Overview: Computer Vision Performance Characterization

Robert M. Haralick*
Department of Electrical Engineering
University of Washington
Seattle, WA, 98195
haralick@ee.washington.edu

Abstract

Computer vision algorithms are composed of different sub-algorithms often applied in sequence. Determination of the performance of a total computer vision algorithm is possible if the performance of each of the sub-algorithm constituents is given. Performance here does not mean against a user criterion. Performance here is in terms of a complete statistical characterization of the random perturbation on the output of a sub-algorithm as a function of the random perturbation on the input to the subalgorithm.

We have done performance characterization for a variety of edge detector and edge-linking schemes and for a maximum *a posteriori* probability corner detector. And the theoretical results agree with the observed experimental results. In addition we have done some theoretical work with the operations of mathematical morphology and we have done an extensive amount of annotation on the RADIUS Model-Board Images in order to be able to permit comparisons between theory and experiment.

1 IU Performance Characterization

In the ideal world for quickly designing algorithms for Image Understanding, there would be for each algorithm step a performance characterization. This performance characterization would be something like a reference data sheet for an integrated circuit. It would give expressions by which the values for the parameters of the distribution for the random perturbation on the output given the parameters of the distribution for the random perturbation on the input, could be determined for any given description for object and background or clutter.

To then design an algorithm that solves a new task, a data set of representative images would be given. From this training set, the appropriate parameters of the random perturbation processes for the objects of interest and for the clutter would be estimated. These estimates combined with the performance characterization expressions for any algorithm step would then be utilized by the Image Understanding engineer. The engineer would input these into an appropriate computational optimization tool to design an algorithm sequence and determine the values for the associated algorithm tuning parameters to solve the new task as well as be able to estimate the performance of the designed algorithm in the specified application domain.

This is the end goal of performance characterization. The theoretical work of performance characterization involves choosing appropriate stochastic models so that expressions may be derived that express the relationship between the parameters governing the output random perturbation process in terms of the parameters of the input random perturbation process, the algorithm tuning parameters and the distribution of the parameters of the image population in the ideal world.[3]

To do this requires first annotating image data sets for the purpose of validating the stochastic models and estimating the values of the random perturbation processes active on the image. We discuss the annotation we have been doing in the next section. Then we discuss the kinds of feature extraction algorithms we have been exploring to give a sense of how the probabilities related to the output random perturbation process can be derived. We are working not only on operations such as edge detection, edge linking and corner detection, but as well we are understanding how finite random sets propagate through mathematical morphology operations and how to calculate uncertainty in 3D inference algorithms using one or more perspective projection images. Related to this work we have developed expressions for calculating the covariance of any computed quantity that minimizes a given expression with or without constraints. The calculated covariance does not depend on how the minimization is done, only on the expression being minimized and the constraint expressions, if any. And we have worked out experimental protocols for validating a stochastic model.

2 Image Data Annotation

To support the estimation of input random perturbation process parameters, we developed a protocol for use in annotating the Radius model board I, J, and K images. On each image we have appropriately labeled all the building and non-building boundaries. As well we have located the coordinates of all ground control points on all images. Finally, we have located all the building vertices of all buildings on all K images.[7] We are currently working on locating all the building vertices of all the I and M images.

*Funding from DARPA contract 92-F1428000-000 is gratefully acknowledged.

From the labeled annotated images we have been able to determine the gradient distribution of the building boundaries and the non-building boundaries. They are different. Likewise we have been able to measure the variance of the perturbation for areas of building boundaries and areas of non-building boundaries. We have also been able to determine the distribution for the angles between adjacent line segments of buildings and non-buildings. And we have been able to determine the line-length distributions for building lines and non-building lines. Knowing this kind of information then influences the parameters in sucessive feature extraction algorithms.

