

Sensor fusion

Antanas Verikas[†]

[†]Intelligent Systems Lab, Halmstad University,
Box 823, 30118 Halmstad, Sweden
antanas.verikas@hh.se

2009

Advantages of multi-sensor systems

- Higher signal-to-noise ratio.
- Increased reliability in the case of sensor failure.
- Reduced uncertainty and increased confidence.
- Increased discrimination power due to more comprehensive information arriving from multiple sensors.
- Improved resolution due to multiple sensors.

Difficulties

- Data in diverse formats, noisy and ambiguous
 - analogue, digital, discrete, textual, imagery
- Data dimensionality and alignment
 - Different coordinate systems, units, frequency, amplitude, timing
- Temporal alignment
 - Synchronisation of data
 - Spatial distribution of sensors demands precise time measurements
 - Data arrival at fusion node may not coincide due to variable propagation delays

Perspectives

Sensor fusion concepts can be characterized from several perspectives:

- Application domain
- Fusion objective
- Sensor type
- Sensor suite configuration
- Fusion level

Application domain

- Defense
- Robotics
- Medical
- Space
- Entertainment

Fusion objective

- Tracking of an object
- Detecting the presence of an object
- Assessing environmental conditions
- Recognizing an object

Sensor type

The major factors dictating the selection of sensors for fusion are:

- the compatibility of the sensors for deployment within the same environment
- the complementary nature of the information derived from the sensors

Sensor suite configuration

The following configurations are by far the most common:

- Parallel Configurations
- Serial Configurations
- Varied combinations of the parallel and serial configurations

The parallel configuration has been the most studied wherein the sensors are making observations concurrently.

Fusion level

- Data fusion
- Feature fusion
- decision fusion

Table: Sensor fusion in terms of input/output provided.

Mode	Example
Data in \rightarrow data out	Fusion of multi-spectral data.
Feature in \rightarrow feature out	Fusion of image and non-image data.
Decision in \rightarrow decision out	Case of non compatible sensors.
Data in \rightarrow feature out	Shape extraction.
Feature in \rightarrow decision out	Object recognition.
Data in \rightarrow decision out	Pattern recognition.

Techniques

- Artificial neural networks
- Fuzzy set theory
- Neuro-fuzzy systems
- Kernel methods (support vector machines)
- Probabilistic methods
 - Bayesian inference
 - Dempster-Shafer theory

Bayesian inference (1)

Suppose H_1, H_2, \dots, H_M are mutually exclusive and exhaustive hypothesis that explain the observed data S . Then

$$p(H_j|S) = \frac{p(S|H_j)p(H_j)}{\sum_i p(S|H_i)p(H_i)} \quad (1)$$

$$\sum_i p(H_i) = 1 \quad (2)$$

where $p(H_i)$ is the a priori probability of the hypothesis being true, $p(S|H_i)$ is the probability of getting the data S , given the H_i is true, and $p(H_j|S)$ is the a posteriori H_j probability given the data S .

Bayesian inference (2)

For data generated by multiple sensors S_1, \dots, S_n

$$p(H_j | S_1 \cap \dots \cap S_n) = \frac{p(H_j)[p(S_1|H_j)p(S_2|H_j)\dots p(S_n|H_j)]}{\sum_i p(H_i)[p(S_1|H_i)p(S_2|H_i)\dots p(S_n|H_i)]} \quad (3)$$

The conditional probability $p(S_i|H_j)$, the probability of observing the signal S_i , given that H_j is true, are usually determined based on experimental results.

Bayesian inference (3)

- Provides the probability of a hypothesis being true, given the evidence.
- Allows incorporating a priory knowledge of the likelihood of a hypothesis being true at all.
- The decision is usually based on the *maximum a posteriori* principle.

Dempster-Shafer theory (1)

Let $\Theta = \{\theta_1, \theta_2, \dots, \theta_M\}$ be a finite set of mutually exclusive and exhaustive atomic hypotheses about some problem domain.

- The probability theory assigns *probabilities* to atomic hypotheses θ_i , while in DS theory, a *basic probability number* $m(A)$ represents one's *belief* in a not necessarily atomic hypothesis A .
- For a compound hypothesis $A \neq \theta_i$, $m(A)$ measures our belief that we are willing to commit to A . The belief cannot be subdivided amongst the subsets of A and is assigned to A at the expense of support $m(\theta_i)$.
- The belief in A and the belief in its negation \bar{A} do not necessarily sum to 1.

