Qualitative model-based multisensor data fusion and parameter estimation using $\infty$-norm Dempster-Shafer evidential reasoning *

Steven Reece
Dept. of Engineering Science University of Oxford
Oxford England OX1 3PJ †

ABSTRACT

This paper is concerned with model-based parameter estimation for noisy processes when the process models are incomplete or imprecise. The underlying representation of our models is qualitative in the sense of Interval Arithmetic and Qualitative Reasoning (QR) and Qualitative Physics from the Artificial Intelligence literature. We adopt a specific qualitative representation, namely that advocated by Kuipers, in which a well defined mathematical description of a qualitative model is given in terms of operations on intervals of the reals.

We investigate an weighted opinion pool formalism for multi-sensor data fusion, develop a definition for unbiased estimation on quantity-spaces and derive a consistent mass assignment function for mean estimators for two state systems. This is extended to representations involving more than two states by utilising the relationships between coarse (i.e. two state) and fine (i.e. N state) representations explored by Shafer. We then generalise the Dempster-Shafer Theory of Evidence to a finite set of theories and show how an extreme theory can be used to develop mean minimum-mean-square-error (MMSE) estimators applicable to situations with correlated noise.

We demonstrate our theory using real data from a mobile robot application which utilises sonar and laser time-of-flight and gyroscope information to disseminate surface curvature.

Keywords: qualitative reasoning, non-linear estimation, Dempster-Shafer theory, data fusion.

1 INTRODUCTION

The relationships between sensory information in data fusion systems is often in the form of quantitative, algebraic models and, in practise, parameter estimation systems incorporating such models suffer numerous problems. For example, complex models can be unwieldy, the interpretation of sensor output might require detailed knowledge of the physics of the sensing process or extensive sensor calibration might be required. Also, sufficient representation of uncertainty must be included to cater for unmodelled or non-determinable components such as temperature changes or hysteresis effects in sensor components. The Kalman filter, uses algebraic differential models which are hardly ever accurate descriptions of the environment at this level of precision and the estimate often diverges (i.e. true state and estimated state differ by a significant multiple of the standard deviation in the estimate). Apart from a few simple cases (for example, when the modelling error is known to be constant analytical solutions to modelling inaccuracies are cumbersome and the alternatives are to mask the deficiencies of the process model by introducing fictitious noise inputs, over-weighting the most recent data (limited memory filtering) or by adaptive noise estimation using observation and estimation residuals. These methods require adjustment for each application by experimentation and are therefore fragile. We believe

*This work is supported by EPSRC grants GR/J/46067 and GR/J/57773.
†Further author information: Phone +44 1865 282180, Fax +44 1865 273008, Email reece@robots.ox.ac.uk.
that the process model itself should reflect the imprecision when information is incomplete and to this end, we explore an alternative modelling representation: the qualitative model.

Often, precise numerical measures are not required when less detailed, (qualitative) information is perfectly adequate to draw the necessary conclusions. In such situations we may represent imprecise information as a range of real values. For example, the linguistic statement “to the left” may be represented as a range of possible bearings \((0, \frac{\pi}{2})\) radians and the statement “monotonically increasing function of \(y\) and \(x\)” would be the set of functions for \(y\) whose derivative with respect to \(x\) lies in \((0, \infty)\). Smaller intervals may be used where more detailed information is available. For example, when the zero return from a gyroscope sensor subject to drift is guaranteed to correspond to a true value within \((-0.01, 0.01)\) rads per sec.

Ideally, we seek to utilise all the information available whether it is incomplete, imprecise or uncertain while maintaining an accurate estimate of state. In this paper, we describe the Qualitative filter for qualitative estimation and sensor fusion, which uses abstract (as opposed to approximate) process and observation models to maintain a qualitative measure of the parameter values while utilising the abstraction properties of the Kalman filter: namely, its ability to maintain an estimate of mean and variance irrespective of the exact form of the underlying noise probability density distributions.

In Section 2, we introduce the requirements for a qualitative data fusion framework with an example drawn from the robot sensing domain. In Section 3, we describe the conditions for unbiased estimation using \(\infty\)-norm opinion pool methods and Dempster-Shafer mass and introduce mass assignment functions for mean estimators analogous to the Kalman filter. We develop the \(\infty\)-norm Dempster-Shafer theory and subsequently derive expressions for the filter prediction and update phases. Finally, in Section 4, we demonstrate the Qualitative filter with real data obtained in the robot sensing domain.

