Proceedings of the 1991 IEEE
International Conference on Robotics and Automation
Sacramento, California - April 1991

Set Membership Approach to the Propagation of Uncertain Geometric Information *

Assumpta Sabater

Matemàtica Aplicada III (UPC)
Colom 1,
08222 Terrassa, Spain

Federico Thomas

Institut de Cibernètica (UPC – CSIC)
Diagonal 647, 2 planta,
08028 Barcelona, Spain

Abstract

The fusion of geometric information is of great significance in multisensorial systems, mainly in robotics applications, where multiple sensors or mobile sensor systems that change their perspective of the environment capture sparse, and sometimes partial, geometric data. These data contain some level of uncertainty and, in general, some level of redundancy. Probabilistic approaches have been used to solve the problem of fusing this information to obtain the best estimate of a given set of parameters describing a collection of geometric features and its final associated uncertainty. Nevertheless, a probabilistic description of errors is not always available and only a bound on them is known. The set-membership approach postulates that a measurement only allows us to establish an uncertainty region in the space of parameters describing a geometric feature. This approach avoids the general assumptions of unbiased and independent measurements taken by the probabilistic approaches.

1 Introduction

Low-level sensing processes extract geometric features such as line segments or surface patches from the sensor data, while high-level sensing processes use symbolic models, geometric templates, and prior heuristic assumptions. While probabilistic approaches have been usually applied for solving the problems arising in the former processes, Artificial Intelligence techniques have been applied for the latter. In this paper, we will concern ourselves to the former kind of processes. See [8] and [6] for recent surveys on sensor fusion.

Given some uncertain information about geometric features in the environment, the problem tacked here is to obtain an estimate for each of these features which satisfies a set of constraints derived from stored object models to which the features belong to, or from the acquired information itself. For example, let us assume a system that extracts straight line segments from images, with some uncertainty in their position, and this information

must be fused to obtain a consistent interpretation of the scene. For this example, the geometric features would be the straight line segments and their ending points, and the constraints would be derived from local spatial relationships – such as alignment, perpendicularity or parallelism – between segments leading to equations relating the parameters that represent the features. An estimate of the actual position of the features must satisfy all the introduced constraints and must be *close* to the original measurement.

The probabilistic approach has dominated much of the work on low-level sensing processes handling uncertainty. Although probabilistic models which assume an uniform distribution inside a range [5] have been used, the computational tractability of low-level sensing processes, under this approach, requires the general assumption that the experimental error is simply an additive term with Gaussian distribution, and the data combination formalism is essentially that of maximum likelihood estimation.

Gaussianity is usually assumed because: (1) the mean and the covariance matrix are sufficient information to completely define the feature density function; and (2) the distributions are easily manipulated only through matrix computations, allowing the development, in general, of fast and efficient algorithms for the manipulation of uncertain geometry. Nevertheless, this approach exhibits some drawbacks. For example, when the flow of incoming information and the number of constraints between the geometric features to estimate are large, there is a heavy computational overhead in the maintenance and updating of large covariance matrices.

Actually, it is difficult to give a complete error analysis because the complexity of the process of extracting low level data. Instead, errors in measurements are assumed to be bounded, so that every measurement leads to an uncertainty set in the space of parameters where the actual value is bound to be.

The essence of the set membership approach, as it has been pointed out in [3], is that certain a priori information is known about the system that helps pare down the space of parameters that need to be considered as estimations.

The underlying set membership principles have been applied in the control and systems science domains. Two main sorts of sets have been used: polytopes [10] and ellipsoids [3]. Our approach makes use of both kinds of sets. We will concentrate ourselves on the propagation of uncertainty over graphs of ge-

ometric constraints. In these graphs, nodes stand for geometric primitives whose uncertainty regions are represented by ellipsoids in their parameter spaces. When a new measurement is acquired, a new uncertainty set is introduced. This set is assumed to be a convex polytope defined by a set of strips, or pairs of parallel hyperplanes. These strips are the elements propagated all over the graph of geometric constraints and fused at each node. An updated ellipsoid is thus obtained as the final uncertainty region for each feature, as well as an estimation of its actual location.