Dempster-Shafer theory (2)

To obtain the total belief in A , ($\text{Bel}(A)$, where $\text{Bel} : 2^\Theta \rightarrow [0, 1]$), we must add the basic probability numbers $m(B)$ for all subsets B of A .

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B) \quad (4)$$

- The subsets B of Θ for which $m(B) > 0$ are called the *focal elements* of the belief function.
- The union of the focal elements is called the *core* of the belief function.
- The belief functions having only one focal element in addition to Θ are called *simple support functions*.

Dempster-Shafer theory (3)

Given two basic probability assignments m_1 and m_2 associated with Bel_1 and Bel_2 induced by two different sources of information over the same Θ can be combined if their cores are not disjoint. The *orthogonal sum* ($m = m_1 \oplus m_2$, $m : 2^\Theta \rightarrow [0, 1]$) is a convenient way to perform such a combination:

$$m(\emptyset) = 0 \quad (5)$$

$$m(A) = K^{-1} \sum_{B \cap D = A} m_1(B)m_2(D), \quad A \neq \emptyset \quad (6)$$

$$K = \sum_{B \cap D \neq \emptyset} m_1(B)m_2(D) \quad (7)$$

Dempster-Shafer theory (4)

- The function m is a basic probability assignment.
- The core of Bel given by m equals the intersection of the cores of Bel_1 and Bel_2 .
- The combination rule is commutative and associative, and it may be generalized to combine multiple evidences.
- Maximum belief over the atomic hypothesis is the most commonly used approach to make decisions based on the DS fusion result.

Fuzzy set theory (1)

- The advantages of the fuzzy set theory lie in the variety of combination operators.
- Fusion by IF-THEN rules, for example

IF x_1 is A_1 AND ... AND x_n is A_n THEN class C_q with z^q (8)

where $A_i (i = 1, \dots, n)$ are fuzzy sets defined over the input variables x_i , C_q is a class label and z^q is a rule weight. Each fuzzy set is represented by a membership function, for example Gaussian:

$$\mu_i = \exp \left(- \frac{(x_i - c_i)^2}{\sigma_i^2} \right) \quad (9)$$

where c_i and σ_i are the center and the width of the Gaussian function, respectively.

Fuzzy set theory (2)

- Fuzzy integration is an example of a successful combination operator.
- The fuzzy integral is interpreted as searching for the maximal grade of agreement between the objective evidence and the expectation.
- A non additive fuzzy measure (FM) is used in the fuzzy, for example Choquet, integral.

Fuzzy set theory (3)

If we assume that $Z = \{z_1, z_2, \dots, z_L\}$ is a non-empty finite set of data from L sources/sensors and g is an FM on Z , the *discrete Choquet integral* of a function $h : Z \rightarrow \mathbb{R}^+$ with respect to g is defined as

$$C_g(h(z_1), \dots, h(z_L)) = \sum_{i=1}^L [h(z_i) - h(z_{i-1})]g(A_i) \quad (10)$$

where indices i have been permuted so that $0 \leq h(z_1) \leq \dots \leq h(z_L) \leq 1$; $A_i = \{z_i, \dots, z_L\}$; $h(z_0) = 0$.

- If decision fusion is considered, $h(z_i)$ stands for the classification result of the i th classifier.

Fuzzy set theory (4)

A set function $g : 2^Z \rightarrow [0, 1]$ is a *fuzzy measure* if

- 1) $g(\emptyset) = 0$; $g(Z) = 1$;
- 2) if $A, B \subset 2^Z$ and $A \subset B$ then $g(A) \leq g(B)$;
- 3) if $A_n \subset 2^Z$ for $1 \leq n < \infty$ and $\{A_n\}$ is monotonic in the sense of inclusion, then $\lim_{n \rightarrow \infty} g(A_n) = g(\lim_{n \rightarrow \infty} A_n)$.

