FILTERING DATA BY USING THREE ERROR THEORIES TOGETHER: THE GUESS FILTER

Olivier STRAUSS, Eric AIME
LIRMM - Université Montpellier II - 161 rue ADA
34392 Montpellier Cedex 5 - FRANCE
Tel. (33)67418562 - Fax (33) 67418500
E. mail : strauss@lirmm.lirmm.fr

ABSTRACT: Most papers dealing with noisy measurements combination try and use a single tool providing the "best" estimation. But, as pointed out by Dubois and Prade [DUB 90], most tools dealing with imperfect information have different purposes. The purpose of this paper is to try and find a way of using three different frameworks together in order to obtain a filter that combines robustness, accuracy, reliability and easy implementation. To perform this estimation, possibility theory is considered to deal with precision while a statistical tool reduces uncertainty. Then, we propose to use the rough sets theory to perform robust and easy computation of the resulting filter. Some results on real data provided by the compass of a submarine robot are presented.

1. INTRODUCTION.

The proliferation of inexpensive sensors and the increasing complexity of robotics tasks requires the perception system to provide more accurate and reliable information. A filter aims to handle with noisy measurements in order to provide such an information.

Several mathematical tools are available to deal with errors in measurements. They are often divided in different categories depending on the mathematical theory they are based on. During the last ten years, probabilistic theory has been widely used to perform filtering. The core of those techniques is to use the past knowledge about the occurrence of an event to infer the chance of occurrence of a similar event in the future [JAZ 70]. As voting procedures, they are very concerned with uncertainty. Conversely, set-based techniques deal with imprecision. In this framework, measures are considered as a set containing the information. Handling this information consist of doing set operations like union, intersection or volume approximations. Union provides the most reliable set while intersection provides the most accurate set that is supposed to include the information to be recovered.

However, measurements provided by sensors are generally contaminated by both imprecision and uncertainty. Using only set-based or probabilistic techniques hides one of the aspects of the problem. Moreover, lack of knowledge on the signal or on the sensor cannot be modeled in an easy way.

As shown in [DUB 88], possibility theory allows the representation of both uncertainty and imprecision in the same framework. The present paper propose the use of this framework for filtering.

2. FILTERING WITH POSSIBILITY.

Let m_k be an uncertain and inaccurate measurement of a signal θ_k provided by a sensor S. Because the sensor is inaccurate, the measurement can be considered as a fuzzy sub-set M_k (generally an interval) of the set Ω of the possible values of θ. From the possibilistic point of view, filtering the measurements M_k consists of finding a possibility distribution $\pi(\theta_k)$ of θ_k on Ω i.e. a fuzzy subset Θ_k of the possible values of θ on Ω at time k.

Let us first consider θ_k to be stationary (i.e. $\theta_k=\theta\ \forall k$). From a set-theoretic point of view, there are two extreme modes of combination depending on the reliability of the sensor. If the sensor is fully reliable, then Θ_k can be obtained by intersecting all the M_k. Conversely, if the M_k are not reliable, union-like combination is to be performed. This statement is linked to the duality precision/reliability. Disjunctive mode of combination increases precision while reliability decreases. The dual effect is obtained by conjunctive mode.

No sensor can be supposed to be fully reliable. The lack of precision obtained by disjunctive mode is undesirable. Now, if θ_k is not stationary, then the influence of the information given by the measurement M_{k-n} for estimating Θ_k has to weaken when n increases.

So, filtering the measurements consists of a dissymmetric combination process. The new measurement M_{k+1} is used to modify the a-priori knowledge on θ i.e. to revise $\mu_{\Theta_k}(\theta) = \pi_k(\theta)$ in the light of $\mu_{M_{k+1}}(\theta)$.

