually quite large, but one pays for this in that the compressed image is not identical to the original one. However it is usually close enough that a human can't tell the difference, because it was developed to remove those features of the image that are not important to the human visual system.

Many image processing operators and techniques are sensitive to the the noise introduced by JPEG. Our method of double neighborhood Markov random fields can remove many of the errors in detection resulting from the image noise and the compression.

Markov random fields are one of the popular techniques for edge relaxation, image restoration and image segmentation. A Markov random field describes the local structure of an image by defining a neighborhood system on the pixels of the image and a probability density function over the labelings of the pixels. For example, to label the pixels of the image as either containing an object boundary $B$ or lying entirely in the interior of an object $I$. The Markov random field describes the probability of various local structures of $I$'s and $B$'s—for example figure 1 shows local neighborhoods at various levels of probability.

A Markov random field, even one with small local neighborhoods, requires numerous probabilities or clique potentials to describe it. Most researchers estimate these probabilities by an ad hoc scheme (Geman and Geman, 1984; Marroquin *et al.*, 1985; Swain, 1990). A natural technique for estimating these marginal probabilities is to use a training set of typical images and compute the sample frequencies for the neighborhoods in these images. These estimated marginal probabilities will represent a true MRF if the images from which they are sampled are MRFs.

If the images are not true MRFs, then the marginals estimated by sampling will not be true MRF marginals. This is important for some techniques, such as simulated annealing, but not important for others, such as edge relaxation, which only desire to find the locally optimal solution. However the assumption that the images are Markov, or nearly Markov, is not unreasonable when doing MRF research. Even so, MRF algorithms may result in adequate image interpretations even when the distribution is not an MRF. We are currently investigating a technique for converting these sampled marginals into true MRF marginals.

Sampling can introduce errors in calculating the probability of rare local neighborhoods. The techniques of Hancock and Kittler (Hancock and Kittler, 1990) can ameliorate these errors. We have previously studied their approach (Milun and Sher, 1991a; Milun and Sher, 1991b; Sher and Milun, 1991; Milun and Sher, 1992b; Milun and Sher, 1992c).

In the literature, for example (Geman and Geman, 1984; Blake, 1989; Szeliski, 1989; Swain, 1990; Milun and Sher, 1991a; Sher, 1991c), researchers dealing with MRFs have generally assumed that the following three items are available to the optimizer: (Note: When we say "optimizer", we refer generally to whatever piece of software that is being used to process the field, such as a simulated annealer or a gradient descent algorithm.)

**I1** The corrupted image

**I2** The MRF probability density function

**I3** A model for the noise

**I1** is trivial, and so much of the research has been concentrated on generating good models and representation schemes for **I2**. And as for **I3**, it has always been assumed that

**A1** The noise has a simple model, such as independent bit errors

**A2** The noise model is accurately given to the optimizer

However, **A1** and **A2** are not accurate for real world applications. **A1** is not necessarily true, and is definitely false in the case of binary images generated by edge detectors, as their errors are almost always correlated in some way. **A2** is actually a very non-intuitive assumption: if one is given an image, it is an extremely non-trivial task to obtain from that image a model of the noise, so that it can be given to the optimizer.

So we now solve the problems of **A1** by modifying **I2**. By sampling neighborhoods from noisy imagery together with true imagery, we encode the noise model right into the MRF probability density function, forming a double neighborhood MRF. This neighborhood structure permits a general model of the noise, and a natural means of estimating the noise model.

Consider the example of the output of an edge detector. Even state of the art edge detectors will not give exactly the correct edges of an image—there will be gaps in the edges, spurious edges, etc. However, let us assume that the true edges of the image are available. We build an MRF with a double neighborhood system, incorporating information from both the noisy and true edges. We build it by sampling, but this time from both images at once—a $3\times3$ neighborhood from the given, noisy output of the edge detector, and the $3\times3$ neighborhood in the same position in the image of true edges. See figure 2 for an example of a neighborhood pair.

The power of the method is that it can be applied to any area in which one has a set of input examples and true output values corresponding to those examples, a limitation being that the information to be extracted must be local in nature.

