Evidence Representation & Combination in Low-level Vision

E-ren Chuang and David Sher

Department of Computer Science

State University of New York at Buffalo

226 Bell Hall

Buffalo, NY 14260.

SUMMARY

This paper proposes statistic as low-level visual evidence, and shows the elegant meaning of summing rule[1] in integrating these evidence. Evidence is provided by a χ^2 test of local visual feature such as edge, or corner. A feature ω is modeled by a $1 \times n$ multivariate normal random vector $\mathbf{X}_\omega = (X_1, \ldots, X_n)'$, denoted by $\mathbf{X}_\omega \sim N_n(\mu_{\mathbf{X}_\omega}, \mathbf{\Sigma}_{\mathbf{X}_\omega})$, where $\mu_{\mathbf{X}_\omega}$ is mean vector which represents the intensity of pixels in a template, and $\mathbf{\Sigma}_{\mathbf{X}_\omega}$ is variance-covariance (or dispersion) matrix which includes the intrinsic structure of the feature ω with additive Gaussion noise [15]. After the transformation

$$f(\mathbf{X}) : (\mathbf{X} - \mu_{\mathbf{X}})^{\mathbf{t}} \mathbf{\Sigma}_{\mathbf{X}}^{-1} (\mathbf{X} - \mu_{\mathbf{X}})$$

this model becomes a random variable with χ^2 distribution of n degree of freedom [16], then χ^2 test is applied to measure the similarity between data and the expectation vector of each model[15].

Let $X = (\mathbf{X} - \mu_{\mathbf{X}})^{\mathbf{t}} \mathbf{\Sigma}_{\mathbf{X}_\omega}^{-1} (\mathbf{X} - \mu_{\mathbf{X}})$, then X is called the squared *Mahalanobis* distance[17] (MD) between the mean vector $\mu_{\mathbf{X}}$ of $\mathbf{X}_\omega$ and the data $\mathbf{x}$. Experimental results show that the relationship between signal-noise-ratio (SNR) and the MD can be approximated by a branch of a hyperbola, and the relationship between the variance of Gaussian noise (VGN) and SNR also can be approximated by a branch of another hyperbola, so MD and VGN has linear relationship. In another words, when the noise level is high, the data will be farther from the ideal model (i.e. *Mahalanobis distance is longer*).

In 1890, William James emphasized summing rule for brain activity[1]. In **Theorem 1**, we prove that the summing rule can not only combine evidence from χ^2 tests, but also integrate the geometric meanings of these addends in the sum. Besides, the sum is also a χ^2 statistic, therefore, the χ^2 test of different degree of freedom can consistently apply on the sum to verify the significance of the sum. We apply χ^2 values and summing rule to Hough transform, instead of accumulating edge points. Experimental results show that these combined statistics signifcantly improve the performance of Hough transform on detecting broken line segments, and avoid making errors. Therefore, summing rule could be one of the powerful operations in to combine the evidence with uncertainty.

Abstract

We discuss using statistics from χ^2 tests and summing rule[1] for evidence representation and combination in low-level vision. Feature likelihoods are represented by statistics which are assured from χ^2 tests, and the relationship between these likelihoods and noise level is linear. Multiple statistics are combined by summing rule into statistics for larger features such as segments and polygons. This is justified because the sum on a set of independent χ^2 random variables is also a χ^2 random variable, and the geometric meaning of the sum is equal to the integration of these addends. Examples show that the performance of Hough transform for detecting line segments is significantly improved if statistics are included. Therefore, statistics which adopted as evidence can not only include uncertainty, but also help to avoid error in Hough transform. The summing rule can combine statistics from χ^2 tests, and also can integrate the meanings of addends into the sum.

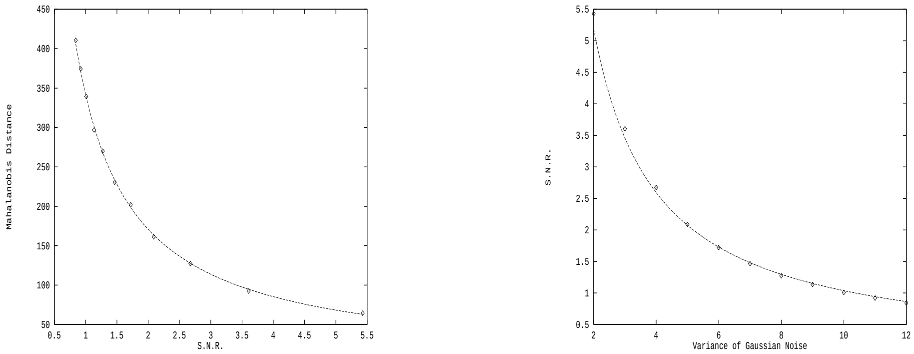
Key words: Evidence Combination χ^2 Statistic Hough Transform Summing Rule

