Decision Considerations Arising from Error Propagation through Belief Calculations

Jeffrey L. Kretsch and *Edward M. Mikhail*
School of Civil Engineering, Purdue University, West Lafayette, IN 47902

ABSTRACT: In digital automated feature extraction, one attempts to extract symbolic information from digital images by assigning symbolic labels to collections of features obtained from the images. By the nature of the processes used in extracting the information from the images, uncertainty exists concerning the properties of the extracted features; some of the features may result from noise and artifacts of the processing. This leads to symbolic uncertainty, meaning an uncertainty that the proper labeling has been applied. One source of error leading to this symbolic uncertainty is metric positional error in features extracted due to the low-level processing and the mensuration process. Here, we outline an approach using error propagation techniques for exploring the relationship of metric uncertainty to symbolic uncertainty. Although this work concentrates on the Dempster-Shafer formalism for representing evidence, it is felt that propagation concepts might be extended to research on other evidence representations as well.

INTRODUCTION

THE MAIN ISSUE of feature extraction to be dealt with in this article concerns uncertainty representation and measure. The low level processes (Ballard and Brown, 1982) used to extract edges, textures, and other low level attributes are highly error-prone, leading to high uncertainty in feature extraction. There are two types of uncertainty, symbolic and metric. Symbolic uncertainty refers to uncertainty about the proper labeling of an object extracted from the scene. Metric uncertainty refers to the error in measurement of the position of a feature. There are three types of metric error: random error, systematic error, and blunders or gross errors to be discussed later. Metric uncertainty has been dealt with for a long time in photogrammetry, at least for points, but symbolic uncertainty is a relatively new problem in image understanding. In this work, the relationship between the two is being explored.

The word "metric" is used in two different ways in this article. In the first sense, metric uncertainty, metric refers to the act of performing a measurement. In another, we will speak of metric in the topological sense, where metric refers to an expression which measures a distance in some space. These metrics are computed from observations of some quantities and then used as sources of evidence. The evidence from different sources is then combined using some formalism to give an updated evidence which is then used to compute something called a belief value to decide on the symbolic labeling. Metric uncertainty is represented by the variance in observations, and symbolic uncertainty by variance in the belief values (as well as the belief values themselves). Metric uncertainty is related to symbolic uncertainty by error propagation.

The variance in the belief value can be used for the following purposes:

- To understand how metric error affects symbolic matching,
- To compare different metrics used as sources of evidence, and
- To help make decisions after belief values have been calculated.

UNDERSTANDING HOW METRIC ERROR AFFECTS SYMBOLIC MATCHING

In digital automated feature extraction, information is extracted from the image and used in combination with other available information to assign symbolic labelings to groups of features in the image. In doing so, attention has not been paid to the fact that positional or metric error exists in the information extracted from the image. If the information is derived by functional relationships from measured quantities (parameters), then it is possible to propagate error from the measured quantities to belief values and other extracted information.

Two systems which derive evidence needed in performing feature extraction through the use of functional relationships are SCERPO (Spatial Correspondence, Evidential Reasoning, and Perceptual Organization), developed at the Courant Institute of Mathematical Sciences under David G. Lowe (Lowe, 1985), and PSEIKI (Production System Environment for Integrating Knowledge from Images), developed at Purdue University under Avinash C. Kak (Andress and Kak, 1988a).

SCERPO is designed to perform object recognition in the absence of a known orientation of the object. It works on the fundamental premise that certain features are nearly viewpoint invariant, such as parallel lines and collinear lines. In addition, features that are in close proximity to one another will appear in close proximity from most viewpoints if the viewpoint distance from both objects is greater than the separation between them. However, by accidental alignment, features which are distant from one another may appear in close proximity. To help account for this, as well as accidental occurrences of colinear and parallel alignment, SCERPO looks for cases of low probability of accidental occurrences. Its metrics are measures of the probability of an accidental occurrence of the proximity or alignment. The expression "(1-metric)" gives the probability that a significant viewpoint invarient feature has been found. SCERPO ranks the features according to these probabilities, and then begins a search process for object models which have these features. If it finds a match, it then searches the image for features predicted by the object model; if it finds them, further updating of the probability is done by Bayesian Updating (Lowe, 1985).