In [DUB 92], several symmetric and dissymmetric aggregation rules are proposed. The core of those techniques is to find a way of changing gradually the combination mode from conjunction to disjunction according to a measurement of the conflict between the measurements. According to possibility theory, this conflict can be measured by the conditional possibility and conditional necessity:

$$\Pi(\Theta_k|M_{k+1}) = \underset{\theta\in\Omega}{\mathrm{Sup}}\left\{\min\left(\mu_{\Theta_k}(\theta), \mu_{M+1}(\theta)\right)\right\} \qquad N(\Theta_k|M_{k+1}) = \underset{\theta\in\Omega}{\mathrm{Inf}}\left\{\max\left(\mu_{\Theta_k}(\theta), 1-\mu_{M+1}(\theta)\right)\right\}$$

In [DEV 93], an application of the more elaborate symmetric aggregation rule has been implemented for filtering. Its performance has been qualitatively compared with that of a Kalman filter. If the signal is fairly corrupted by contaminated noise, the conflicts seem to be ill-handled and this filter is rather unstable. This is due to the use of a symmetrical rule to perform a dissymmetric problem: the same weight cannot be given to the a priori knowledge and the new measurement.

To overcome these problems, we propose to combine a voting process with a non voting process and to avoid normalization of the possibilistic distribution. The voting part of the process is performed by a first order statistical process. This new process allows the representation of confidence and uncertainty knowledge, at two different levels. It is performed in two steps.

First, a symmetrical combination rule is applied to provide a deduced set D_{k+1}:

$$\mu_{D_{k+1}}(\theta) = \min\left(\min\left(\mu_{\Theta_k}(\theta), \mu_{M_{k+1}}(\theta)\right), \eta\right) + \min\left(\max\left(\mu_{\Theta_k}(\theta), \mu_{M_{k+1}}(\theta)\right), 1-\eta\right)$$

where $\eta = \Pi(\Theta_k|M_{k+1}) + N(\Theta_k|M_{k+1}) - \Pi(\Theta_k|M_{k+1}).N(\Theta_k|M_{k+1})$ is an estimate of the conflict between the prediction and the measurement.

Then, D_{k+1} is used to update Θ_k to give the final distribution:

$$\mu_{\Theta_{k+1}}(\theta) = \alpha.\mu_{\Theta_k}(\theta) + (1-\alpha).\mu_{D_{k+1}}(\theta) \quad \text{with } \alpha\in[0,1].$$

This last update is performed by a first order exponential statistical filter on the possibility distribution instead of being applied on the value itself. α can be fixed a priori if the noise has been clearly identified or be updated by a fuzzy rule or Baye's rules depending on the a priori knowledge available on the process.

A problem still remains: computation of such rules leads to the need of an appropriate data structure.

3. APPROXIMATION OF THE POSSIBILITY DISTRIBUTION.

Two classical ways are mainly available to perform the computation of this filter. First of all, the fuzzy set can be considered as an union of fuzzy intervals. The second possibility consist in sampling the membership function on Ω. Let us see the advantages and the drawbacks of each methods.

Assuming that the measurements are fuzzy intervals, the union-like and intersection-like operations performed in the filtering process also provide fuzzy intervals. The advantage of such a representation is accuracy. The drawback is that the computation time highly depends on the number of elements of the list. This limitation is not acceptable if the filter is to be implemented on a real process. Moreover, non-consistent intervals have to be removed in order to avoid an exponential memory request. If such rules are used, then the benefit of accuracy of this representation vanishes. Otherwise, the memory can explode.

Hence, a sampled representation is generally preferred. It consists of assuming that the fuzzy membership can be represented by membership function on sparse elements of Ω. In fact, this representation is an incomplete and imprecise representation of this membership function. Moreover, the data structure do not take into account this coarsening.

In order to deal with this rough representation and handle computation in a proper manner, we propose to use the recent theory of fuzzy rough sets proposed by Dubois and Prade in [DUB 90]. This will be performed by decomposing the possibility distribution on a fuzzy partition of Ω.

First of all, let us have a brief overview of fuzzy rough sets. Fuzzy sets and rough sets both deal with imprecision. But while the poor definition of boundaries of sub-classes are properly modeled by fuzzy sets, rough sets are more concerned with the objects in a set being indiscernible. The key idea of using rough decomposition is to make the intrinsic indiscernibility of sampled computation part of the computation itself.