1 Introduction

It is the current trend to include uncertainty in low-level vision [2] because there are shading, blurring, and noise in real images. The Hough transform (HT)[3], Dempster-Shafer theory of evidence[4] and Bayesian network[5] are notable methods for evidence accumulation and combination. The HT projects evidence, boundary points in an image, back to some cells in a parameter space of a pattern[6,7]. For example, the line, $y = ax + b$, has two parameters, a, and b. For each edge point, $(x_i, y_i), i = 1, \ldots, m$, the votes in the cells $c(a_i', b_i')$, i=1,. . . , n, are incremented by one if $y = a'_i x + b'_i$. Techniques are devised to search cells with more votes than a threshold[8,9]. These cells indicates a possible instance of the pattern. The performance of HT is analyzed in[10,11,12]. Although HT can be viewed as a procedure of evidence gathering, the evidence, edge points, does not include how accurate they are.

Bayesian formalism is another important branch of evidence aggregation. Szeliski[13] chose a particular power spectrum in Fourier domain for *a prior* in Bayes' rule. Sher[14] applies Bayes' rule to compute the conditional probability of a event, and obtains the likelihood ratio of this event. Pearl[5] applies a Bayesian formalism for evidence aggregation and propagation in a hierarchy of hypotheses. The likelihood ratio, α, is used to represent the degree of confirmation or disconfirmation of an event e. The original likelihood ratio of a hypothesis, S, in the hierarchy is updated with respect to α_S. α_S is passed up and down to the father and the successor of S respectively, and their belief is also updated accordingly. The probabilistic justification of the process is based on the assumption, conditional independence: $P(e|S, h_i) = P(e|S)$, $h_i \in S$, where h_i is one of S's constituents. The above process has the the following advantages: distributed computation in the hierarchy, meaningful parameter α, and transparent inferences.

In this paper, evidence is represented by statistics from χ^2 hypothesis tests of local visual features instead of probabilities which is usually applied to predict the outcome of an experiment. Statistics are combined by summing rule[1]. These combined statistics are verified again by χ^2 tests. It has demonstrated by examples that the performance of



(a) (b)

Figure 1: In (a), the relationship between signal-noise-ratio (SNR) and *Mahalanobis* distance (MD) is shown by small boxes. This relationship is approximated by a branch of the hyperbola $SNR \times MD = 341.51$, plotted by a dashed line. In (b) is depicted the relationship between the variance of gaussian noise (VGN) and SNR, which is a branch of another hyperbola $SNR \times VGN = 10.37$.

Hough transform on line segment detection can be significantly improved if χ^2 statistics and summing rule are included.

2 Evidence Representation

In statistics, there are tools to measure how close the outcome of an experiment to a model. The closeness (in the sense of probability) between the model and the outcome is represented by the statistic from those tools. Therefore, we chose statistics as evidence in low-level vision. Evidence is provided by a χ^2 test of local visual feature such as edge, or corner. A feature ω is modeled by a $1 \times n$ multivariate normal random vector $\mathbf{X}_\omega = (X_1, \ldots, X_n)'$, denoted by $\mathbf{X}_\omega \sim N_n(\mu_{\mathbf{X}_\omega}, \mathbf{\Sigma}_{\mathbf{X}_\omega})$, where $\mu_{\mathbf{X}_\omega}$ is mean vector which represents the intensity of pixels in a template, and $\mathbf{\Sigma}_{\mathbf{X}_\omega}$ is variance-covariance (or dispersion) matrix which includes the intrinsic structure of the feature ω with additive Gaussion noise [15]. After the transformation

$$f(\mathbf{X}) : (\mathbf{X} - \mu_{\mathbf{x}})^{\mathbf{t}} \mathbf{\Sigma}_{\mathbf{x}}^{-1} (\mathbf{X} - \mu_{\mathbf{x}})$$

this model becomes a random variable with χ^2 distribution of n degree of freedom [16], then χ^2 test is applied to measure the similarity between data and the expectation vector of each model.