PSEIKI is designed to take image input and model input and establish a labeling of image features by model features. Several full descriptions are available (Andress and Kak, 1988a; 1988b). Unlike SCERPO, this system assumes that a fairly good sensor exterior orientation exists along with knowledge of the expected scene. After a successful labeling is performed, the exterior orientation can be updated. PSEIKI stores a model of the expected scene represented in a hierarchical framework, starting from vertices and proceeding to build edges, faces, and objects, and then the union of all objects defines the scene. It extracts edges from the image, groups them into expected faces, and then gets evidence for labeling by computing metrics for possible correspondences between image and model edges.

In Figure 1, we illustrate the geometry of a match of image and model features in PSEIKE. The distances D_{perp} and D_{par} are

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 56, No. 6, June 1990, pp. 927–931.

0099-1112/90/5606–927$02.25/0

used in the metrics for collinearity, noncollinearity, and initial labeling.

The first metric makes use of a collinearity function (Andress and Kak, 1988a):

$$\text{collinearity } (E_1, E_2) = \frac{(D_{max} - D_{perp})}{D_{max}} \cos(\theta) \qquad (1)$$

The second metric is a non-collinearity function (Andress and Kak, 1988a):

$$\text{noncollinearity } (E_1, E_2) = \frac{D_{perp}}{D_{max}} \text{Scale } (E_1) \sin(\theta) \qquad (2)$$

where D_{max} is the largest value which D_{perp} or D_{par} can take and E_1 and E_2 refer to Edge 1 and Edge 2, respectively.

For computing initial basic probability assignments (BPA) a modification of the first metric is used:

$$\text{ES\_collinearity} = \frac{D_{max} - D_{perp}}{D_{max}} \times \frac{D_{max} - D_{par}}{D_{max}} \times \cos\theta \qquad (3)$$

The scale (E_1) is a scale factor ranging for 0 to 1, which is a function of the length of E_1, to limit the amount of disconfirmatory evidence due to noise. Both metrics range from 0 to 1. If E_2 is a model edge (i.e., E_a; for models, letter subscripts are used instead of numbers). On the other hand, the noncollinearity function provides evidence that E_1 does not match E_a. These two metrics are computed for each pairing of edges grouped together in the image with each element of the candidate face. The BPAs for each edge are then assigned: Collinearity gives evidence in favor of a match; non-collinearity gives evidence that a match is wrong.

Basically what PSEIKI does is to hypothesize initial groupings, which it tests by first computing initial BPAs, selecting the most likely labeling, and then testing the labeling for corresponding faces and objects. Evidence is propagated up the hierarchy by the simple scheme that if edges are properly labeled, then that is evidence that the face is also properly labeled. Similarly, correctly labeled faces comprise evidence that an object is correctly labeled. The process is fully described in Andress and Kak (1989b). For the representation of evidence, PSEIKI uses the Dempster-Shafer formalism (Shafer, 1976).

Uncertainty often arises in cases where multiple sources of evidence, sometimes conflicting, exist. In these cases, a method is needed for combining these different pieces of evidence to, in the end, provide some updated belief value. One such mechanism is the Dempster-Shafer Theory of Evidence. This does not concern itself with the origin of the evidence, but only that the evidence be expressed in a way that it has certain properties; for example, it is expressed as a number between 0 (no support) and 1 (complete support). This formalism is distinguished by the fact that it is not required that all belief be committed to either a proposition or its negation, as for example, Bayesian theory requires. The uncommitted portion of belief represents our ignorance. In some cases in which we feel we should not have ignorance due to the nature of evidence being collected, we might associate this uncommitted portion to doubt that we have been considering the proper set of possibilities.

SCERPO is not concerned about ignorance because "There seems to be no need for an estimate of ignorance when calculating rankings for a search process" (Lowe, 1985). This is one reason Bayesian statistics were used.