2 THE ROBOT SENSING DOMAIN

This paper is concerned with data fusion when knowledge of physical systems is incomplete. A qualitative representation is one that captures distinctions that make an important, categorical difference, and ignores others. For example, a natural set of qualitative regions for describing the orientation of a feature relative to a robot is defined by the following landmarks (see Figure 2):

\[
\text{in front of} \ldots \text{left of} \ldots \text{behind} \ldots \text{right of}.
\]

The sensors we investigate (sonar, laser time-of-flight) offer a 2-D view of the environment (1-D signal propagation and 1-D rotation of sensor head). Therefore, the robot identifies features through the intersection of a plane with the environment and Leonard 10 shows that feature curvature is a good discriminant in such circumstances. Hoffman 6 contemplates encoding shape by curvature and concludes that a qualitative description of curvature suffices to capture the essential discriminants of the environment. However, since curvature is a second order derivative it is noisy 1.

In this section, we investigate how qualitative information about curvature can be computed and we demonstrate how a mobile robot, equipped with a sonar tracking sensor, a laser time-of-flight range sensor and a gyroscope (see Figure 1) can distinguish feature curvature using qualitative models inter-relating its sensor cues. In later sections, we will investigate the problem of noise filtering using qualitative models.
In Figure 2, the robot (indicated by $\Delta$) passes in front of a curved surface (shaded). For specular (i.e. mirror-like) surfaces the robot sees only the reflection normal to the surface (i.e $\mu = 0$) $^8, ^{10}$. The robot tracks the feature by following this normal reflection (see Figure 3). If the sonar turns with rate $\dot{\phi}$ while tracking a beacon with curvature $\frac{1}{r}$ lying a distance $R$ from the robot and at a bearing $\theta$ relative to the forward motion of the robot then $^{17}$:

$$\dot{\phi}(r + R) = -\dot{R}\tan \theta.$$  \tag{1}

The robot is able to determine the surface curvature $r$ of a feature in the environment by fusing qualitative information about $\tan \theta$, $\dot{R}$ and $\dot{\phi}$. Table 1 shows qualitative behaviours consistent with Equation (1). $^1$ To illustrate the interpretation of this table consider the first entry in conjunction with the convex surface depicted in Figure 3. When the feature lies left-frontal (i.e. $\text{Qmag}(\tan(\theta)) = +$) and the range $R$ is decreasing (i.e. $\text{Qdir}(R) = -$) and the gyro is turning anti-clockwise (i.e. $\text{Qdir}(\phi) = -$) then the feature must be a convex surface. This table can be constructed from Equation (1) using qualitative knowledge of the quantitative algebraic building blocks $^9$. For example, qualitative multiplication $\text{Qmult}$: $[\text{Qmag}(A) = +] \land [\text{Qmag}(B) = +] \Rightarrow [\text{Qmult}(A, B) = +]$ or qualitative negation $\text{Qminus}$: $[\text{Qmag}(A) = +] \Rightarrow [\text{Qminus}(A) = -]$. As is evident from Table 1, however, the sonar/gyroscope

<table>
<thead>
<tr>
<th>$\text{Qmag}(\tan \theta)$</th>
<th>$\text{Qdir}(R)$</th>
<th>$\text{Qdir}(\phi)$</th>
<th>$\text{Qmag}(\frac{1}{r})$</th>
<th>Surface Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>Convex</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+/−</td>
<td>Concave/Convex</td>
</tr>
<tr>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+/−</td>
<td>Concave/Convex</td>
</tr>
<tr>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>Convex</td>
</tr>
<tr>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+/−</td>
<td>Concave/Convex</td>
</tr>
<tr>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+/−</td>
<td>Convex</td>
</tr>
<tr>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>Convex</td>
</tr>
<tr>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+/−</td>
<td>Concave/Convex</td>
</tr>
<tr>
<td>+/0/−</td>
<td>+/−</td>
<td>0</td>
<td>0</td>
<td>Plane</td>
</tr>
</tbody>
</table>

Table 1: Curvature inferred from sonar and gyroscope cues.

Figure 3: Normal incidence reflection behaviour. Close-concave behaviour is observed when the robot moves between a concave surface and the focal point $F$ of the surface.

\footnote{The qualitative magnitude and derivative operators are:

$\text{Qmag}(X) = \text{sign}(X) \subseteq (-\infty, \infty)$

$\text{Qdir}(X) = \text{sign}(X) \subseteq (-\infty, \infty)$.}
Combination is unable to disambiguate curvature types when \( R + r > 0 \) (i.e. when \( \frac{R \tan \mu}{r} < 0 \)). Sonar cues for *far-concave* surfaces exhibit identical qualitative behaviour to *convex* surface cues (see Figure 3). For specular objects the robot is unable to gain any further information by scanning the object using the sonar sensor. However, if we use a range sensor with a shorter signal wavelength, small blemishes on the surface of object reflect the signal back towards the sensor for any incident angle \( \mu \). Since the wavelength of laser light (typically \( 800 \text{nm} \)) is significantly shorter than that of ultra-sound (typically \( 0.5 \text{cm} \)), imperfections in the surface cause the laser signal to scatter. The qualitative behaviour of range against bearing for a complete scan of the surface using laser light is able to disambiguate *far-concave* and *convex* surfaces when \( R + r > 0 \). To see this, if \( \mu \) is the reflectance angle of the laser signal relative to the feature normal (see Figure 2), then it can be shown that:

\[
\frac{\partial^2 R}{\partial \phi^2} = R \tan(\mu)^2 + R \left[ 1 + \tan(\mu)^2 \right] \left[ 1 + \frac{R}{r \cos \mu} \right].
\]