3 Feature Extraction

We have been exploring feature extraction algorithms for edges,[5] edge-linking, boundary detection, and corner detection.[8] We have been able to analytically propagate the random perturbation of the image input data through edge detection, linking, and boundary extraction as well as corner detection. As a result of this theoretical work, we have defined the basis for a better building boundary detection algorithm than has been previously available.

With regard to edge detection, it is well known that for edge boundaries having gradient magnitude g, the distribution for the estimated squared gradient magnitude $\hat{g}^2$ times $N^2(N-1)(N+1)/\sigma^2$ is a non-central χ^2 with 2-degrees for freedom and with non-centrality parameter g^2/σ^2, when the estimated gradient magnitude is computed over a $N \times N$ neighborhood, N odd. And if the clutter is modeled by independent Gaussian Noise with variance σ^2, then the edge detection false alarm rate p_f is given by

$$p_f = Prob(\chi_2^2 > \frac{T^2 N^2(N-1)(N+1)}{\sigma^2})$$

where T is the gradient threshold.

We have derived the theoretical expression for the distribution of the estimated gradient direction $\hat{\theta}$ as a function of the true gradient magnitude, g, the true gradient direction, θ, the estimated gradient magnitude, $\hat{g}$, the noise variance, σ^2, and the neighborhood size, $N \times N$, N odd. $\hat{\theta}$ is distributed as a Von Mises distribution:

$$P(\hat{\theta}|\theta, \sigma, g, \hat{g}, N) = \frac{1}{2\pi I_0(\kappa)} exp^{(\kappa cos(\hat{\theta}-\theta))}$$

where

$$\kappa = \frac{g\hat{g}}{\sigma^2} \frac{N^2(N-1)(N+1)}{12}$$

These results imply that even for signal to noise ratios for which the gradient magnitude estimates are reasonable close to the true gradient magnitude, the estimated gradient direction can have excessively large variance. Therefore, special care must be taken when employing estimated gradient directions in edge-linking or boundary detection algorithms. It is this fact which has led us to combine the estimated gradient magnitude in narrow elongated directed neighborhoods for determining a more stable estimate of the gradient direction.

We have also examined the hystersis edge-linking performed by the Canny edge detector. We have determined that the probability p_c that an edge is correctly detected can be expressed in terms of the high threshold T_1, the low threshold T_2, the cumulative probability distribution P for the estimated gradient g and the neighborhood size W.

$$p_c = 1-P(g < T_1)+[1-P(g < T_1)^{W-1}]P(T_2 < g < T_1)$$

Having these and related kinds of probabilities, we can determine the length distribution for detected chains of edge pixels that are really boundary chains and the length distribution for detected chains of edge pixels that are due to clutter. This then makes it possible to select an intelligent (optimal) chain length threshold in terms of the prior distribution of true edge chain lengths, the noise variance of the clutter, and the noise variance around the true edges.

We have developed a corner detector which estimates a corner to be that point on the input digital arc whose *a posteriori* probability of being a corner is the maximum among all the points on the arc segment around the point. The detection procedure involves sliding a context window of specified length over the sequence of pixels constituting the arc segment, and doing a line segment fit for each of the two segments starting at the hypothesized corner and ending at one or the other end of the context window. The *a posteriori* probability of a corner is then calculated by using the prior information about corner included angle and adjusting the previously calculated fitting parameters so that the parameter values are found that maximize the *a posteriori* probability of the point being a corner.

We have theoretically determined the false alarm and misdetect rate performance of the corner detector analytically and in and extensive set of experimental trials found these results to be comparable to those measured experimentally.