2 Probabilities and uncertainty sets

We defined an *observation* as a given spatial configuration of a set of sensors and their surrounding environment. We will denote $x^i \in \Re^n$ the vector of parameters, or *measurement*, representing a geometric feature obtained in observation i, n being its dimension.

The parametrization chosen for each feature are very important for two reasons:

- The volume of the uncertainty region associated with an estimation may change if the reference frame of the geometric feature is changed. We should find a set of parameters, so that the volume of the uncertainty region is the same over different coordinate systems.

- As it will be seen later, we are going to deal with constraints between features which can be translated into equations linking their representing parameters. A proper choice of the parameters can lead, in some cases, to linear constraints.

It is not always possible to find parametrizations satisfying these constraints. This problem is outside the scope of this paper.

All measurements are assumed to be contaminated by noise, that is

$$x^i = x_r^i + n^i,$$

x_r^i being the actual value of x and n^i some aditive perturbation. Under the probabilistic approach, n^i is assumed to be, in general, an unbiased ramdom Gaussian variable with covariance matrix A_i. Moreover, for different measurements, these perturbations are also assumed to be uncorrelated, that is

$$E\{n^i \cdot (n^j)^T\} = 0, \quad i \neq j,$$

where $E\{\cdot\}$ denotes the expectation operator. Note that, for Gaussian distributions, this implies independency between measurements. The independency assumption is the hardest assumption taken, which leads to very optimistic final uncertainties, as it will be shown in an example below.

The set of points in $\Re^n$ defined by

$$(x - x^i)^T A_i^{-1}(x - x^i) \leq k$$

is an ellipsoid for any positive k, which is called the *uncertainty ellipsoid* associated with x^i. Depending on the value of k, this leads to a region where the actual value of x is most likely to be. In the set membership approach, there is also an uncertainty region associated with each measurement x^i, where the actual value, x_r, is assumed to be with absolute certainty. This provides a connection between both approaches upon which to compare them.

Often we only have partial information on feature parameters (infinite uncertainty in one or more degrees of freedom), whence the distribution function is improper and the covariance matrix is not strictly defined. In the set membership approach partial information is represented through degenerated uncertainty regions, that is unbounded sets which extend till infinity in some directions.

One of the basic tools of a multisensor system is the fusion of different measurements, taken at different observations, of a unique feature to obtain a consensus. Obviously, this consensus should have less uncertainty than the measurements themselves.

Let us assume that we have sensory information obtained from p different observations about the same geometric feature. The problem consists in fusing x^i, $i = 1 \ldots p$, to obtain a consensus $\hat{x}$ which *best* estimates x_r. Several criteria have been used to obtain this estimate, but all of them assume statistical independency between measurements (see the appendix for a brief survey). The most widely used formula for obtaining this consensus is:

$$\hat{x} = \sum_{i=1}^{p} \{\sum_{j=1}^{p} (A_j)^{-1}\}^{-1} (A_i)^{-1} x^i, \tag{1}$$

whose covariance matrix is:

$$\tilde{A} = \{\sum_{i=1}^{p} (A_i)^{-1}\}^{-1} \tag{2}$$

While using the same optimum (1), other covariance matrices derived from linear interpolation have been used instead of (2) (see [7]):

$$\tilde{A} = \frac{1}{p-1} \cdot \sum_{i=1}^{p} \frac{\sum_{j=1, j \neq i}^{p} |\hat{x} - x^j| A_i}{\sum_{k=1}^{p} |\hat{x} - x^k|} \tag{3}$$

This tends to reduce the optimistic results obtained through the application of (2) due to the independency assumption taken. As an example, *fig. 1* depicts two measurements and their uncertainty ellipsoids. The estimation obtained and the associated uncertainty ellipsoid using (2) – *fig. 1a* – and (3) – *fig. 1b* – are shown in dotted lines.