The goal of the Dempster-Shafer theory is to provide a means to represent and pool different sources of evidence. Generally, the final belief values assigned to different possibilities as a result of using Dempster's Rule of Combination are used in

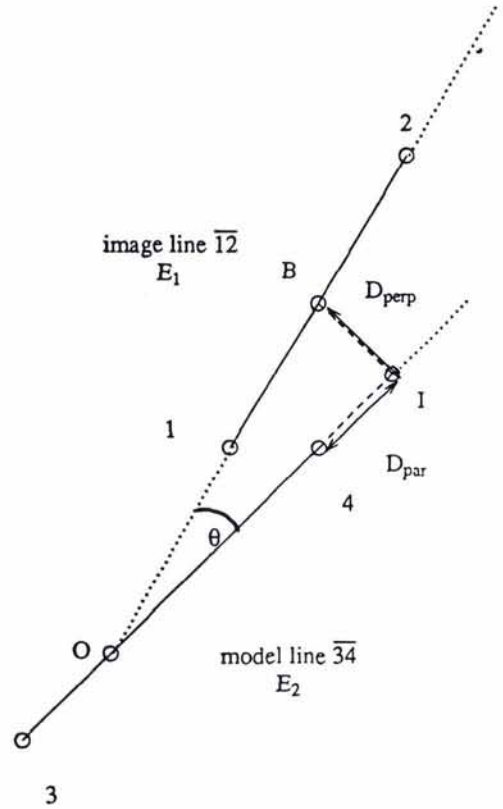

FIG. 1. Geometry of an example match in PSEIKI.

making some decision on which possibility to accept. In the simplest case, one might simply accept the possibility for which the belief value is highest. This would, in effect, assume that the evidence sources were reliable and exact. Suppose, however, that, in addition to the evidence sources, there also exists a value for the variance of each piece of evidence. This may occur if the evidence is a value computed from observed quantities. We will look at how the variance can help in making a decision.

Up to this point we have assumed that there is no error in our extracted parameters [θ, D_{perp}, D_{par}]. In reality, of course, there must be error in the parameters. Let us assume we have errors in the vertices of our extracted edges, and further, the model is also imperfectly known, so we have errors there as well.

We propagate the error from the vertices to the parameters and then from the parameters to the metrics:

$$\Sigma_{xx} = \mathbf{J}_{xl} \Sigma_{ll} \mathbf{J}_{xl}^T \qquad (4)$$

$$\Sigma_{mm} = \mathbf{J}_{mx} \Sigma_{xx} \mathbf{J}_{mx}^T \qquad (5)$$

where $\mathbf{J}_{xl}$ and $\mathbf{J}_{mx}$ are Jacobians, Σ_{ll} is the covariance matrix of the parameters, and Σ_{mm} is the covariance matrix of the metric.

For establishing the initial labelings, PSEIKI calculates ES_collinearity for the whole set of possibilities, and then assigns that label which has the highest belief value. The accompanying error propagation would provide a variance for that value as well.

One can continue the error propagation through higher levels by propagating through Dempster's Rule, as is shown in Kretsch and Mikhail (1989). Having determined a variance for the belief values, the next sections deal with how the variance is used.

COMPARISON OF DIFFERENT METRICS

The central issue of interest is the behavior of a computer vision system with respect to metric error. This error can be divided into three components: random error, systematic error, and blunders.

Random error is that unpredictable, unavoidable component. It is this error which we propagate as variance. Systematic error is predictable error due to shortcomings in the mathematical model. In the case of PSEIKI, it is really systematic error that "drives" the process, because it is trying to find a specific systematic error in position which will be used to update the position of a robot. It is important, however, that this systematic error in position not be so severe as to prevent the proper labeling needed to determine the robot's position. Also, there are other sources of systematic error which may be misinterpreted and lead to error in position. Finally, there are blunders, and in low-level processing there are certainly a lot of those. There appear to be some very interesting aspects of blunders which may occur in PSEIKI which contrast to blunders encountered in other fields of photogrammetry. PSEIKI itself provides some means of dealing with these blunders (Kretsch and Mikhail, 1989).