Let $X = (\mathbf{X} - \mu_{\mathbf{x}})^{\mathbf{t}} \mathbf{\Sigma}_{\mathbf{x}_{\omega}}^{-1} (\mathbf{X} - \mu_{\mathbf{x}})$, then X is called the squared *Mahalanobis* distance[17] between the mean vector $\mu_{\mathbf{0}}$ of $\mathbf{X}_{\omega}$ and the data $\mathbf{x}$. In Fig. 1(a), the relationship of signal-noise-ratio (SNR) and the *Mahalanobis* distance (MD) is depicted. The *Mahalanobis* distances are provided by from χ^2 tests of 125 degree of freedom between an ideal horizontal edge of length 25 and data with difference gaussian noise. This relationship can be described by a branch of the hyperbola SNR $\times$ MD = 341.51. In Fig. 1(b), we also depict the relationship between the variance of the gaussian noise (VGN) and SNR. It is also a branch of another hyperbole VGN $\times$ SNR = 10.37, therefore, MD = 32.93 VGN, which means MD and VGN has linear relationship. In another words, when the noise level is high, the data will be farther from ideal model (i.e. *Mahalanobis* distance is longer).

3 Evidence Combination

In 1890, William James emphasized summing rule for brain activity[1]. He wrote: "*The amount of activity at any given point in the brain cortex is the sum of the tendencies of all other points to discharge into it*". We adopt the summing rule as the main operation of evidence combination so that χ^2 test can coherently apply to the sum, shown in theorem 1.

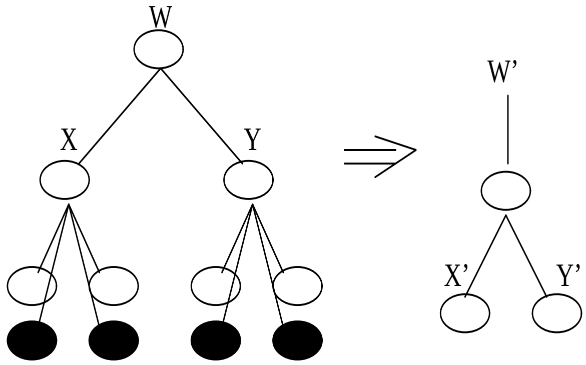


Figure 2: The pictorial display of Theorem 1.

Theorem 1. if $W = X + Y$, where X and Y are random variables (r.v.) transformed from two uncorrelated random vectors $\mathbf{X} \sim N_m(\mu_\mathbf{x}, \mathbf{\Sigma_x})$, and $\mathbf{Y} \sim N_n(\mu_\mathbf{y}, \mathbf{\Sigma_y})$ by $(\mathbf{X} - \mu_\mathbf{x})^\mathbf{t}\mathbf{\Sigma_x^{-1}}(\mathbf{X} - \mu_\mathbf{x})$, $(\mathbf{Y} - \mu_\mathbf{y})^\mathbf{t}\mathbf{\Sigma_y^{-1}}(\mathbf{Y} - \mu_\mathbf{y})$, respectively, then W has χ^2_{m+n} distribution.

proof:

Let $\mathbf{W} = (\mathbf{X}, \mathbf{Y}) \sim \mathbf{N}[(\mu_\mathbf{x}, \mu_\mathbf{y}), \mathbf{\Sigma_w}]$, where $\mathbf{\Sigma_w} = \begin{pmatrix} \mathbf{\Sigma_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma_y} \end{pmatrix}$ By **Gauss-Jordan elimination**, it can be easily proved that the inverse matrix $\mathbf{\Sigma_w}^{-1} = \begin{pmatrix} \mathbf{\Sigma_x^{-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma_y^{-1}} \end{pmatrix}$,

$$\Rightarrow W = (\mathbf{W} - \mu_\mathbf{w})^\mathbf{t}\mathbf{\Sigma_w^{-1}}(\mathbf{W} - \mu_\mathbf{w}) = \chi^2_{m+n}$$

$$= [(\mathbf{X} - \mu_\mathbf{x}), (\mathbf{Y} - \mu_\mathbf{y})]^\mathbf{t} \begin{pmatrix} \mathbf{\Sigma_x^{-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma_y^{-1}} \end{pmatrix} [(\mathbf{X} - \mu_\mathbf{x}), (\mathbf{Y} - \mu_\mathbf{y})]$$

$$= (\mathbf{X} - \mu_\mathbf{x})^\mathbf{t}\mathbf{\Sigma_x^{-1}}(\mathbf{X} - \mu_\mathbf{x}) + (\mathbf{Y} - \mu_\mathbf{y})^\mathbf{t}\mathbf{\Sigma_y^{-1}}(\mathbf{Y} - \mu_\mathbf{y})$$