## 2 Background

MRF models are of continuing interest as a basis for image analysis. Their main use has been in image reconstruction and segmentation. Various estimators have been used to consider a reconstructed field as a "good fit" to an MRF probability field, for example, MAP (maximum *a posteriori*) (Besag, 1974; Besag, 1986; Geman and Geman, 1984; Goutsias, 1989; Jeng and Woods, 1990; Karssemeijer, 1990; Geman *et al.*, 1990; Goutsias, 1991; Johnson *et al.*, 1991; Tan *et al.*, 1992; Bedini and Tonazzini, 1992; Lugosi, 1992; Banerjee and Rosenfeld, 1992), and MPM (most probable marginals) (Marroquin, 1985;

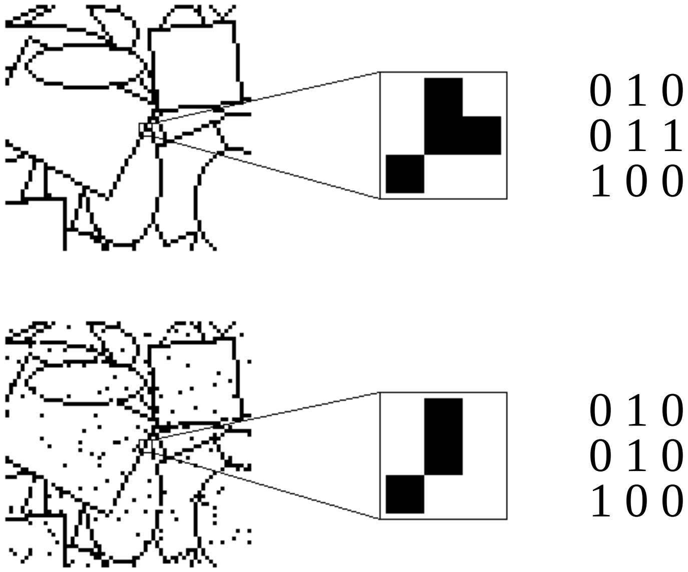


Figure 2: Example of a double neighborhood probability. The uncorrupted image is shown above, and the noisy, corrupted image is shown below. Enlarged is one neighborhood pair. The existence of this particular neighborhood pair would be recorded in the probability density function in position $010010100010011100_2$ (reading the corrupted image as the high bits) which is decimal 75932.

Marroquin *et al.*, 1985; Sher, 1991b; Geiger and Girosi, 1991); and various methods of generating these estimates have been used, for example, simulated annealing (Geman and Geman, 1984; Bedini and Tonazzini, 1992; Blake, 1989; Metropolis *et al.*, 1953; Basu *et al.*, 1989; Lakshmanan and Derin, 1989), Highest Confidence First (Chou, 1987), ICM (Besag, 1974; Besag, 1986; Geman *et al.*, 1990; Karssemeijer, 1990) and refinements and modifications to parts of these methods (Swain, 1990; Sher, 1991b; Johnson *et al.*, 1991). However, not much attention has been paid to the methods of deriving the probability density functions.

Simulated annealing is a technique that overcomes the requirement that the search space be convex, and has been extensively studied in the domain of image reconstruction and edge detection (Jeng and Woods, 1990; Blake, 1989; Geman and Geman, 1984; Marroquin *et al.*, 1987). Simulated annealing converges with high probability to the MAP probable labeling. But in all the simulated annealing work, the authors assume the existence of a density function, usually in the form of clique potentials, with which to anneal. We do not use clique potentials for defining our MRFs.

We generate the MRF probability density functions by sampling. But there is another way to generate it for various domains. Sher (Sher, 1991a; Sher, 1991c) approaches this problem of generating the density function by theoretically dropping objects of known shape into an image, and calculating what the probabilities of the various neighborhoods would be.

Lakshmanan and Derin's (Lakshmanan and Derin, 1989) goals are similar to ours. They try for a data-driven approach, with almost no external parameters. However they still require the noise to be independent. They estimate the field from the image itself, using two interleaved simulated annealing processes. They do not derive the field from a training set as we do. Manjunath and Chellappa's approach (Manjunath and Chellappa, 1991) is similar to Lakshmanan and Derin's, but they show that a nearest neighbor technique works almost as well as simulated annealing.

Veijanen (Veijanen, 1991) discovers Markov random fields by searching for parameters of a field that would generate the same cooccurance matrix as provided by sample images. This work is similar in goals and techniques to our own but we feel this approach is unnecessarily indirect.

Figure 3: Some instances of generated gray scale random images.

Kim and Yang (Kim and Yang, 1991) use the error backpropagation technique from neural networks to estimate the MRF parameters.