A fuzzy partition of Ω is a family (a_i) of N fuzzy sub-sets of Ω. Some properties are requested for a partitioning:

i) The (a_i) are supposed to cover Ω enough : $\operatorname{Inf}_{\theta\in\Omega}\left\{\max_{i=1\ldots N}\left(\mu_{a_i}(\theta)\right)\right\} > 0$

ii) The (a_i) are disjoint: $\operatorname{Sup}_{\theta\in\Omega}\left\{\min\left(\mu_{a_i}(\theta),\mu_{aj}(\theta)\right)\right\} < 1\ \forall(i,j)\in[1,N]^2$

iii) The (a_i) provide a uniform partition: $\sum_{i=1\ldots N}\mu_{a_i}(\theta) = 1\ \forall\theta\in\Omega$.

Then, a rough decomposition of the fuzzy subset Θ on the fuzzy partition (a_i) is given by the mean of N pairs (Π_i,N_i) with:

$$\Pi_i(\Theta) = \Pi(\Theta|a_i) = \operatorname{Sup}_{\theta\in\Omega}\left\{\min\left(\mu_\Theta(\theta),\mu_{a_i}(\theta)\right)\right\}$$

$$N_i(\Theta) = N(\Theta|a_i) = \operatorname{Inf}_{\theta\in\Omega}\left\{\max\left(\mu_\Theta(\theta),1-\mu_{a_i}(\theta)\right)\right\}$$

Π_i can be seen as the degree of possible membership of (a_i) in Θ, while N_i is a degree of certain membership. Fig. 1 illustrates of such a concept on a triangular decomposition.

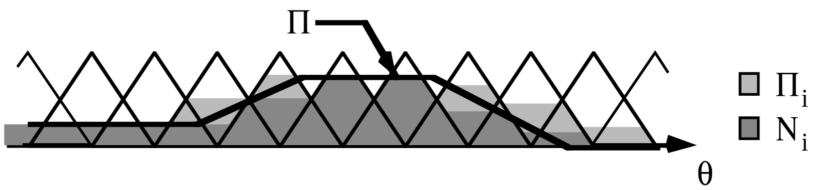


Fig. 1: decomposition of a fuzzy set on a fuzzy partition.

Triangular decomposition is the most commonly used for different reasons. It is easy to compute. It satisfies all the expected properties (i,ii,iii), and it is the most neutral way of representing a sparse knowledge. A triangular set is a fuzzy number [DUB 81] with a linear form function. A triangular fuzzy number can be represented by two parameters: spread and mean.

Finally, a desirable property can be added: $\operatorname{Inf}_{\theta\in\Omega}\left(\max_{i=1\ldots N}\left(N(M_k|a_i)\right)\right) > 0$ that limits the loss of information due to decomposition.

Some properties still hold on rough decomposition on a triangular fuzzy partition:

$$\Pi((P\cap Q)|a_i) = \min\left(\Pi(P|a_i),\Pi(Q|a_i)\right) \text{ and } N((P\cup Q)|a_i) = \max\left(N(P|a_i),N(Q|a_i)\right)$$

But other properties don't:

$$\Pi((P\cup Q)|a_i) \le \max\left(\Pi(P|a_i),\Pi(Q|a_i)\right) \text{ and } N((P\cap Q)|a_i) \ge \min\left(N(P|a_i),N(Q|a_i)\right)$$

Because we aim at computing the upper and lower bounds of the distributions, if no equality is available we take the upper bound for possibility and the lower bound for necessity.

Then, the formulation of the filter becomes:

$$\begin{cases}\Pi(\Theta_k\cap M_{k+1}|a_i) = \min\left(\Pi(\Theta_k|a_i),\Pi(M_{k+1}|a_i)\right)\\ \Pi(\Theta_k\cup M_{k+1}|a_i) \approx \max\left(\Pi(\Theta_k|a_i),\Pi(M_{k+1}|a_i)\right)\end{cases} \text{ and } \begin{cases}N(\Theta_k\cap M_{k+1}|a_i) \approx \min\left(N(\Theta_k|a_i),N(M_{k+1}|a_i)\right)\\ N(\Theta_k\cup M_{k+1}|a_i) = \max\left(N(\Theta_k|a_i),N(M_{k+1}|a_i)\right)\end{cases}$$