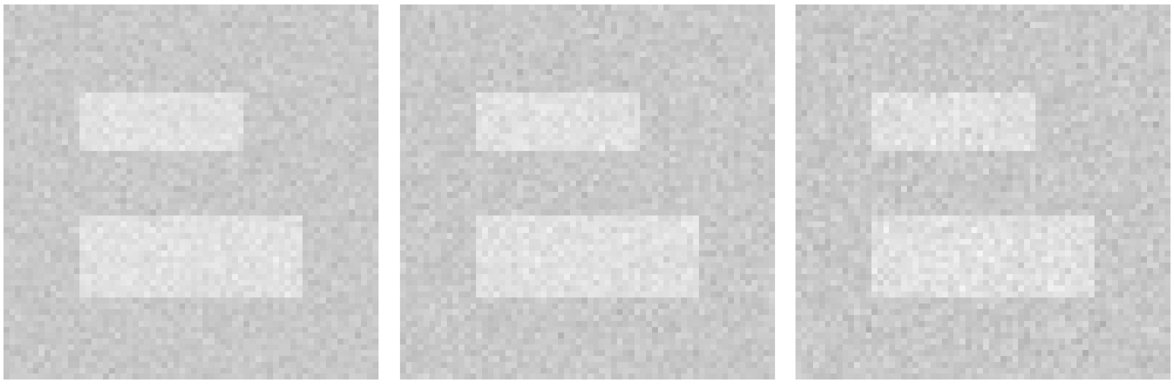
$$= X + Y = \chi^2_m + \chi^2_n$$

In the proof, it has also been shown that the random vector $\mathbf{W}$ models a bigger feature which contains two features modeled by random vectors $\mathbf{X}$ and $\mathbf{Y}$, shown in Fig. 2. The random variable W, which has χ^2_{m+n} distribution and is transformed from $(\mathbf{W}-\mu_{\mathbf{w}})^{\mathbf{t}}\mathbf{\Sigma}_{\mathbf{w}}^{-1}(\mathbf{W}-\mu_{\mathbf{w}})$, is the sum of X and Y transformed from $(\mathbf{X}-\mu_{\mathbf{x}})^{\mathbf{t}}\mathbf{\Sigma}_{\mathbf{x}}^{-1}(\mathbf{X}-\mu_{\mathbf{x}})$ and $(\mathbf{Y}-\mu_{\mathbf{y}})^{\mathbf{t}}\mathbf{\Sigma}_{\mathbf{y}}^{-1}(\mathbf{Y}-\mu_{\mathbf{y}})$ respectively. Therefore, we can apply χ^2 test on W, i.e. making decision upon the sum of these statistics provided from the χ^2 tests on X and Y.

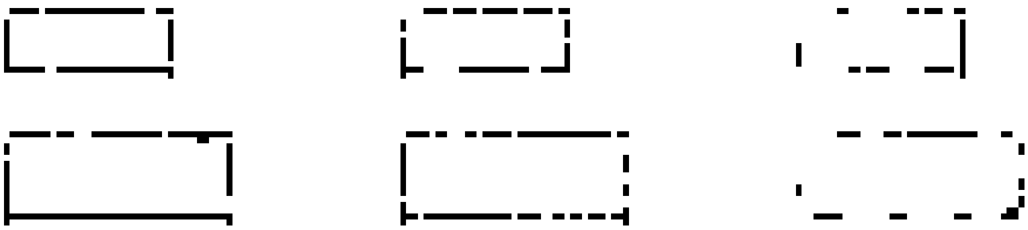
When a significance level is chosen, a threshold value is determined accordingly by looking up χ^2 table. However, it is very possible that the degree of freedom of a χ^2 distribution, n, is very large so that the threshold value is beyond the scope of tabulated χ^2 distribution. This problem can be solved by transforming W to a r.v. $Z=\sqrt{2X}-\sqrt{2n-1}\sim N(0,1)$ so that the threshold value is determined by looking up the table of standard normal distribution.