Since for all scenarios \( R, \tan^2 \mu \) and \( \cos \mu \) are positive, then we can see that when \( r + R > 0 \), \( \frac{\partial^2 R}{\partial \phi^2} < 0 \) if and only if the surface is *concave* \((r < 0)\). Thus, when \( R + r > 0 \), the laser is able to resolve the ambiguity in the sonar qualitative model. However, using qualitative information alone, the laser is unable to disambiguate *planar* and *convex* surfaces. Hence, we must use the sonar and laser sensors together to determine the surface curvature. Thus, not only does data fusion eliminate noise but it also helps to overcome ambiguity problems in the qualitative models.

In the remainder of this paper, we develop efficient methods for filtering noise in systems employing qualitative representations.

### 3 THE QUALITATIVE FILTER

Unlike quantitative model-based estimation systems where parameter point estimates are inter-related using precise quantitative models, no such precision exists within the qualitative approach. Consider the way in which estimates \( \hat{x} \) and \( \hat{y} \) for two variables \( x \) and \( y \) are combined using the Kalman filter (see Figure 4). A function \( f \) (i.e. process or observation model) transforms \( \hat{x} \) into the \( y \)-space and the point estimates \( f(\hat{x}) \) and \( \hat{y} \) are combined by weighted averaging. When \( f \) is not known precisely we no longer have the requisite point estimates since the best we can say about \( f(\hat{x}) \) is that it must take some value between \( Y_0 \) and \( Y_1 \).

We solve this problem by transforming our point estimate \( \hat{x} \) into a Dempster-Shafer mass estimate for the region \((X_0, X_1)\). This mass value is then assigned to the region \((Y_0, Y_1)\) according to the imprecise description of \( f \) (that any point in the region \((X_0, X_1)\) maps into \((Y_0, Y_1)\)) and fused with mass estimates generated by \( \hat{y} \) using the Dempster rule of combination.

<table>
<thead>
<tr>
<th>Qmag(( \frac{1}{x-\mu} ))</th>
<th>Qmag(( \frac{\partial^2 R}{\partial \phi^2} ))</th>
<th>Qmag(( \frac{\partial \phi}{\partial \phi} ))</th>
<th>Surface Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>convex</td>
</tr>
<tr>
<td>0</td>
<td>+</td>
<td>0</td>
<td>plane</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>-</td>
<td>close-concave</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>far-concave</td>
</tr>
</tbody>
</table>

Table 2: Curvature inferred from sonar, gyroscope and laser cues.

**Figure 4:** Hypothetical function \( f \) relating two variables \( x \) and \( y \).
3.1 Notation

At this point, we must introduce the concepts used in the theory developed in later sections.

The true state $x$ of the system is a vector $x = (x_0, \ldots, x_n)$ in which the parameter $W_i$ takes the quantitative value $x_i$.

A qualitative value is an interval of the reals. A state $Q$ of a set of parameters $\{W\}$ is a vector of qualitative values $Q = (q_0, \ldots, q_n)$, in which the parameter $W_i$ takes the qualitative value $q_i$. The quantity space of a parameter is the set of possible qualitative values the parameter can take.

A Finite State Machine (FSM) takes the role of the process and observation models of the Kalman filter and describes the allowable evolutions of a system. A finite state machine has an associated set $T \subseteq S \times S$ where $S$ is the set of states. Thus, $(Q_i, Q_j) \in T$ iff there exists a single-step transition of the FSM from $Q_i$ to $Q_j$. The set of successor states $R(Q)$ from any state $Q$ is defined:

$$R(Q) = \{ Q'| (Q, Q') \in T \}.$$  

A basic path in FSM is a sequence of states $Q_0, Q_1, \ldots, Q_m$ such that $\forall i \in 0, \ldots, m-1, (Q_i, Q_{i+1}) \in T$. A compound path which ends at $Q_m$ is the union of basic paths, each of which end at $Q_m$. A path (denoted $\omega$) is either a basic path or a compound path.

A single observation is denoted $z_i$ and a set of $t+1$ observations is denoted $z_{0:t}$ (which may include observations from different sensors).

The mass assigned to a state $Q$ by the set of $t+1$ observations $z_{0:t}$ along path $\omega$ ending at $Q$ is denoted $m_{\omega}(Q|z_{0:t})$.

The operator $E$ is the expectation over observation ensembles $z_{0:t}$.