Another important and exciting approach to training a visual feature discrimination is that of Kleinberg (Kleinberg, 1990). His technique uses a training set of examples and counter examples and information about the coherence of the data to generate a set of low quality classification algorithms. Taken together these low quality algorithms can act as a high quality, robust classification algorithm.

Many successful character recognition systems rely on training phases. Some successful examples are Kushnir's work (Kushnir *et al.*, 1985), and Baird's work (Kahan *et al.*, 1987; Lam and Baird, 1986). The early work on this approach was surveyed by Nagy (Nagy, 1982). Hull used statistics gathered from images of words to demonstrate constraints on word recognition systems (Hull, 1987). These successes motivated our approach.

# 3 Results

Our test data generator for edge detection enhancement creates pairs of gray scale images. Each pair consists of randomly placed, randomly sized, circles of random solid color, and an edge map representing those pixels through which an edge passes. Some examples of these images are shown in figure 3. Our training set contains 128 such pairs of gray scale images, and our test set contains 128 different pairs. Each image was corrupted with Gaussian noise of standard deviation 16 to each of these images—we call this set $G$. The clean and noisy images were corrupted with JPEG compression down to 1.5 bits per pixel (bpp)—generating sets $J$ and $GJ$ respectively. We then ran a thresholded, thinned Sobel (Duda and Hart, 1973) edge detector on the 3 sets of corrupted images. We trained the double neighborhood MRF on the pairs of true edge images and elements of $G$, $J$ and $GJ$ to generate 3 MRFs. An example image from each of the 3 sets can be found in figure 4.

A gradient descent algorithm was applied to the $G$, $J$ and $GJ$ versions of the test set, at a variety of Hancock-Kittler smoothing levels. Figure 5 shows the results on an image from $J$. Our methods remove a significant number of the false edges. Figure 6 shows the results on two images from $GJ$. Our methods remove a large number of the false edges.

In order to gauge the performance of our method, we used two different measures. One is the metric of the number of bits that differ between the reconstructed image and the true image. The results from this do not, however, always match the visual sense of what is best. The second measure we used is Pratt's Figure of Merit (Pratt, 1978). This is a number in range 0 to 1, which represents how good an edge detector is, with 1 being perfect. It takes both edge position and accuracy into account.

Figure 4: Noisy images. From left to right: a noiseless image; an image with standard deviation 16 Gaussian noise added (from set $G$); an image compressed by JPEG to 1.5 bpp (from set $J$); both JPEG and Gaussian noise (from set $GJ$); the true edge in this image

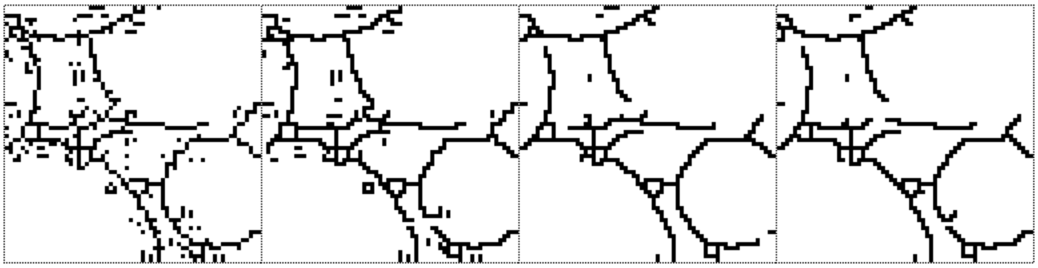

Figure 5: Sample run of double neighborhood MRF used for Edge Enhancement on an artificial image of circles which was compressed by JPEG to 1.5bpp (from Set $J$). The 4 images shown are, in left to right order: the output of the thresholded, thinned Sobel; a reconstruction with no HK smoothing; a reconstruction with 0.010 HK smoothing; and a reconstruction with 0.100 HK smoothing.