$$\Pi(D_{k+1}|a_i) = \min\left(\Pi(\Theta_k\cap M_{k+1}|a_i),1-\eta_k\right) + \min\left(\min\left(\Pi(\Theta_k\cup M_{k+1}|a_i),1-\eta_k\right),1-N(\Theta_k|a_i)\right)$$

$$N(D_{k+1}|a_i) = \min\left(N(\Theta_k\cap M_{k+1}|a_i),1-\eta_k\right) + \min\left(\min\left(N(\Theta_k\cup M_{k+1}|a_i),1-\eta_k\right),1-\Pi(\Theta_k|a_i)\right)$$

with $\eta = \Pi(\Theta_k|M_{k+1}) + N(\Theta_k|M_{k+1}) - \Pi(\Theta_k|M_{k+1}).N(\Theta_k|M_{k+1})$

and $$\begin{cases} \Pi\left(\Theta_k \mid M_{k+1}\right) \approx \sup\limits_{(i,j)=[1,N]^2} \left\{ \min\left(\Pi(\Theta_k \mid a_i), \Pi(M_{k+1} \mid a_j), \Pi(a_i \mid a_j) \right) \right\} \\ N\left(\Theta_k \mid M_{k+1}\right) \approx 1 - \sup\limits_{(i,j)=[1,N]^2} \left\{ \min\left(N(\Theta_k \mid a_i), \Pi(M_{k+1} \mid a_j), \Pi(a_i \mid a_j) \right) \right\} \end{cases}$$

4. DEFUZZIFICATION.

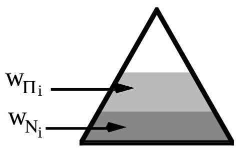
The use of such a filter in a real process implies the ability to provide one real crisp value for the obtained distribution. Such a transformation is also needed to allow some comparisons between this filter and other filters. This process is known in fuzzy control as "defuzzification".

For monomodal statistical processes, like Kalman filter, this transformation is trivial because the best estimate is part of the distribution representation. For multi-modal processes, even if those are statistical processes, this transformation is not problem-free.

Literature provides several ways of performing a defuzzification. Because the possibility distribution can be seen as a sparse knowledge, the main problem addressed here is to find a compromise between all these pieces of information. In general, the so-called barycenter method is used. This method consists in calculating the barycenter of the obtained fuzzy sets. However some problems arise. As a matter of fact, if the fuzzy set is made of the union of two disjuncted intervals (e.g. $[21,25] \cup [29,31]$), this approach will provide a value (here 25.625) which does not belong to any fuzzy set (i.e. is not possible according to the knowledge). Conversely, the maximum method consists of finding the peak in the fuzzy subset. In case there is more than one peak, a mean value is provided or a selection rule is applied. This method has a flaw: it provides a highly biased value.

In order to avoid most of the drawbacks of each method, we use the method proposed in [DUB 88] consisting of associating a probability distribution to the given possibility distribution. This process is performed in three steps:

1) two weights are associated to each cell taking into account the shape of the cell. w_{Π_i} (rsp. w_{N_i}) is computed by using the conditional possibility (rsp necessity) of the cell. Then the weight is computed using a mean operation.



$$w_i = M\left(w_{\Pi_i}, w_{N_i}\right)$$

Fig. 2: weight associated to each cell

2) the weights are sorted by order: $w'_i \geq w'_{i+1}$ and $w'_{N+1} = 0$

3) the probability distribution is then obtained by: $P_i = \sum\limits_{k=i}^{N} \frac{1}{k}\left(w'_i - w'_{i+1}\right)$

Then the probability is used to weight each cell to provide the estimate. This method can be improved by using robust statistics [HUB 81].

Finally, three measurements are available to qualify the fuzzy estimation: a confidence is given with the possibility and the necessity, while the fuzzy cardinal gives a measurement of the precision.