3.2 The P-Norm Estimator Framework

We advocate a weighted information pooling paradigm to data fusion. The total evidence assigned to state $Q$ due to the observation set $z_{0:t}$ is the sum over the evidence offered by each alternative path leading to $Q$. The evidence assigned to each path is reinforced by each observation. We introduce a general class of estimators with free parameters which assign weights to paths (i.e. parameter $p \geq 0$) and individual observations (i.e. parameters $K_i \geq 0$). For any state $Q$ and all paths $\omega$ which end at $Q$ the p-norm estimate $m(Q|z_{0:t})$ assigned to $Q$ due to observations $z_{0:t}$ is:

$$m(Q|z_{0:t}) = k \left[ \sum_\omega \prod_{i=0}^t m_{\omega}^K \cdot p(Q|z_i) \right]^{1/p}$$  \hspace{1cm} (2)

where $k$ is a normalisation constant so that $\left[ \sum_\omega m^p(Q|z_{0:t}) \right]^{1/p} = 1$. Bayesian methods and the Dempster rule of combination belong to the set of I-norm estimators. If for some state $Q$, $\lim_{t \to \infty} m(Q|z_{0:t}) = 1$, then $m$ is said to have converged to $Q$. If $m(Q|z_{0:t})$ converges then, it must do so to a state $Q$ which contains the true state $x$ otherwise the estimator is deemed to be biased. For the class of $\infty$-norm estimators (i.e. $p \to \infty$), if $\omega$ and $\omega'$ are paths which end at $Q$ and $Q'$ respectively after $t$ observations, then an $\infty$-norm estimator $m(Q|z_{0:t})$ is unbiased if:

$$\exists Q : x = Q, \forall Q', \omega. E \left( \log \frac{m_{\omega}(Q|z_{0:t})}{m_{\omega'}(Q'|z_{0:t})} \right) \geq 0.$$  

We initially obtain a general unbiased estimator for binary quantity-spaces $\{Q, \neg Q\}$ and then build estimators for $N$ region spaces from this. For a landmark $l$ separating the decision regions $Q$ and $\neg Q$, the following mass assignment function is an estimator for the mean of an arbitrary pdf $f(z|x)$:

$$m_{mean}(Q|z) = \begin{cases} 1, & z \in Q \\ \exp(-|z-l|), & z \in \neg Q. \end{cases}$$  \hspace{1cm} (3)

\footnote{The FSM is generated from qualitative constraints (confluences in the QR literature) by, for example, QSIM.}

\footnote{The weights $p$ and $K_i$ must be positive so that the evidence is assigned to the estimate with appropriate polarity.}
One immediate conclusion is that the mean estimator log likelihood ratio \( \log \frac{m_{\text{mean}}(Q|z)}{m_{\text{mean}}(\neg Q|z)} \) between binary states \( Q \) and \( \neg Q \) converges, on average, more slowly the closer \( x \) is to the landmark \( l \). The choice of \( K_i \) is therefore crucial to the convergence rate of the estimator and in Section 3, we describe the form for \( K_i \) which maintains the most certain estimates.

To extend our estimation framework from binary quantity-spaces to \( N \) region spaces, we show that filtering in a quantity-space comprising \( N - 1 \) landmarks can be formulated as the conjunction of \( N - 1 \) binary decision problems, each comprising a single landmark. We use the notion of coarse frames of discernment from the Dempster-Shafer Theory of Evidence.

A coarse frame of discernment \( Y \) comprises propositions which are supersets of propositions in \( \Theta \) (\( \Theta \) is subsequently called a refined frame). For each member \( v \in Y \) the refining operator \( \omega_Y : 2^Y \rightarrow 2^\Theta \) gives the subset \( \omega_Y(v) \) of \( \Theta \) consisting of those possibilities into which \( v \) can be split.

Each of the \( N - 1 \) coarse frame mass assignments \( m_Y(Q|z) \) obtained from an observation using Equation (3) can be seen as a source of information pertaining to the true state of the system and the true state must have common consensus for all sources. Therefore, the likelihood of \( x \) belonging to a state in \( \Theta \) is represented by the conjunction of the \( N - 1 \) coarse frame mass distributions. To illustrate, the three coarse frames in Figure 5 denoted \( A \), \( B \), and \( C \) are coarsenings of the underlying frame of discernment.