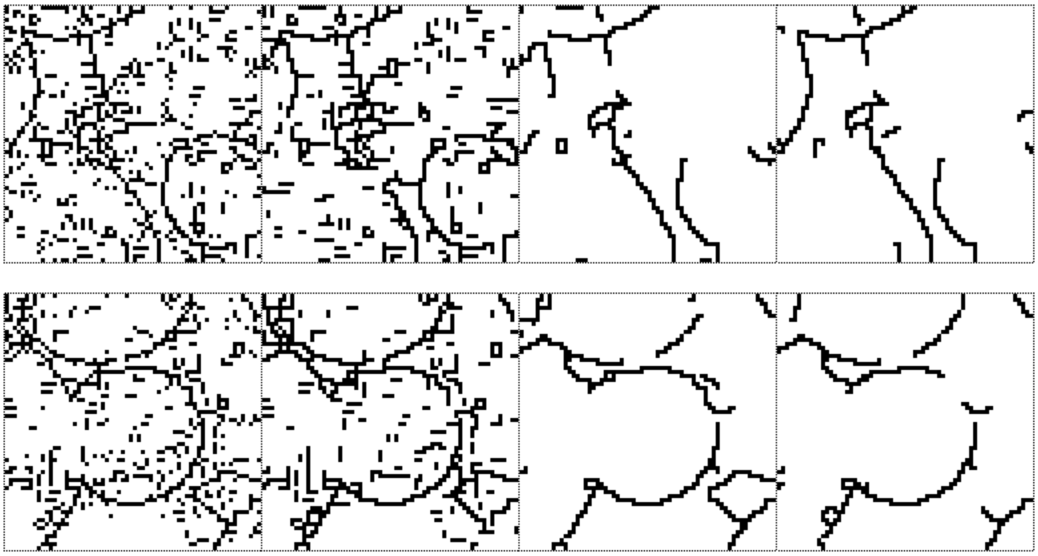

Figure 6: Sample runs of double neighborhood MRF used for Edge Enhancement on an artificial image of circles with standard deviation 16 noise added which was compressed by JPEG to 1.5bpp (from Set $GJ$). The 4 images shown in each strip are, in left to right order: the output of the thresholded, thinned Sobel; a reconstruction with no HK smoothing; a reconstruction with 0.010 HK smoothing; and a reconstruction with 0.100 HK smoothing.

Using these two measures, we found that our technique of reconstruction shows an improvement over the thinned Sobel. We have found that using the Hancock-Kittler smoothing technique gained a significant improvement, both in a numerical sense, and by visual inspection. Especially in the case of the $GJ$ set (compressed noisy images), the improvement is notable: the thinned Sobel has an mean of about 650 bits incorrect with a standard deviation of about 100; our method, with Hancock-Kittler smoothing, has a mean of just below 300 bits incorrect with a standard deviation of about 75. See table 1 for detailed numerical results.

# 4 Conclusions

We have applied our technique of double neighborhood Markov random fields to the output of the thinned Sobel edge detector, applied to lossily compressed images. Double neighborhood MRFs are useful in this domain, because the errors are correlated, and double neighborhood MRFs incorporate a general noise model within their probability density function. Our method improves the quality of the resulting edge detection algorithm. We plan to apply double neighborhood MRFs to gray scale images in the future.

| Noise Level | Thinsobel | Double neighborhood MRF | | | | | |
|---|---|---|---|---|---|---|---|
| | | no HK | HK=.060 | HK=.010 | HK=.025 | HK=.050 | HK=.100 |
| Pratt's Figure of Merit | | | | | | | |
| No noise | 0.7641 | 0.8813 | 0.8865 | 0.8865 | 0.8879 | 0.8882 | 0.8874 |
| Noise with $\sigma = 16$ (Set $G$) | 0.5679 | 0.6353 | 0.6301 | 0.6329 | 0.6356 | 0.6338 | 0.6387 |
| JPEG at 1.5bpp (Set $J$) | 0.6563 | 0.6875 | 0.7058 | 0.7137 | 0.7231 | 0.7369 | 0.7516 |
| Noise and JPEG (Set $GJ$) | 0.4300 | 0.4665 | 0.5320 | 0.5336 | 0.5227 | 0.5299 | 0.5242 |
| Bits of Error | | | | | | | |
| No noise | 168.47 | 146.46 | 138.09 | 138.44 | 133.87 | 132.71 | 130.97 |
| Noise with $\sigma = 16$ (Set $G$) | 238.88 | 222.49 | 222.27 | 219.44 | 217.68 | 215.63 | 215.43 |
| JPEG at 1.5bpp (Set $J$) | 352.73 | 322.70 | 240.76 | 237.70 | 234.52 | 232.32 | 226.70 |
| Noise and JPEG (Set $GJ$) | 648.80 | 590.23 | 301.04 | 297.27 | 290.23 | 286.81 | 278.18 |

Table 1: Numerical results for the test set of images of circles. The standard deviations of the above means are roughly the following:

| | Thinsobel and no HK | | HK> 0 | |
|---|---|---|---|---|
| | Pratt | Bits | Pratt | Bits |
| No noise | 0.025 | 38 | 0.025 | 34 |
| Noise with $\sigma = 16$ (Set $G$) | 0.10 | 56 | 0.11 | 58 |
| JPEG at 1.5bpp (Set $J$) | 0.05 | 84 | 0.08 | 62 |
| Noise and JPEG (Set $GJ$) | 0.075 | 100 | 0.10 | 73 |

## References


(Banerjee and Rosenfeld, 1992) Saibal Banerjee and Azriel Rosenfeld. MAP estimation of context free grammars. *Pattern Recognition Letters*, 13:95–101, Feb 1992.