The combination law derived from (2) is commutative and associative, which are desirable properties for any combination law of competitive information. Associativity leads to modularity in the combination when the incoming information from $p + 1$ sources must be fused. In other words, the fusion of $\hat{x}_p$ and x^{p+1} is the same as the one obtained from the fusion of $x^1, \ldots, x^p$, and x^{p+1}. Thus, the information provided by sensors with fast response can be fused before those from sensors with slower response.

In general, there are some geometric relationships known a priori between the geometric features. For example, the vertex of a trihedron is on its adjacent faces and edges, and the edges on their adjacent faces. These are geometric relationships that are independent of the position of the trihedron. Let us assume that some information has been obtained about the edges and
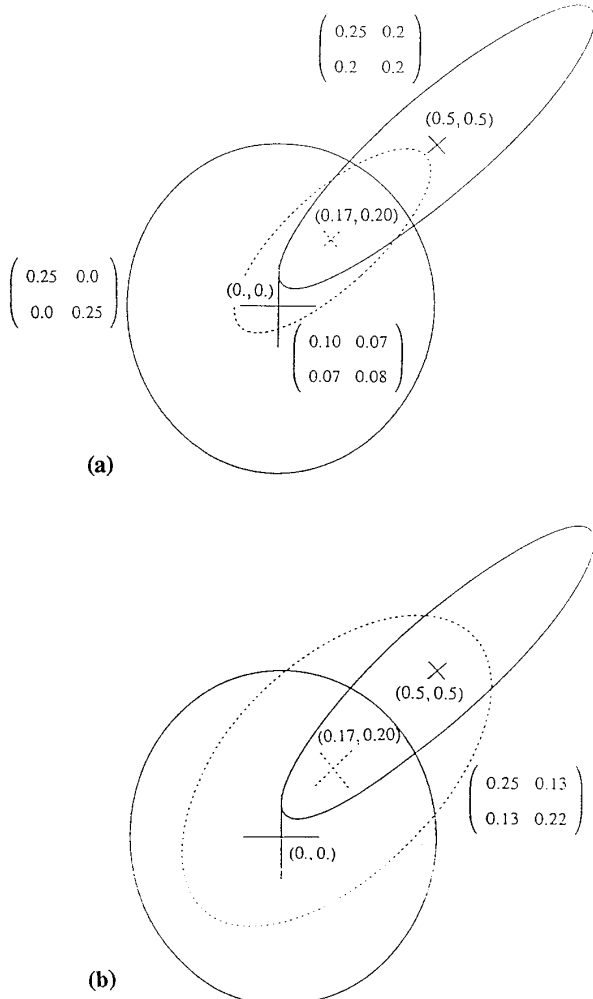
$$\begin{pmatrix} 0.25 & 0.2 \\ 0.2 & 0.2 \end{pmatrix}$$

(0.5,0.5)

(0.17,0.20)

$$\begin{pmatrix} 0.25 & 0.0 \\ 0.0 & 0.25 \end{pmatrix}$$

(0.,0.)

$$\begin{pmatrix} 0.10 & 0.07 \\ 0.07 & 0.08 \end{pmatrix}$$

(a)

(0.5,0.5)

(0.17,0.20)

$$\begin{pmatrix} 0.25 & 0.13 \\ 0.13 & 0.22 \end{pmatrix}$$

(0.,0.)

(b)

Figure 1

the vertex they intersect. This leads to a graph of geometric constraints with three nodes, which can be parametrized in 3D using three or four parameters each, and three edges. Obtaining the best estimate of the edges involves dealing with covariance matrices of 15×15.

Features and geometric constraints define a graph, called *graph of geometric constraints*, whose nodes stand for geometric features – stored as parameter vectors – and arcs, for constraints – stored as vectorial functions.

Geometric constraints allow us to propagate information, since information about a feature gives information about other features in relation with it through constraints.