![Figure 5: The three binary frames of discernment \( Y \) corresponding to the four region refined frame \( \Theta \).](image)

\[ \Theta = \{Q_0, Q_1, Q_2, Q_3\} \] such that:

\[
\omega_A(\{A\}) = \{Q_0\} \quad \omega_A(\{\neg A\}) = \{Q_1, Q_2, Q_3\}
\]

\[
\omega_B(\{B\}) = \{Q_0, Q_1\} \quad \omega_B(\{\neg B\}) = \{Q_2, Q_3\}
\]

\[
\omega_C(\{C\}) = \{Q_0, Q_1, Q_2\} \quad \omega_C(\{\neg C\}) = \{Q_3\}
\]

Since, for any individual binary coarse frame \( v \in Y \),

\[
\omega_Y(\omega_Y(v)) = \omega_Y(v)
\]

we have the following refined frame assignments from the mass assigned to the coarse frames by consensus:

\[
m_\Theta(\{Q_0\}) = km_A(\{A\})m_B(\{B\})m_C(\{C\})
\]

\[
m_\Theta(\{Q_1\}) = km_A(\{\neg A\})m_B(\{B\})m_C(\{C\})
\]

\[
m_\Theta(\{Q_2\}) = km_A(\{A\})m_B(\{\neg B\})m_C(\{C\})
\]

\[
m_\Theta(\{Q_3\}) = km_A(\{\neg A\})m_B(\{\neg B\})m_C(\{\neg C\})
\]

where \( k \) is a normalisation factor. Normalising by the mass assigned to the central state \( m_\Theta(\{Q_1\}) \), say, we obtain:

\[
\frac{m_\Theta(\{Q_0\})}{m_\Theta(\{Q_1\})} = \frac{m_A(\{A\})}{m_A(\{\neg A\})}, \quad \frac{m_\Theta(\{Q_2\})}{m_\Theta(\{Q_1\})} = \frac{m_B(\{\neg B\})}{m_B(\{B\})}, \quad \frac{m_\Theta(\{Q_3\})}{m_\Theta(\{Q_1\})} = \frac{m_C(\{\neg C\})}{m_C(\{C\})}.
\]

In general, we have, for singleton states \( Q \) and \( Q' \) in \( \Theta \):

\[
\log \frac{m_\Theta(Q|z)}{m_\Theta(Q'|z)} = \sum_Y \log \frac{m_Y(\omega_Y^{-1}(\{Q\})|z)}{m_Y(\omega_Y^{-1}(\{Q'\})|z)}
\]

We next combine the \( \infty \)-norm filtering formalism developed so far with the Dempster-Shafer Theory of Evidence to produce the Qualitative filter. A straightforward interface is not possible as we seek the maximum mass assignment according to the \( \infty \)-norm independent opinion pool formalism developed in the previous section. To address this problem, we develop the \( \infty \)-norm Dempster-Shafer Theory of Evidential Reasoning.
3.3 The \( \infty \)-norm Filter Cycle

In accordance with most quantitative filters, the Qualitative filter consists of a prediction stage and an update stage: states are inferred and assigned mass during the prediction stage. Mass estimates from new observations are transferred using the observation transfer function (i.e. FSM) and fused with the predicted state mass values in the update stage using the \( \infty \)-norm Dempster rule of combination. Conceptually, the \( p \)-norm generalisation of the Dempster-Shafer theory is no different to the original Dempster-Shafer theory. The \( p \)-norm formalism generates an infinite class of theories of which the Dempster-Shafer theory is a special case. For any mass assignment function \( m \) on the frame of discernment \( \Theta \), we have\(^{16}\):

\[
\forall A \subseteq \Theta \quad m_p(A) \geq 0 \quad m_p(\emptyset) = 0 \quad \left[ \sum_{A \subseteq \Theta} m_p(A)^p \right]^{\frac{1}{p}} = 1.
\]

with the following belief and plausibility functions:

\[
Bel_p(A) = \left[ \sum_{B \subseteq A} m_p(B)^p \right]^{\frac{1}{p}} \quad Pl_p(A) = \left[ \sum_{B \cap A \neq \emptyset} m_p(B)^p \right]^{\frac{1}{p}}
\]

and the following Dempster Rule of Combination:

\[
m_p(A|z_x, z_y) = k \left[ \sum_{A \cap X \cap Y} m_{\infty}(X|z_x)^{p} m_{\infty}(Y|z_y)^{p} \right]^{\frac{1}{p}}
\]

As can be seen, the original Dempster-Shafer theory corresponds to \( p = 1 \). In the limit \( p \to \infty \) the Dempster combination rule is an \( \infty \)-norm information pool, a requirement from Section 3.2. Thus, the \( \infty \)-norm Dempster-Shafer theory is appropriate for our needs:

\[
\forall A \subseteq \Theta \quad m_{\infty}(A) \geq 0 \quad Bel_{\infty}(A) = \max_{B \subseteq A} m_{\infty}(B) \quad m_{\infty}(A|z_x, z_y) = k \max_{A \subseteq X \cap Y} m_{\infty}(X|z_x)m_{\infty}(Y|z_y)
\]

\[
m_{\infty}(\emptyset) = 0 \quad Pl_{\infty}(A) = \max_{B \cap A \neq \emptyset} m_{\infty}(B)
\]

\[
\max_{A \subseteq \Theta} m_{\infty}(A) = 1.
\]

In the 1-norm formalism all evidence pointing towards a proposition contributes to the belief in that proposition whereas, in the \( \infty \)-norm formalism only the least corrigible evidence contributes.