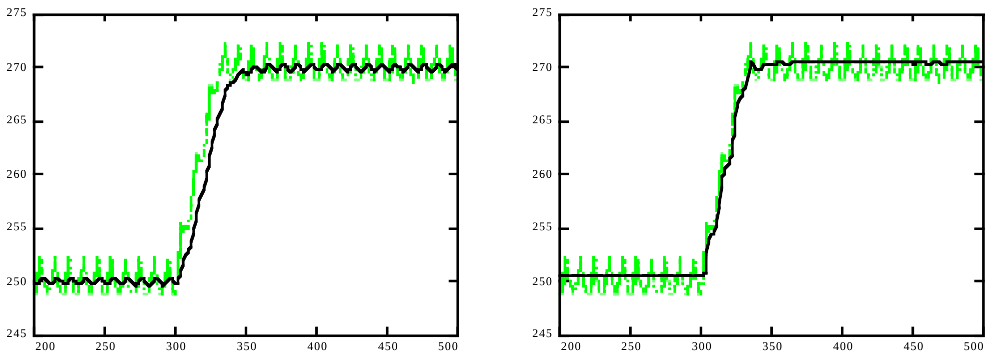
Because it works on hypothesis verifications at different levels, we called this filter: “Guess filter”.

5. EXPERIMENTAL RESULTS.

This algorithm has been tested on simulated and real data. The above examples are real signal provided by the compass mounted on the submarine vehicle OTTER*. The noise on this compass has been clearly identified and has been represented in both frameworks: possibility and probability (bounded and non-bounded error).

* OTTER is a semi autonomous submarine vehicle developped by the Monterey Bay Aquarium Resarch Institute.

This example allows a qualitative comparison between a first order Kalman filter and the Guess filter. It demonstrates the ability of the latter to distinguish sudden transition and random noise. The parameters of the Kalman filter have been chosen to provide a good compromise between the ratio signal/noise and the delay due to sudden transition. The behavior difference between the KF and the GF is due to the error models on which they work . Our experiments have demonstrated that this filter seems to be more *robust* to the hypothesis on noise distribution.



Kalman Filter   Guess Filter

Fig. 3: qualitative comparison on a real signal

6. CONCLUSION

The present paper has proposed a new way of filtering a signal by using three theories together: possibility-, probability- and rough-sets theory. Using these three frameworks together allows to deal with all the error problems in signal processing. First, possibility theory is a general enough framework to allow the representation of both uncertainty and imprecision in a quite simple manner. Using set-theoretic rules deals with current imprecision, with compromise rules aggregating past knowledge and observation from conjunctive to disjunctive mode depending on their conflicts. Then, using a statistical exponential smoothing filter palliates the lack of knowledge on the process and on the error model variations. Finally, approximation due to computation and sampling are taken into account with the rough-subset theoretical framework. This last framework makes the error in computation part of the computation itself. The remaining work is concerned with the optimization of the computation of the filter and its utilization for sensor data fusion.

REFERENCES

[DEV 93] S. DEVEUGHELE and B. DUBUISSON – “Handling Suden State Transition” – IEEE International Conference on Systems, Man and Cybernetics – vol.4, (pp 680…686) – Le Touquet, France– oct. 1993.

[DUB 81] D. DUBOIS and H. PRADE – “Additions of Interactive Fuzzy Numbers” – IEEE transaction on Automatic Control – vol.26, n°4 (pp 926…936) – 1981.

[DUB 88] D. DUBOIS and H. PRADE – “Possibility Theory: an Approach to the Computerized Processing of Uncertainty” – Plenum Press – New York – 1988.

[DUB 90] D. DUBOIS and H. PRADE – “Rough Fuzzy Sets and Fuzzy Rough Sets” – International Journal of General Systems – n°17 (pp 191…200) – 1990.

[DUB 92] D. DUBOIS and H. PRADE – “Combination of Fuzzy Information in the Framework of Possibility Theory” – Data Fusion in Robotics and Machine Inteligence (pp. 481…505)– Academic Press – New York – 1992.

[HUB 81] P. HUBER – “Robust Statistics” – Wiley ed. – New York – 1981.

[JAZ 70] A. Jazwinski – “Stochastic Processes and Filtering Theory” – Academic Press – New York – 1970.