A geometric constraint between two features with parameters $x \in R^m$ and $y \in R^n$ may be given as a vectorial expression $h_j(x, y) = 0$. In general, these functions do not define x or y in a deterministic sense. This is because the two related features may

have different degrees of freedom as described by their different dimension. If functions h_j are considered in a stochastic sense, then this lack of constraint can be represented as partial information [4]. When propagating partial information, the uncertainty regions become degenerate, that is, there is infinite uncertainty or no information in one or more degrees of freedom.

Any constraint $h_j(x, y) = 0$ can be linealized for small errors, so that

$$\frac{\partial h_j}{\partial x}(x_0, y_0)(x - x^i) + \frac{\partial h_j}{\partial y}(x_0, y_0)(y - y^i) = 0.$$

where x_0 and y_0 are previous estimates (see [12] for examples of geometric constraints and their linearizations). Using a Gaussian model, this equation leads to

$$\frac{\partial h_j}{\partial x} \Lambda_x \left(\frac{\partial h_j}{\partial x} \right)^T = \frac{\partial h_j}{\partial y} \Lambda_y \left(\frac{\partial h_j}{\partial y} \right)^T$$

where Λ_x and Λ_y are the covariance matrices of x and y, respectively, and $\frac{\partial h_j}{\partial x}$ and $\frac{\partial h_j}{\partial y}$ are the Jacobian matrices obtained in (x_0, y_0).

If a set membership approach is used and the uncertainty regions for both geometric features are

$$(x - x^i)^T E_x (x - x^i) \le 1 \quad \text{and} \quad (y - y^i)^T E_y (y - y^i) \le 1,$$

then

$$\frac{\partial h_j}{\partial x} E_x^{-1} \left(\frac{\partial h_j}{\partial x} \right)^T = \frac{\partial h_j}{\partial y} E_y^{-1} \left(\frac{\partial h_j}{\partial y} \right)^T$$

holds.

This matrix equation permits the information propagation from feature x to feature y. Actually, if $x^i \in R^m$ and $E_x \in R^{m \times m}$ are known and we have a previous estimation of y, $E_y \in R^{n \times n}$ can be obtained from:

$$E_y = \left(\frac{\partial h_j}{\partial y} \right)^T \left(\frac{\partial h_j}{\partial x} E_x^{-1} \left(\frac{\partial h_j}{\partial x} \right)^T \right)^{-1} \frac{\partial h_j}{\partial y}.$$

See [11] for a proof of this last step.

The information about y given by these equations is complete if $\text{rank} \left(\frac{\partial h_j}{\partial y} \right) = n$, or partial if $\text{rank} \left(\frac{\partial h_j}{\partial y} \right) < n$. If the information is complete E_y describes a non degenerate ellipsoid, otherwise an ellipsoidal cylinder.

If $h(x, y) = 0$ does not allows us to completely determine y^i, we can take as y^i the point y in the manifold $h_j(x^i, y) = 0$ with shortest Euclidean distance to the previous estimation of y.

3 Finding a tight uncertainity bound

Given several measurements and their associated uncertainty regions, the best uncertainty bound that we can obtain for the estimation is the intersection of these regions.

Let $x_1 \ldots x_n$ be a set measurements and $\mathcal{E}_1 \ldots \mathcal{E}_n$ their associated uncertainty sets. These sets are assumed to be possibly degenerate hyperellipsoids, that is:

$$\mathcal{E}_i : (x - x^i)^T E_i (x - x^i) \le 1 \tag{4}$$

Notice that $\mathcal{E}_i$ can also be a strip as defined above. Then, any estimation of x, $\hat{x}$, must simultaneously belong to each of these sets, that is

$$\hat{x} \in \hat{\mathcal{E}} = \bigcap_{i=1}^{n} \mathcal{E}_i.$$

We could intersect all sets in (4) in parameter space, and in this case end up with a minimal membership set. Nevertheless, even in the case the initial uncertainty sets are strips, so that the resulting region is a polytope in $\Re^m$, high computational complexity algorithms have to be implemented [10].