Fuzzy set theory (5)

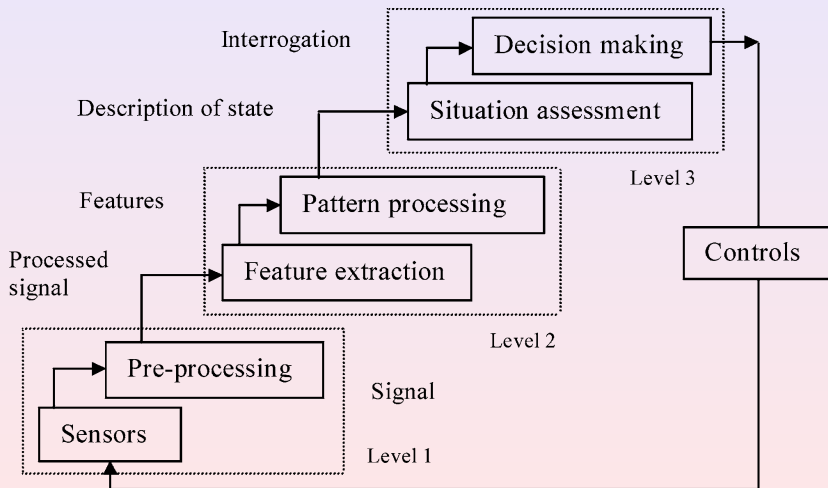
An ordinary FM of a union of two disjoint subsets cannot be directly computed from the FMs of the subsets. The so called λ -FM allows such computation:

$$g(A \cup B) = g(A) + g(B) + \lambda g(A)g(B) \quad (11)$$

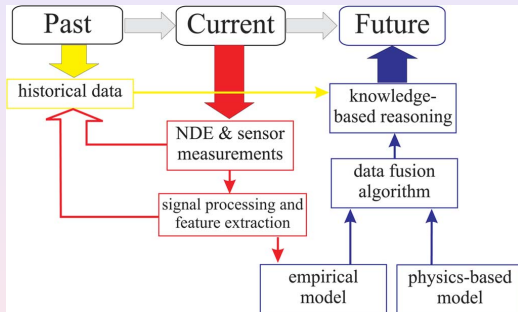
When g is the λ -FM, the values of $g(A_i)$ can be computed recursively. If $g^i = g(\{z_i\})$, then

- $g(A_1) = g(\{z_1\}) = g^1$
- $g(A_i) = g^i + g(A_{i-1}) + \lambda g^i g(A_{i-1})$, for $1 < i \leq L$.

A waterfall fusion model



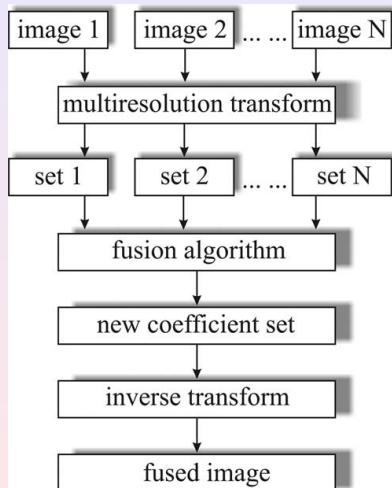
Diagnostic-prognostic monitoring



- The prediction of future is based on the knowledge of current conditions, the physical model of a process, and past events. (NDE=nondestructive evaluation).

- At the signal level, the fusion of redundant information enhances the reliability of the system.
- At a higher level, the fusion of physical and numerical models provides an input to the diagnostic and prognostic reasoners.

Multiresolution image fusion (1)



Two aspects of the fusion scheme:

- Multiresolution algorithms for image decomposition and reconstruction.
- The combination rule for the coefficient sets in the transform domain.

Multiresolution image fusion (2)

- The input image is decomposed into a set of spatial frequency bandpass sub-images.
- An image $I(x, y)$ can be represented as a sum over a collection of basic functions $g_i(x, y)$

$$I(x, y) = \sum_i y_i(x, y)g_i(x, y) \quad (12)$$

where i is the resolution level and $y_i(x, y)$ are the transform coefficients.

- The coefficients can be computed by projecting the image onto a set of projection functions $h_i(x, y)$:

$$y_i(x, y) = \sum_{x, y} h_i(x, y)I(x, y) \quad (13)$$

- The functions $h_i(x, y)$ are the translated and dilated copies of one another.