(Basu *et al.*, 1989) Sujit Basu, M. Mohan, and Vijay K. Agarwal. Estimation of cloud microstructural parameters from satellite imagery: A Markov random field approach. *Mausam*, 40(3):243–246, July 1989.

(Bedini and Tonazzini, 1992) Luigi Bedini and Anna Tonazzini. Image restoration preserving discontinuities: the Bayesian approach and neural networks. *Image and Vision Computing*, 10(2):108–118, March 1992.

(Besag, 1974) J.E. Besag. Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal Royal Statistical Society, Series B*, 36:192–236, 1974.

(Besag, 1986) J.E. Besag. On the statistical analysis of dirty pictures. *Journal Royal Statistical Society, Series B*, 48:259–302, 1986.

(Blake, 1989) Andrew Blake. Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(1):2–12, January 1989.

(Chou, 1987) Paul Chou. Multi-modal segmentation using Markov random fields. In *IJCAI*, July 1987.

(Duda and Hart, 1973) Richard O. Duda and Peter E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, Inc., New York, London, Sydney, Toronto, 1973.

(Geiger and Girosi, 1991) D. Geiger and F. Girosi. Parallel and deterministic algorithms from MRF's: Surface reconstructions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-13(5):401–413, May 1991.

(Geman and Geman, 1984) Stuart Geman and Donald Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, November 1984.

(Geman *et al.*, 1990) D. Geman, S. Geman, C. Graffigne, and P. Dong. Boundary detection by constrained optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):609–628, July 1990.

(Goutsias, 1989) J. K. Goutsias. Mutually compatible Gibbs random fields. *IEEE Transactions on Information Theory*, 35(6):1233–1249, November 1989.

(Goutsias, 1991) J. K. Goutsias. A theoretical analysis of Monte Carlo algorithms for the simulation of Gibbs random field images. *IEEE Transactions on Information Theory*, IT-37(6):1618–1628, November 1991.

(Hancock and Kittler, 1990) E.R. Hancock and J. Kittler. A label error process for discrete relaxation. In *Proceedings of the 10th international conference on pattern recognition*, pages 523–527. IEEE, June 1990.

(Hull, 1987) J.J. Hull. *A computational theory of visual word recognition*. PhD thesis, SUNY at Buffalo, Department of Computer Science, 1987.

(Jeng and Woods, 1990) F.C. Jeng and J.W. Woods. Simulated annealing in compound Gaussian random fields. *IEEE Transactions on Information Theory*, 36(1):94–107, January 1990.

(Johnson *et al.*, 1991) V. E. Johnson, W. H. Wong, X. Hu, and C. T. Chen. Image restoration using gibbs priors: Boundary modeling, treatment of blurring and selection of hyperparameter. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-13(5):413–425, May 1991.

(Kahan *et al.*, 1987) Simon Kahan, Theo Pavlidis, and Henry S. Baird. On the recognition of printed characters of any font and size. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(2):274–288, March 1987.

(Karssemeijer, 1990) N. Karssemeijer. A relaxation method for image segmentation using a spatially dependent stochastic model. *Pattern Recognition Letters*, 11:13–23, 1990.

(Kim and Yang, 1991) Il Young Kim and Hyun Seung Yang. Markov random field based image labeling with parameter estimation by error backpropagation. *IEICE Transactions on Information and Systems*, E74(10):3513–3521, October 1991.

(Kleinberg, 1990) E. M. Kleinberg. Stochastic discrimination. *Annals of Mathematics and Artificial Intelligence*, 1:207–239, 1990.

(Kushnir *et al.*, 1985) M. Kushnir, K. Abe, and K. Matsumoto. Recognition of handprinted Hebrew characters using features selected in the Hough transform space. *Pattern Recognition*, 18(2):103–114, 1985.