The accumulated inequality derived from (4) holds and defines the following set:

$$\hat{\mathcal{E}} : \sum_{i=1}^{n} \lambda_i (x - x^i)^T E_i (x - x^i) \leq \sum_{i=1}^{n} \lambda_i$$

for any non-negative set of weights λ_i. Since $\hat{\mathcal{E}}$ is the smallest possible set, it must be true that $\hat{\mathcal{E}} \subseteq \hat{\mathcal{E}}$. Actually, the exact computation of the ellipsoid with minimum volume leads to a problem that can only be solved by using numerical methods.

Then, let us assume that we are only to find the uncertainty region obtained from the fusion of two measurements whose uncertainty regions are an ellipsoid $\mathcal{E}_1$ and a strip $\mathcal{S}$ (fig. 2a). To this end, let $\mathcal{E}_1$ be an ellipsoid, that is

$$\mathcal{E}_1 : (x - x^1)^T E_1 (x - x^1) \leq 1, \qquad (5)$$

where $x^1 \in \Re^n$ is its center and E_1 is a $n \times n$ symmetric positive defined matrix, and let $\mathcal{S}$ be a strip, that is

$$\mathcal{S} : |A^T x - b| \leq 1, \qquad (6)$$

where $b \in \Re$ and $A \in \Re^n$ is a normal vector to the hyperplanes.

All the ellipsoids containing the intersection $\mathcal{E}_1 \cap \mathcal{S}$ satisfy

$$(x - x^1)^T E_1 (x - x^1) + \lambda |A^T x - b| \leq 1 + \lambda,$$

where λ is a positive weight. This inequality can also be expressed as

$$\hat{\mathcal{E}} : (x - \hat{x})^T \hat{E} (x - \hat{x}) \leq 1$$

where

$$\left.\begin{array}{rcl} E_0 & = & E_1 + \lambda AA^T \\ \hat{E} & = & E_0 / K \\ \hat{x} & = & E_0^{-1}(E_1 x^1 + \lambda bA) \\ K & = & 1 + \lambda - (x^{1T} E_1 x^1 + \lambda b^2) + \hat{x}^T E_0^{-1} \hat{x} \end{array}\right\}$$

$\hat{x}$ being the center of $\hat{\mathcal{E}}$.

It is better to compute directly the matrix E_0^{-1} and $\hat{E}^{-1}$ from E_1^{-1}, instead of E_0 and $\hat{E}$, since only the inverse matrices are needed in latter fusions. Thus, it can be shown [3] that:

$$E_0^{-1} = E_1^{-1} - \lambda \frac{E_1^{-1} AA^T E_1^{-1}}{1 + \lambda G}$$

where

$$\left.\begin{array}{rcl} K & = & 1 + \lambda - \frac{\lambda \epsilon^2}{1 + \lambda G} \\ \hat{x} & = & x^1 + \lambda E_0^{-1} \epsilon A \\ G & = & A^T E_1^{-1} A \\ \epsilon & = & b - A^T x_1 \end{array}\right\}$$

The ellipsoid which smallest volume, including the intersection of the ellipsoid and the strip, is the one that minimizes $1/\det \hat{E}$. Thus, we should find, by derivation, the corresponding weight λ. It has been proved elsewhere [3] that this value for λ is the most positive root of the equation:

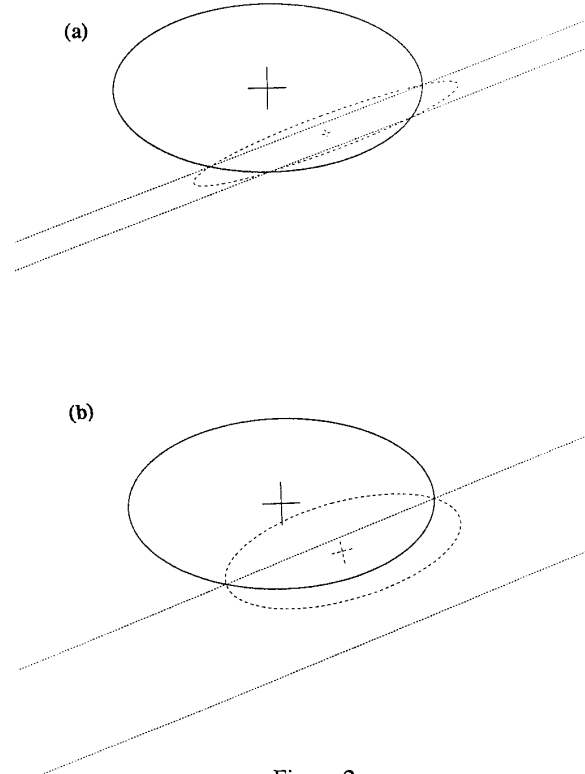$$(n - 1)G^2 \lambda^2 + (2n - 1 + \epsilon^2 - G)G\lambda + n(1 - \epsilon^2) - G = 0, \quad (7)$$

(a)

(b)

Figure 2

The thus obtained ellipsoid, which will be denoted by $(\mathcal{E}_1 \cap \mathcal{S})$, contains the set defined by (5) and (6) in C^n. If this set is embedded in $\Re^n$, the obtained ellipsoid is the tightest one bounding the intersection, as shown in fig. 2a. Otherwise, the ellipsoid obtained is not the tightest possible (fig. 2b). Nevertheless, this is not an important drawback since it can be proved that $(\mathcal{E}_1 \cap \mathcal{S}) \subset (\mathcal{E}_1 \cup \mathcal{S})$. In other words, no new elements of uncertainty are ever introduced.

Two important situations arise: (a) if K is negative, there is no intersection between the ellipsoid and the strip. This provides an easy way of detecting inconsistencies; and (b) if there is no positive value for λ satisfying equation (7), the intersection of the ellipsoid and the strip is the ellipsoid itself, that is, the measurement acquired does not reduce the uncertainty in the system.

4 Updating a graph of constraints

A process for updating a graph of geometric constraints would consists in obtaining a set of estimations that satisfy all the constraints inside the uncertainty regions associated with them. Before obtaining an estimation, the resulting uncertainty regions will be found.

Each time a sensory data is acquired, the corresponding measurement is assumed to be inside a set of strips. These strips must be propagated and fused with the uncertainty ellipsoids to obtain the updated ellipsoids of uncertainty. Let us assume that only a new uncertainty strip for feature q is introduced and, as a result of this, we want to update the uncertainty ellipsoids for all the features in the graph. This can be carry out by propagating this strip all over the spanning trees rooted at q. This suggests the important role of cycles in graphs of geometric constraints.

Since no previous assumption about the probability distributions inside the uncertainty regions has been taken, any point inside these regions would be a good estimation for the corresponding feature. Thus, any set of values inside the regions satisfying the constraints would be a good estimation. Nevertheless, the center of the uncertainty regions are considerer as the best estimation [3]. In our case, since these values must satisfy all the constraints, a good estimation would be that which minimize the sum of the Mahalanobis distances to the centers of the uncertainty regions, that is, the estimations $\hat{x}_i$ would be the values for x_i that minimize

$$\sum_{i=1}^{p} (x_i - c_i)^T E_i (x_i - c_i)$$

c_i being the centers of the regions defined by $\mathcal{E}_i$, subject to the constraints h(.). Thus, in general, the uncertainty regions will not be centered around the estimations.

5 Conclusions

An alternative approach for the propagation of uncertain geometric information, based on the ideas presented in [3] and extended to deal with graphs of geometric constraints, has been presented. This method avoids the independency assumption taken by the probabilistic approach, being of great interest for those situations in which no probabilistic description of errors is available and only bounds on them are known.