4 Experimentation

In this section, it is shown that statistics fom χ^2 tests can not only include uncertainty, but also avoid errors, and it can significantly improve the performance of Hough transform on detecting line segments. In order to reduce computation time and the size of Hough Transform parameter space, we only apply 2×7 edge detectors in horizontal and vertical direction on an image with gaussian noise, shown in Fig. 3(a), 3(b), and 3(c). From left to right, the SNR of them are 1.17, 0.95, 0.74, respectively. Given 0.1 probability that a true edge point will not be detected, 21.064 is chosen as a threshold by looking up χ^2 table of 14 degree of freedom. The edge points in Fig. 3(a), 3(b), and 3(c) are shown in 3(d), 3(e), and 3(f), accordingly.

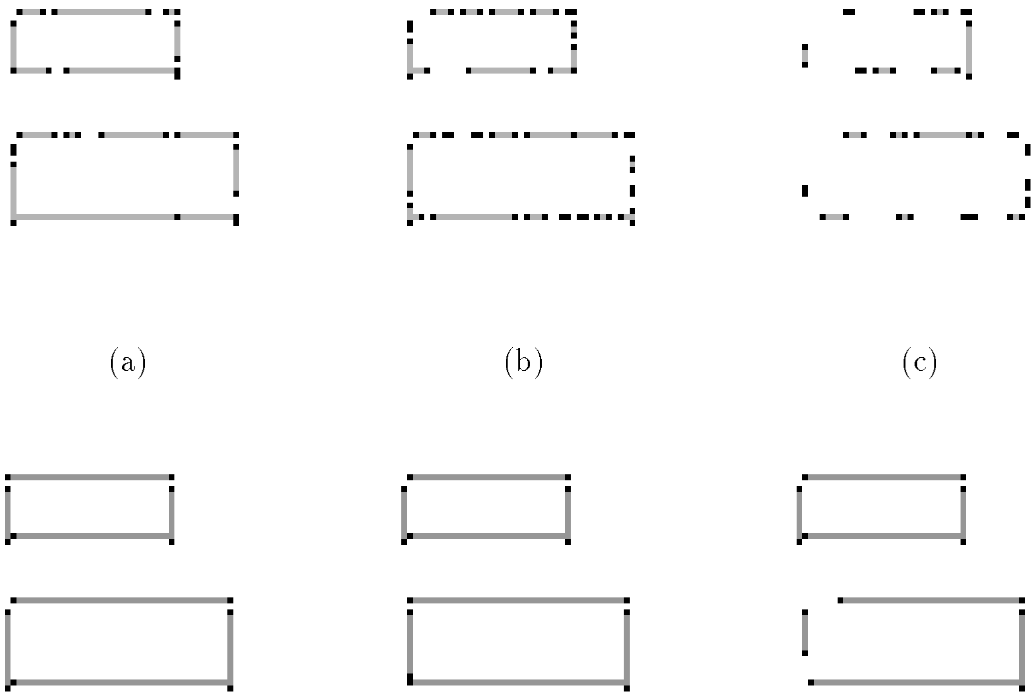


(a) (b) (c)



(d) (e) (f)

Figure 3: Three identical images with different signal-to-noise ratio 1.17, 0.95 and 0.74 are shown in (a), (b), (c), respectively. The results from χ^2 tests for feature detection are shown is (d), (e), (f) correspondingly.



(a) (b) (c)

(d) (e) (f)

Figure 4: The experimental results, shown in (d), (e), and (f), of the Hough transform which includes the statistics from χ^2 tests is better the results of traditional Hough transform on line segment detection, shown in (a), (b), and (c).

In Hough transform, edge points are evidence. These points are the results from edge detector after thresholding, so the accuracy of each edge point is not included. This effects the performance of Hough transform. For example, the traditional technique to find a line segment is to go along the line which corresponding cell in the parameter space of Hough transform has more votes than a threshold. Unfortunately, by this method, it is hard to verify whether these segments in the line belong to a broken segment, or they are colinear. If we apply the above technique on Fig. 3(f) which is the result of our edge detector on Fig. 3(c), and set the threshold of Hough transform to be six, then three vertical line segments are missing, and many pieces of line segments are found, shown in Fig. 4(c).