The update stage of the filtering cycle is captured by the \( \infty \)-norm Dempster rule of combination. For the prediction stage of the cycle the observation and process model transfer functions can be clearly seen to obey, for some possible future state \( X \):

\[
m(X|z_{t_0}^{t}) = \max_{X=R(Q)} m(Q|z_{t_0}^{t})
\]

where \( R(Q) \) are the successor states of \( Q \). Figure 6 depicts the filter cycle diagrammatically.

In the next section we describe how, given a finite set of observations, the qualitative state which constrains the true state \( x \) can be identified.
3.4 Meta-Reasoning and the Decision Rule

Confidence measures in the log likelihood ratio are required to indicate when a rational decision can be made. As described in Section 3.2, we can decide \( x \in Q \) for a finite set of observations whenever:

\[
\exists Q : x \in Q, \forall Q', \forall' . \quad E \left( \log \frac{m_{\psi}(Q|z_0^t)}{m_{\psi'}(Q'|z_0^t)} \right) \geq 0.
\]  

(6)

In quantitative data fusion systems, covariance is a natural measure of confidence and, analogously, we can estimate the expected value in Equation (6) from the empirically determined \( \log \frac{m_{\psi}(Q|z_0^t)}{m_{\psi'}(Q'|z_0^t)} \) and the variance of this value. If \( Q = \arg\max_Q m_{\psi}(Q|z_0^t) \) and we choose to make the decision \( x \in Q \) when \(^{10} \):

\[
\forall Q' \neq Q, \forall' . \quad \log \frac{m_{\psi}(Q|z_0^t)}{m_{\psi'}(Q'|z_0^t)} > G \sqrt{\text{Var} \left( \log \frac{m_{\psi}(Q|z_0^t)}{m_{\psi'}(Q'|z_0^t)} \right)}
\]  

(7)

then, we can show using Chebyshev’s inequality that:

\[
\Pr \left( E \left( \log \frac{m_{\psi}(Q|z_0^t)}{m_{\psi'}(Q'|z_0^t)} \right) \geq 0 \right) \geq 1 - \frac{1}{G^2}
\]

Thus, we can find the qualitative region \( Q \) such that \( x \in Q \) to arbitrary probability of correctness \( 1 - \frac{1}{G^2} \).

The variance \( \text{Var} \left( \log \frac{m_{\psi}(Q|z_0^t)}{m_{\psi'}(Q'|z_0^t)} \right) \) for observations \( z_0^t \) is calculated recursively from the covariances of the mass assigned to subsets of \( z_0^t \). It is straightforward to show that the covariance of the log likelihood ratio for two observations \( z_1 \) and \( z_2 \) for binary coarse frames is directly related to the covariance of the observations themselves:

\[
\text{Cov} \left( \log \frac{m_{\text{mean}}(Q|z_1)}{m_{\text{mean}}(\neg Q|z_1)}, \log \frac{m_{\text{mean}}(Q|z_2)}{m_{\text{mean}}(\neg Q|z_2)} \right) = \text{Cov}(z_1, z_2).
\]

The variance of the log likelihood ratio in the refined frame \( \Theta \) can be obtained using Equation (5). For two regions \( Q_i \) and \( Q_{i+n} \) in the refined frame \( \Theta \) (see Figure 5) we have:

\[
\text{Var} \left( \log \frac{m_{\psi}(Q_i|z)}{m_{\psi}(Q_{i+n}|z)} \right) = n \text{Var} \left( \log \frac{m_{\psi}(Q|z)}{m_{\psi}(\neg Q|z)} \right).
\]  

(8)

Before continuing to develop the recursive covariance update rule, we introduce some notation for readability sake:

\[
M_0^t = \log \frac{m_{\psi}(Q|z_0^t)}{m_{\psi'}(Q'|z_0^t)} \quad V_0^t = \text{Var} \left( \log \frac{m_{\psi}(Q|z_0^t)}{m_{\psi'}(Q'|z_0^t)} \right) \quad C = \text{Cov} \left( \log \frac{m_{\psi}(Q|z_0^t)}{m_{\psi'}(Q'|z_0^t)}, \log \frac{m_{\psi}(Q|z_0^{t-1})}{m_{\psi'}(Q'|z_0^{t-1})} \right)
\]

\[
M_0^{t-1} = \log \frac{m_{\psi}(Q|z_0^{t-1})}{m_{\psi'}(Q'|z_0^{t-1})} \quad V_0^{t-1} = \text{Var} \left( \log \frac{m_{\psi}(Q|z_0^{t-1})}{m_{\psi'}(Q'|z_0^{t-1})} \right)
\]

\[
M_t = \log \frac{m_{\psi}(Q|z_t)}{m_{\psi'}(Q'|z_t)} \quad V_t = \text{Var} \left( \log \frac{m_{\psi}(Q|z_t)}{m_{\psi'}(Q'|z_t)} \right)
\]