Under the described approach, when a new sensory data is acquired, a set of strips are obtained, propagated and fused to obtain the updated ellipsoids associated with each feature. Then, the hypothesis about the location of the involved geometric features can be easily updated. Also, inconsistencies are easily detected. This means fast rejection of erroneous data.

Finally, a hardware implementation of the described method, extending the implementation presented in [2], would be of great interest for robotic systems with massive incoming sensory information, such as hands with multiple tactile sensors and vision systems extracting geometric features in real time.

6 Appendix

This appendix contains a brief survey of criteria for fusing competitive information under the probabilistic approach.

Let us assume that we have a set of measurements obtained from p different observations. The problem consists in fusing x^i, $i = 1 \ldots p$, to obtain a consensus $\hat{x}$ which *best* estimates x_r. The value of this consensus can be obtained according to the following optimality criteria:

[1] **Maximum total probability** (see [7]). This criterion assumes that the optimum $\hat{x}$ corresponds to the maximum of

$$\sum_{i=1}^{p} f_i(x^i \mid x) \tag{8}$$

where $f_i(x^i \mid x)$ is the density of probability of a feature being x^i given that the measurement is x, hence

$$f_i(x^i \mid x) = \frac{1}{(2\pi)^{\frac{n}{2}} \mid A_i \mid^{\frac{1}{2}}} e^{-\frac{1}{2}[X^i - X]^T (A_i)^{-1} [X^i - X]} =$$

$$N_X(x^i, A^i) \tag{9}$$

To find the maximum of (8), (9) can be substituted into (8) and the resulting expression derived in x. This leads to a transcendental equation which can be solved using numeric methods [7]. This criterion has, in general, very little sense since (8) may have more than one maximum.

[2] **Weighted minimum squares** (see [9]). According to this criterion, the optimum $\hat{x}$ is the value of x that minimizes the cuadratic error:

$$\varepsilon(x) = \sum_{i=1}^{p} (x - x^i)^T A_i^{-1} (x - x^i) \tag{10}$$

The solution can be easily obtained by taken the derivative of (10) with respect to x, and set it to zero, taking into account that:

$$\frac{\partial}{\partial x}[(x - x^i)^T A_i^{-1} (x - x^i)] = 2A_i^{-1}(x - x^i) \tag{11}$$

[3] **Maximum likelihood** (see [1]). Assuming that all sensory data have been obtained from independent sources, we have

$$f(x^1 \wedge x^2 \wedge \ldots \wedge x^p \mid x) = \prod_{i=1}^{p} f_i(x^i \mid x) \tag{12}$$

Taking (12) as a function in x, $f(x^1 \wedge x^2 \wedge \ldots \wedge x^p \mid x)$ is called the likelihood of x with respect to the sensory information $\{x^1, x^2, \ldots, x^p\}$. The estimation of maximum likelihood of x is the value of $\hat{x}$ that maximizes (12).

It is easier to work with the logarithm of the likelihood, than the likelihood itself. Since the logarithm is monotonically increasing, the value of $\hat{x}$ that maximizes the algorithm of the likelihood also maximizes the likelihood. This optimum is obtained by taking the logarithm of (12) and taking its derivative with respect to x using (11).

[4] **Minimum volume of uncertainty** (see [11]). Let us assume that $\hat{x}$ can be obtained as a linear combination of the sensorial information in different observations, then:

$$\hat{x} = \sum_{i=1}^{p} W_i x^i \tag{13}$$

where $W_i \in \Re^{n \times n}$ is a weighting matrix. For the estimation to be unbiased:

$$E\{\hat{x}\} = \sum_{i=1}^{p} W_i E\{x^i\} = \sum_{i=1}^{p} W_i x_r = x_r \tag{14}$$

then

$$\sum_{i=1}^{p} W_i = I \in \Re^{n \times n} \tag{15}$$

The volume of the uncertainty ellipsoid associated with $\hat{x}$ depends on the values chosen for the elements in the weighting matrix. Minimizing this volume, subject to the constraint (15), is easy through Lagrange's multipliers.