(Lakshmanan and Derin, 1989) S. Lakshmanan and H. Derin. Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(8):799–813, August 1989.

(Lam and Baird, 1986) S. W. Lam and H. S. Baird. Performance testing of mixed-font, variable-size character recognizers. Technical Report Technical Report No. 126, AT&T Bell Laboratories Computing Science, November 1986.

(Lugosi, 1992) Gabor Lugosi. Learning with an unreliable teacher. *Pattern Recognition*, 25(1):79–85, 1992.

(Manjunath and Chellappa, 1991) B. S. Manjunath and R. Chellappa. Unsupervised texture segmentation using Markov random field models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-13(5):478–482, May 1991.

(Marroquin *et al.*, 1985) J. Marroquin, S. Mitter, and T. Poggio. Probabilistic solution of ill-posed problems in computational vision. In Lee S. Bauman, editor, *Proceedings: Image Understanding Workshop*, pages 293–309. Science Applications International Corporation, December 1985.

(Marroquin *et al.*, 1987) J. Marroquin, S. Mitter, and T. Poggio. Probabilistic solution of ill-posed problems in computational vision. *Journal Am. Stat. Assn.*, 82:76–89, 1987.

(Marroquin, 1985) Jose Luis Marroquin. Probabilistic solution of inverse problems. Technical Report TR 860, MIT Artificial Intelligence Laboratory, September 1985.

(Metropolis *et al.*, 1953) Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, June 1953.

(Milun and Sher, 1991a) D. Milun and D. B. Sher. Improving sampled probability distributions for Markov random fields. Technical Report 91-08, S.U.N.Y. at Buffalo, July 1991.

(Milun and Sher, 1991b) Davin Milun and David Sher. Improving Markov random field reconstructions by using noisy probabilities. In *6th International Conference on Image Analysis and Processing*, September 1991.

(Milun and Sher, 1992a) D. Milun and D.B. Sher. Learning structural and corruption information from samples for markov random field edge detection enhancement. In *The SPIE 1992 International Symposium on Optical Applied Science and Engineering: Neural and Stochastic Methods in Image and Signal Processing*, San Diego, 1992.

(Milun and Sher, 1992b) Davin Milun and David Sher. Improving sampled probability distributions for Markov random fields. *Submitted to Pattern Recognition Letters*, 1992.

(Milun and Sher, 1992c) Davin Milun and David Sher. Learning structural and corruption information from samples for Markov random field binary image reconstruction. In *11th ICPR: Conference on Image, Speech and Signal Analysis*, volume 3, pages 513–516, The Hague, The Netherlands, September 1992.

(Nagy, 1982) George Nagy. Optical character recognition: theory and practice. In P. R. Krishnaiah and L. N. Kanal, editors, *Handbook of Statistics*, volume 2, pages 621–649. 1982.

(Pratt, 1978) W. K. Pratt. *Digital Image Processing*. Wiley-Interscience, New York, 1978.

(Sher and Milun, 1991) David Sher and Davin Milun. Improving Markov random fields by using noisy probabilities. *Submitted to the The Annals of Mathematics and Artificial Intelligence*, 1991.

(Sher, 1991a) David Sher. Deriving the distribution of edge neighborhoods from a library of occluding objects. In *Bar-Ilan Symposium on the Foundations of Artificial Intelligence*, June 1991.

(Sher, 1991b) David Sher. Minimizing the cost of errors with a Markov random field. *Pattern Recognition Letters*, 12(2), February 1991.

(Sher, 1991c) D.B. Sher. A technique for deriving the distribution of edge neighborhoods from a library of occluding objects. In *6th International Conference on Image Analysis and Processing*, September 1991.

(Swain, 1990) Michael J. Swain. Parameter learning for Markov random fields with highest confidence first estimation. Technical Report TR350, Department of Computer Science, University of Rochester, August 1990.

(Szeliski, 1989) R. Szeliski. *Bayesian Modeling of Uncertainty in Low-level Vision*. Kluwer Academic Publishers, 1989.

(Tan *et al.*, 1992) H. L. Tan, S. B. Gelfand, and E. J. Delp. A cost minimization approach to edge detection using simulated annealing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-14(1):3–19, January 1992.

(Veijanen, 1991) A. Veijanen. A simulation-based estimator for hidden Markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-13(8):825–829, August 1991.

(Wallace, 1991) Gregory K. Wallace. The JPEG still picture compression standard. *Communications of the ACM*, 34(4):30–44, April 1991.