Discussion of propagation of variance in belief calculations would be of limited value if the belief functions were subjective in nature. However, in the case of PSEIKI, the belief functions are calculated based on observations. Thus, a variance in the observations will result in a variance in the belief values. This can be useful in providing a means of relating positional metric uncertainty to symbolic uncertainty.

The first step in the PSEIKI labeling process is the assignment of the initial labelings, using the ES_collinearity metric: The first step leads to the first decision made by PSEIKI, and perhaps the most important, deciding what the initial labelings should be. The immediate question arises of how one is to represent metric uncertainty in the image. Another question is where one may obtain these metric uncertainties.

At this point we do not have known variances. We have decided to assume different covariance matrices and then simulate their respective effects on the belief calculations. Two possibilities for the metric uncertainty representation were considered. In the first possibility, we might use a parameter covariance matrix for θ, D_{perp}, and D_{par}. In this case we would assume that we are directly measuring θ, D_{perp}, and D_{par}. With this method it seems hard to get off-diagonal terms of the covariance matrix. In the second possibility, we assume that the vertices are the observations. This is consistent with the hierarchical nature of PSEIKI, in which the lowest level of the hierarchy is identification of the vertices. We assign covariance matrices to the vertices, and then proceed to do error propagation from these to the parameters to get a covariance matrix for θ, D_{perp}, and D_{par}.

There is a philosophical problem with this second representation. The extracted segments presumably come from or correspond to a segment on some model edge. But we have no way of knowing to what point on the model edge our image vertex corresponds. All we know about the expected value of the position of our image vertex is that it should be on the model edge if the match is correct, and if there were no systematic error. Given this state of affairs, what does a variance in position of a vertex mean? We are assuming here that the extraction of the vertex with its corresponding position is an observation to which we are assigning a variance. This variance is related to the image quality which we define loosely as being a measure of the resolution and sharpness of the image.

We decided to use simulation to deal with the research issues. Basically most of the simulation involved introducing a Σ_{ll} matrix constructed from assumed σ_x^2 and σ_y^2 for each vertex and a correlation coefficient ρ between the vertex x and y coordinates, and then propagating this covariance matrix through the belief calculations to get a $\sigma_{ES\text{-}com}^2$ (variance of ES_collinearity). We worked in an arbitrary unit system, in which we picked variances in the range from 0 to 0.10, in a roughly 3 by 3 unit area. If the arbitrary unit is inches, for example, this approximates an image frame. In addition to variances, other parameters such as the length of the image segment, the tuning parameter D_{max}, and ρ were varied to study their relationship to $\sigma_{ES\_COM}^2$ and ES_collinearity. The studies performed are fully described in Kretsch and Mikhail (1989).

In addition to the variance and the expected value of a metric, another important statistical property is the distribution. We assume here that the vertices have normal distributions. Finding the actual distributions of the vertex points would involve a deep study of the edge detection, skeletonization, and data reduction's effect on the stochastic properties of the vertices.

Two methods for finding the distribution of the metric were considered. In the first, we attempt to propagate the distributions of the vertices to find the distribution of the metric. The method is described in Anderson (1958) and Mikhail (1976). In this method, we first establish the functional relationships and the inverse relationships. In this case, it turns out that the inverse relationships do not exist.

So we decided to use the second technique, Monte Carlo simulation, to find the distribution of the metric. The specific metric we chose to look at is ES_collinearity. In this case, we took as our starting point the example in Figure 1.

We generated a normally distributed function with the same mean and variance as the ES_collinearity example in Figure 1 and plotted it with a dashed line along with the distribution of ES_collinearity which we computed by Monte Carlo methods (Kretsch and Mikhail, 1989) and show this in Figure 2. In this figure, the difference between the two is very noticeable. The distribution of ES_collinearity has less scatter than a normal distribution, i.e., its variance is reduced, resulting in a narrower and higher distribution, as would be expected.

Of course, one of the most important issues we wished to address was the relationship of metric uncertainty to symbolic uncertainty. We decided to explore this through simulation for several reasons. First, we have no knowledge of what the true variances of the image edge positions are. Because at least some of this error derives from low level image processing, a very detailed study would be required to find out. Second, through simulation we have control over the error, being sure that only the errors which we have introduced enter into our result.