4 Random Perturbation Propagation

We have explored the propagation of random perturbations through a variety of different kinds of algorithms and in terms of basic models that are applicable to multiple kinds of algorithms. For example, many vision algorithms employ some sort of optimization. The result that is calculated is a value that minimizes a given criterion function. Whether the problem has the form of an unconstrained minimization of constrained minimization, we have derived expressions for the covariance of the computed result as a function of the covariance of the input data. Our result does not depend on knowing any explicit form for how the optimization was done.[1]

In addition we have explored propagating random perturbations through morphological operations of dilation, erosion, opening, and closing. Here, our results are in terms of bounding the probability that a random set is contained in a given set. Here our work is with perturbations that act as additive or act as subtractive noise in binary images.[2]

5 Validation

Since we have an emphasis on estimating the parameters of a random perturbation process, there is the issue of validating the model of the random perturbation process. Here our questions have been relative to how to test that the observed data follows the random perturbation model we are using, whether that data be image input data, or whether that data be some intermediate form that is calculated by a vision algorithm. To do this we have developed experimental protocols using both real and synthetically generated image data.[3]

6 Inverse Perspective Inference

Inverse perspective inference has to do with infering the 3D geometry of objects based on models or partial models and observed image data in one or more images. Because this is a geometry based area, we work with points and lines and conics. Partial object models are specified in terms of relationships existing between points, lines, and conics. At this time we are incorporating all distance, and angle relationships that can exist between points and points, points and lines, points and planes, lines and lines, lines and planes and planes and planes.[4] The general form of the problem for this inference is then a non-linear constrained least squares problem. An example of this kind of a problem is: given the 2D observed perspective projection of vertices of some building, where the faces of the building are assumed to be planer and the angles between the intersecting building faces are given, and given the exterior orientation for all the images in the observation image data set, and given the covariance for all the observations and all the exterior orientation parameters, estimate the 3D location of all building vertices , where the estimated 3D positions satisfy the constraints in terms of which 3D points must lie in the same plane and the intersecting planes having the user specified angles.

7 CD-ROM

Because of the importance and the usefulness of fully annotated data sets, we are collecting all the annotations which we have done and organizing them to be issued on a CD-ROM. In addition to the Radius images and their annotations, we include derived information such as the exterior orientation for all images, the government supplied 3D ground control points, and the additional building vertices that we have located by triangulation. Also derived information about the random perturbation processes will be included. Such information will be relative to noise variances, gradient distributions, line-length distributions, angle distributions, and gap distributions for building and non-building objects.

Formats for this kind of information will be consistent with IUE specifications and for that kind of information permitted by Radius will appear in a Radius format as well.

8 Conclusions

Exploring computer vision algorithms from this point of view leads to insights about the design of better algorithms as well as making possible the optimal settings of algorithm tuning parameters to optimize the expected value of a specific user criterion function at the last algorithm stage. This becomes possible because the expected performance is just the average of the criterion function taken over the distribution that describes the output random perturbation. And it is this distribution that is determined from the theoretical work.[6]

We are also led to a shift in emphasis: we must pay more attention to having good models for the ideal world and good models for the perturbation of the ideal world. Therefore, we must commit ourselves to more data gathering of input images and suitable annotate those images so that estimates can be made of the parameters governing the input random perturbation process.

References

[1] R. Haralick, *Propagating Covariance In Computer Vision*, ICPR-94, Jerusalem, 1994.

[2] R. Haralick, S. Chen, and X. Zhuang, *Finite Random Sets And Morphology*, ICPR-94, Jerusalem, 1994.

[3] R. Haralick, *Performance Characterization Protocol in Computer Vision*, IUW94.

[4] X. Liu, A. Bedekar, and R. Haralick, *Optimzation Methods For Estimating 3D Object Parameters*, IUW94.

[5] V. Ramesh and R. Haralick, *An Integrated Gradient Edge Detector: Theory and Performance Evaluation*, IUW94.

[6] V. Ramesh and R. Haralick, *A Methodology For Automatic Selection of IU algorithm TUning Parameters*, IUW94.

[7] K. Thornton, D. Nadadur, V. Ramesh, X. Liu, X. Zhang, A. Bedekar, R. Haralick, *Groundtruthing The RADIUS Model-Board Imagery*, IUW94.

[8] X. Zhang, R. Haralick. V. Ramesh, A. Bedekar, *A Bayesian Corner Detector: Theory And Performance Evaluation*, IUW94.