By Equation (7), we make a decision whenever \( \frac{(M_0^t)^2}{V_0^t} \geq G^2 \). As mentioned in Section 3.2, the rate of convergence of our estimates decreases the closer the true state is to the landmark. Therefore, it is important to maintain minimal uncertainty in the estimate. Here, we use the flexibility of our observation weights \( K_1 \) and \( K_2 \) in Equation (2) for two observation sets to minimise \( \frac{V_t^0}{M_0^{t-1}} \). Without loss of generality, we can assign \( K_1 = W - K \) and \( K_2 = K \) in Equation (2) with \( W \geq K \geq 0 \) so that \( M_0^t = (W - K)M_0^{t-1} + KM_t \). Unfortunately, the value of \( K \) corresponding to the critical (minimum) value of \( \frac{(M_0^t)^2}{V_0^t} \) is stochastic which violates our conditions for unbiased
estimation. Alternatively, we can maximise \( \frac{E[M^t_0]}{V_0} \) and obtain \( K = \frac{W E[M^t_0]}{sV_0^0 + V_1 - C[1 + S]} \), where the zero-noise decision strength \( S \) between observation \( z_1 \) and observations \( z_0^t \) is defined to be \( S = \frac{E(M^t_0)}{E(M^t_0)} \). Thus, we have 16:

\[
(V_0^t)_{\text{mmse}} = W^2 \frac{(V_0^t - C^0)(V_1 + S^2 V_0^t - 2CS)}{(V_0^t - 1 + V_1 - C[1 + S]^2) - V_0^t}
\]

and:

\[
(M_0^t)_{\text{mmse}} = W \frac{SM_t V_0^t - M_t^t V_1 - C[M_t + S M_0^t]}{V_0^t - 1 + V_1 - C[1 + S]}
\]

Analogously to the Kalman filter, evidence is weighted inversely to its variance (i.e. consider \( S = 1 \) and \( C = 0 \)). The zero-noise decision strength \( S \) reflects the fact that observations of a true state lying further from a landmark produce a faster converging estimate and, therefore, should have more weight. The filter is optimal when \( S = \frac{E(M_t)}{E(M^t_0)} \), but, since the expected values of the masses are not known apriori, obtaining optimal values for \( S \) is problematic. This is an area of ongoing research.

In the next section we demonstrate the qualitative estimator in operation via an example drawn from the robot sensing domain introduced in Section 2.

### 4 The Robot Sensing Domain Revisited

Returning to the sensing domain introduced in Section 2, we implement a MMSE mean estimator as a Dempster-Shafer network and demonstrate this with an application in the domain of mobile robot feature recognition which was described in Section 2.

A network methodology provides a way of breaking the evidence that bears on a large problem into smaller items of evidence that bear on smaller parts of the problem so that these smaller problems can be dealt with one at a time. As such, this has applications in multi-sensor fusion when the problem state is too large to maintain, or when information is not available to maintain a complete state model or when decentralised (i.e. sensor parallel) computations are more appropriate in tasks which allow localisation of effort. The basic idea of local computation for propagating probabilities is due to Judea Pearl. Shenoy 19 extends this idea to belief functions in Belief Networks. In this section, we show that the Qualitative filter can be implemented as a belief network, thus demonstrating its applicability to large systems.

In the following experiment data from sparsely calibrated sensors is used to determine the curvature of some feature in the environment using the qualitative models described in Section 2. The calibrated readings (i.e. the landmarks) and the quantity-spaces for this domain are shown in Figure 7 4 and the belief network is shown in Figure 8 (information flow is in the directions indicated). Estimates for \( Qdir(\phi) \) and \( Qmag(\dot{R} \tan \theta) \) are maintained using the MMSE mean estimator (Equations (9) and (10)). The sonar mass estimates for \( Qmag(\dot{R} \tan \theta) \) are propagated to the \( \frac{1}{\mu_{R^m}} \) node and combined with the estimate for \( Qdir(\phi) \) from the gyroscope thus:

\[
m_{\frac{1}{\mu_{R^m}}}(A) = \max_{Qmult(A,B) = Qminus(B)} \{m_{\phi}(B)m_{\dot{R} \tan \theta}(C)\}.
\]

\footnote{In order to filter to landmark values, each landmark is “blurred” and represented as a narrow interval on the real line (e.g. \( Qmag(\phi) = 0 \Rightarrow \phi \in (-0.01, 0.01) \)).}
The sonar+gyroscope estimate for \( Q_{\text{mag}} \left( \frac{1}{r \frac{\partial R}{\partial \phi}} \right) \) and the laser curvature estimate \( Q_{\text{mag}} \left( \frac{\partial^2 R}{\partial \phi^2} \right) \) are then passed to the \( r \) node which fuses them according to Table 2 in Section 2.