We concluded that the belief metric chosen responds linearly

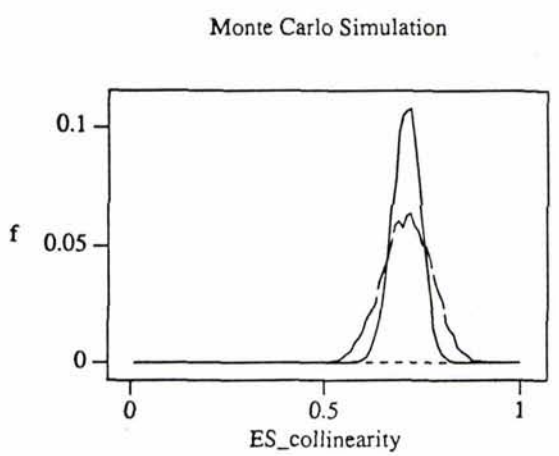

FIG. 2. Distribution of ES_collinearity along with normal distribution.

to changes in the image and model variances. Indeed, analysis of the mathematical derivations shows that the metric chosen only affects the J_{mx} terms, which help determine the slope of the line. So any metric will have a linear relationship, the difference between metrics will lie in the slopes of their respective σ^2_{metric} to positional σ^2 relationships.

To understand the variance of extracted feature parameters, we must look very closely at the digital processes used to extract those parameters. The processes to be examined are the preprocessing phase where pixel to symbol conversion takes place and the data reduction module where deletion of some vertices take place.

MAKING DECISIONS AFTER BELIEF VALUES HAVE BEEN CALCULATED

If one choice clearly stands out from the rest, then the variance of the possibility selected should make no difference. If there is one outstanding candidate for selection, or none, then the variance may mean nothing at all. The case of interest is when two belief values are nearly the same, as Dungeon (1988) has pointed out.

Suppose we have two possibilities, A and B, and Bel(B) > Bel(A) where Bel(B) is the belief value for B and Bel(A) is the belief value for A. Normally, under these circumstances, B would be selected. Different selections might be made if other selection criteria were used.

For example, suppose it was desired to select that possibility which was most likely to exceed a certain threshold T. Although B exceeds A, that is, B probably exceeds A, the probability of A exceeding T can be greater than the probability of B exceeding T.

This is an example of where knowledge of the variance of a random variable, in this case, the belief values Bel(A) and Bel(B), provide additional information which can impact our decision. To actually perform the calculation, the distribution functions must be known. If we assume that the distribution is normal, the mean and variance completely define it. Because we generally assume that our observations are normally distributed, the assumption is not always unreasonable. However, as the next example will show, for some decision criteria, the distribution does not matter. Where distribution does matter, we must propagate distributions through the metrics and uncertainty calculations.

Suppose we normally select that possibility with the highest belief value. Could the variance and distribution information alter this initial decision strategy? Let us first treat our calculated beliefs as we would an observation, by assuming in the absence of any other information that our calculation represents the expected value of the belief, and that these belief values are independent.

$$E(\mathrm{Bel}(X)) = \mathrm{Bel}\ (X)\ \text{(calculated)} \tag{6}$$

and for our case

$$E(\mathrm{Bel}(B)) > E(\mathrm{Bel}(A)) \tag{7}$$

Given this, it can be shown (Kretsch and Mikhail, 1989) that

$$P[\mathrm{Bel}(B) > \mathrm{Bel}(A)] > .5 \tag{8}$$

Thus, if one desires the choice with the probability of having the highest belief value, B is the proper choice, given the assumptions above, regardless of the distribution.

INTERPRETATION OF RESULTS

What does the variance in the belief values tell us about our symbolic matching? In the case of SCERPO, the main effect would be on the search strategy. SCERPO would seem to be rather robust with respect to metric error for two reasons; first, from the results above, the ranking is not affected by distribution and, hence, variance; and second, even if an error in ranking is made, it is likely that this will decrease the efficiency of the search, but not the outcome unless the error leads to settling on a wrong but not inconsistent solution.