If we apply the following Hough transform to detect line segments, which parameter space is a set $S = \{(x, y, d, l) | \; x, y, d, l \in \mathbf{I}\}$, where x, y are the coordinates of a starting point of a line segment, d and l are the orientation and length of the segment. In our technique, each cell of S accumulates the χ^2 values from the associated edges, instead of counting the number of edge points. The results from applying this procedures is shown in Fig. 4(f). The threshold for the accumulated statistic in S is determined as follows: Suppose we apply this 2×7 detector on an image which contains a horizontla edge of length 20, and the accumulated statistic is 210.64 if every return values of χ^2 tests are one the edge of the chosen threshold 21.064. After transforming the χ^2 statistic to the standard normal distribution, we have 3.82 as the threshold for these transformed statistics. Besides, we also implement two filters to eliminate internal and overlapped line segments.

5 Conclusions

This paper proposes statistic as low-level visual evidence, and shows the elegant meaning of summing rule in integrating these evidence. These statistics, which are called *Mahalanobis* distance (MD), are obtained from χ^2 tests. Experiment results show the relationships between *Mahalanobis* distance and signal-and-noise (SNR), and variance of gaussian noise and SNR can be described by branches of hyperbolas, so the relationship of MD and VGN is linear, which means the MD between ideal feature and data is farther while the noise level increases.

The simple and meaningful summing rule has been adopted for evidence combination. In **theorem 1**, it has been proved that the sum on a set of χ^2 random variables is also a χ^2 random variable. Those random variables in the set are transformed from a set of uncorrelated random vectors with multivariate normal distribution. Therefore, χ^2 test can still apply on the sum to verify the significance of the sum.

In addition, the geometric meanings of those addends are reserved in the sum. This suggests that only a fixed number of small models are needed, then computation time is reduced. After applying χ^2 tests on the image, multiple statistics from the tests of small features, such as edges and corners, are combined by summation into statistics for larger features such as segments and polygons. These combined statistics signifcantly improve the performance of Hough transform on detecting broken line segments, and avoid making errors. Therefore, summing rule could be one of the powerful operations in to combine the evidence with uncertainty.

References

[1] William James. *Psychology.* Holt, New York, 1890.

[2] Steven W. Zucker. The emerging paradigm of computational vision. *Annual Review of Computer Science*, 2, 1987.

[3] J. Illingworth and J. Kittler. A survey of the hough transform. *Computer Vision, Graphics, and Image Processing*, 44, 1988.

[4] G. Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, Princeton, N.J., 1976.

[5] Judea Pearl. On evidential reasoning in a heirarchy of hypotheses. *Artificial Intelligence*, 28, Feb. 1986.

[6] J. O'Rourke. Dynamically quantized spaces for focusing the hough transform. In *The 7th IJCAI, Vancouver*, pages 737–739, 1981.

[7] C. M. Brown. Hierarchical cache accumulators for sequential mode estimation. Technical report, Dept. of Computer Science, University of Rochester, July 1983.

[8] J. Illingworth and J. Kittler. The adaptive hough transform. *IEEE Trans. Pattern. Anal. Machine Intell.*, 9, 1987.

[9] Victor J. Milenkovic. Multiple resolution search techniques for the Hough transform in high dimensional parameter spaces. In A. Rosenfeld, editor, *Techniques for 3-D Machine Perception*, pages 231–255. Elsevier Science Publishers B. V. (North Holland), 1986.

[10] H. Maitre. Contribution to the prediction of performances of hough transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:669–674, 1986.

[11] Stephen D. Shapiro and Anthony Iannino. Geometric construction for predicting hough transform performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1:310–317, 1979.

[12] D.J. Hunt, L.W. Nolte, and W.H. Ruedger. Performance of the hough transform and its relationship to statistical signal detection theory. *Computer Vision, Graphics, and Image Processing*, 43:221–238, 1988.

[13] Richard Szeliski. *Bayesian Modeling of Uncertainty in Low-level Vision.* Kluwer Academic Publishers, 101 Philip Drive, Assinippi Park, Norwell, MA 02061, USA, 1989.

[14] David B. Sher. A probabilistic approach to low-level vision. Technical Report TR 232, Department of Computer Science, University of Rochester, October 1987.

[15] E-Ren Chuang and David Sher. A χ^2 vision system for feature detection and information combination. In *Machine Vision Architectures, Integration, and Applications.* SPIE - The International Society for Optical Engineering, Nov. 10, 1991.

[16] Robb J. Muirhead. *Aspects of multivariate statistical theory.* John Wiley and Sons, Inc., 1982.

[17] Richard O. Duda and Peter E. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, 1973.