[5] **Optimum in Kalman's filter sense** (see [13]). According to Kalman filtering theory, to find an optimal solution consists in obtaining the weighting matrix W that minimizes

$$a^T V[x] a = a^T W Y W^T a \tag{16}$$

where

$$Y = \begin{pmatrix} J_1 \Lambda_1 J_1^T & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & J_p \Lambda_p J_p^T \end{pmatrix} \in \Re^{pn \times pn} \tag{17}$$

for any constant vector $a \in \Re^n$, since the dynamics is the identity and the state is static.

The optimum – for criteria [2], [3], [4] and [5] – is (1), and the associated covariance matrix, (2). Thus, (1) is an optimum in many senses. Moreover, it can be shown that $\hat{x}$ converges with probability 1 to x_r as $p \rightarrow \infty$. The likelihood function (12) is asymptotically Gaussian in x, that is

$$\prod_{i=1}^{p} f(x^i \mid x) \simeq N_y(\hat{x}, \Lambda).$$

References

[1] R.M. Bolle and D.B. Cooper, "On Optimally Combining Pieces of Information, with Application to Estimating 3-D Complex Object Position from Range Data," *IEEE Trans. on PAMI*, Vol. 8, No. 5, pp. 619-638, Sep. 1986.

[2] J.R. Deller, "A 'Systolic Array' Formulation of the Optimal Bounding Ellipsoid Algorithm," *IEEE Trans. on ASSP*, Vol. 37, No. 9, Sep. 1989.

[3] J.R. Deller, "Set Membership Identification in Digital Signal Processing," *IEEE ASSP Magazine*, pp. 4-20, Oct. 1989.

[4] H.F. Durrant-Whyte, "Uncertain Geometry," in *Geometric Reasoning*, edited by D. Kapur and J.L. Mundy, MIT Press, 1989.

[5] W.E.L. Grimson and T. Lozano-Perez, "Model-Based Recognition and Localization from Sparse Range or Tactile Data," *Inter. Journal of Robotics Research*, Vol. 3, No. 3, pp. 3-35, 1984.

[6] J.K. Hackett and M. Shah, "Multi-Sensor Fusion: A Perspective," *IEEE Inter. Conf. on Robotics and Automation*, 1990.

[7] R.C. Luo, M.H. Lin and R.S. Scherp, "Dynamic Multi-Sensor Data Fusion System for Intelligent Robots," *IEEE Journal of Robotics and Automation*, Vol. 4, No. 4, pp. 386-396, Aug. 1988.

[8] R.C. Luo and M.G. Kay, "Multisensor Integration and Fusion in Intelligent Systems," *IEEE Trans. on Systems, Man and Cybernetics*, Vol. 19, No.5, Sep./Oct. 1989.

[9] L. Matthies and A. Shafer, "Error Modeling in Stereo Navigation," *IEEE Journal of Robotics and Automation*, Vol. 3, No. 3, pp. 239-248, June 1987.

[10] M. Milanese and G. Belaforte, "Estimation Theory and Uncertainty Intervals Evaluation in the Presence of Unknown but Bounded Errors: Linear Families of Models and Estimates," *IEEE Trans. Automatic Control*, Vol. AC-27, pp. 408-414, 1982.

[11] Y. Nakamura and Y. Xu, "Geometrical Fusion Method for Multi-sensor Robotic Systems," *IEEE Inter. Conf. on Robotics and Automation*, Scottsdale, Arizona, May 1989.

[12] J. Porrill, "Optimal Combination and Constraints for Geometrical Sensor Data," *Inter. Journal of Robotics Research*, Vol. 7, No. 6, pp. 97-113, Dec. 1988.

[13] R.C. Smith and P. Cheeseman, "On the Representation and Estimation of Spatial Uncertainty," *Inter. Journal of Robotics Research*, Vol. 5, No. 4, pp. 56-68, 1986.