For PSEIKI, we might start to answer the question by looking at the effect of variance in the vertex positions. An error in position can only make a correct match appear worse. But an error in position could increase the belief value assigned an incorrect match.

A more important result that may be derived from the variance in the belief value is the threshold that should be used in determining if a label is correct or not. Both SCERPO and PSEIKI use a threshold at some point in their evidence accumulation process. The lower the variance, the higher the threshold should be. For zero variance, meaning perfect metric feature extraction —i.e., no errors in positions of vertices between the image and the model in world space—the threshold should be 1.

At this point we have developed the necessary error propagation theory and using it have simulated error propagation through the assignment of initial BPAs in PSEIKI. The results of these simulations have given us some idea of how the metrics are affected by error in the imagery. This helps meet one of our first goals, that is, determining the behavior of the metrics.

Perhaps the most important contribution of this work is that it forced a re-evaluation of what considerations must go into choosing a metric. We feel that considering the stochastic property of the metrics using error propagation techniques is a definite contribution in both choosing metrics and understanding their performance.

CONCLUDING REMARKS

In this work we have explored the effect of metric error on symbolic error. We have demonstrated that this can be done by implementing an example on the first step of PSEIKI.

Further work can be done at both ends. On the one hand, a study is needed to identify what the true metric error is. This study might be done in two ways or a combination of both: first, to carefully study and model the low level image processing techniques to see what error they introduce; and second, attempt to determine error (accuracies and precision) by empirical studies. In the combination of the methods, we might verify our theoretical studies of low-level processing by empirical studies. By empirical studies, we mean measuring known control features under known circumstances to determine what error has been introduced by the system.

One of the possible future enhancements of PSEIKI is incorporating the ability to change model panels, and use this ability for automatic target recognition (Andress and Kak, 1988a). It is this potential capability which makes PSEIKI interesting for mapping purposes. The ability of PSEIKI to establish a labeling of corresponding feature elements in imagery and then provide an overall belief value for the match could be used to automate some of the processes of extracting those cultural features of interest that can be represented by models. Examples of such features may be buildings and power transmission towers. Much work remains to be done in this area.

This approach of error propagation may also be applied to other representations, such as the maximum entropy method. Different interpretations from those used in Dempster-Shafer analysis may have to be developed, but the issue of relationship between metric uncertainty to symbolic uncertainty will remain.

At any rate, error propagation has made available more information which can be used in the feature extraction process. This information combines the development of photogrammetric data reduction developed over many years with the emerging image understanding field.

ACKNOWLEDGMENTS

The authors wish to thank Professor Avi Kak, Professor Jeff Wright, Professor Howard Turner, and Dr. Keith Andress for their support in this work. The first author also wishes to thank the Defense Mapping Agency for allowing a leave of absence while this work was carried out, and the American Society for Photogrammetry and Remote Sensing and Wild-Heerbrugg Corporation for their support through the Wild-Heerbrugg Photogrammetric Award.

REFERENCES

Anderson, T. W., 1958. *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, New York, N.Y.

Andress, K. M., and A. C. Kak, 1988a. *The PSEIKE Report - Version 2*, Technical Report, TR-EE-88-9, School of Engineering, Purdue University.

———, 1988b. *Evidence Accumulation & Flow of Control in a Hierarchical Reasoning System, AI Magazine*, Vol 9, No. 2, pp. xxx–xxx.

Ballard, D. H., and C. M. Brown, 1982. *Computer Vision*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Dudgeon, D. E., 1988. Extraction and Recognition of Silhouettes from Range Imagery, Special Seminar, Purdue University, 26 February 1988.

Kretsch, J. L., and E. M. Mikhail, January 1989. *Applications of Artificial Intelligence to Digital Photogrammetry*, Purdue Technical Report, CE-PH-89-1.

Lowe, D. G., 1985. *Perceptual Organization and Visual Recognition*, Kluver Academic Publishers, Boston, Mass.

Mikhail, E. M., 1976. *Observation and Least Squares*, Harper & Rowe, Publishers, Inc., New York, N.Y.

Shafer, G., 1976. *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, New Jersey.

(Accepted 18 January 1990)