The data obtained from the sensors, while observing a convex surface, is shown in Figure 9. The following observation errors (standard deviations) were assumed for the sonar sensor and gyroscope: 0.01 metres for \( R \), 0.02 rads for \( \theta \) and 0.1 rads for \( \delta \phi \). Figure 10 shows the second standard deviation limits for the expected log likelihood for the various curvature types discerned by the sonar+gyroscope combination. The decision that the surface is convex or far-concave can be made with 75% confidence after 3 observations have been received from each sensor at which point the following mass values were obtained for \( Q_{\text{mag}} \left( \frac{1}{r \frac{\partial R}{\partial \phi}} \right) \):

\[
\log \frac{m_g \{ \text{convex, far-concave} \}}{m_g \{ \text{planar} \}} = 0.07 \pm 0.02 \\
\log \frac{m_g \{ \text{convex, far-concave} \}}{m_g \{ \text{close-concave} \}} = 0.11 \pm 0.02.
\]

To determine surface curvature from the laser data we identify two subsets of data points \( S_1 \) and \( S_2 \), (shown as shaded regions in Figure 9) such that the data in \( S_1 \) have greater range values that those in \( S_2 \). We then evaluate the mean bearing distance \( l \) between points within the same section and the variance of this distance. We can then determine the curvature from

\[
n(l_1 - l_2) = \begin{cases} Q_{\text{mag}}(E(l_1 - l_2)) > 0 & \iff Q_{\text{mag}}( \frac{\partial^2 R}{\partial \phi^2} ) = + \\ < 0 & \iff Q_{\text{mag}}( \frac{\partial^2 R}{\partial \phi^2} ) = - \end{cases}
\]

We assume that the laser sensor range information is accurate to 0.05\( m \) and the bearing inaccuracy is comparatively negligible \( ^{^{^{15}}} \). Thus, from a single scan, the following estimates for \( Q_{\text{mag}} \left( \frac{\partial^2 R}{\partial \phi^2} \right) \) were obtained from the laser sensor data:

\[
\log \frac{m_l \{ \text{close-concave, convex, planar} \}}{m_l \{ \text{far-concave} \}} = \log \frac{m_{\partial^2 R}}{m_l \{ \} \{ + \}} = 0.35 \pm 0.05.
\]

Combining laser and sonar estimates using Equations (9) and (10) (with \( W = 2 \) and \( K = 1 \)) we obtain:

\[
\log \frac{m \{ \text{convex} \}}{m \{ \text{close-concave} \}} = \log \frac{m_{\partial^2 R}}{m_l \{ \text{close-concave} \}} + \log \frac{m_l \{ \text{close-concave, convex, planar} \}}{m_l \{ \text{convex, far-concave} \}} = 0.07 \pm 0.02
\]

\[
\log \frac{m \{ \text{convex} \}}{m \{ \text{far-concave} \}} = \log \frac{m_{\partial^2 R}}{m_l \{ \text{convex, far-concave} \}} + \log \frac{m_l \{ \text{close-concave, convex, planar} \}}{m_l \{ \text{far-concave} \}} = 0.35 \pm 0.05
\]

\[
\log \frac{m \{ \text{convex} \}}{m \{ \text{planar} \}} = \log \frac{m_{\partial^2 R}}{m_l \{ \text{planar} \}} + \log \frac{m_l \{ \text{close-concave, convex, planar} \}}{m_l \{ \text{far-concave} \}} = 0.11 \pm 0.02
\]

from which the surface is correctly inferred to be convex.

5 CONCLUSIONS

In this paper, we have addressed the problem of data fusion and parameter estimation in systems which use qualitative models. We maintain that the more sensors available the better the estimate both for overcoming
qualitative model ambiguity and for filtering noise in the data.

We have developed a minimum-mean-square-error decision framework and we have explored four levels of (meta-) reasoning about uncertainty in this framework: the transformation of observation to mass via the mass assignment function, the uncertainty (variance) in this mass due to the random nature of the observations, the relationship of variance to mass for decision purposes and lower bounds on the reliability of the decision by Chebyshev's inequality. Thus, we overcome the problem of dependent belief functions cited by, amongst others, Liu and Murphy. 

Our framework supports mixed estimator type fusion and can be used in situations when the mean of one sensor output corresponds with the median of the output from a different sensor. We show elsewhere that our approach to estimation can be extended to stochastic process models. Our framework is amenable to network implementation and this has been demonstrated in the robot sensor fusion domain.
REFERENCES


