Πολυτεχνείο Κρήτης

Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών

Fusion techniques pros and cons

Data fusion: uses only raw data.

Information fusion: uses preprocessed data.

Goal of fusion: obtain a lower detection error probability and a higher reliability by using data from multiple distributed sources.

Three non-exclusive techniques:

1) Data association.
2) State estimation.
3) Decision fusion.

Classification categories for data fusion:

1) Classification based on the relations between the Data sources. The classification criteria could be complementary (the same image from different point of view), redundant (same images, but from different sources to increase confidence) and cooperative (combined information to create new, for example audio and visual information).

2) Dasarthy's classification. There are five categories:
   a) Data in-data out.
   b) Data in-feature out.
   c) Feature in-feature out.
   d) Feature in-decision out.
   e) Decision in-decision out.

3) Classification based on Abstraction level. We have four levels of abstraction:
   a) Signal level (provided from the sensors).
   b) Pixel level (could be used to improve image processing tasks).
   c) Characteristic (employs features that are extracted from the images or signals).
   d) Symbol or decision (the image is represented as symbols).

   The Information fusion addresses three levels of abstraction:

a) Low level fusion (raw data as input to the data fusion process. Lower signal to noise ratio, more accurate data than individual sources).
b) Medium level fusion (characteristics or features are fused to obtain features that could be employed for other tasks).
c) High level (takes symbolic representations as sources and combines them to obtain decision, for example Bayesian methods).
d) Multiple level fusion (combines some or all the above levels of abstraction).

4) JDL data fusion classification. Modification of the five level of abstraction.
   a) Level 0: preprocessing phase (includes fusion in signal and pixel level). It reduces the amount of data and maintains useful information for the higher levels.
   b) Level 1: object refinement (employs the processed data from level 0. Output of this stage is object discrimination and tracking. Input information turns into consistent data structures).
   c) Level 2: situation assessment (aims to identify the likely situations given the observed events and data. Also, aims to perform high level inference and identify patterns. Output a set of high-level inference).
   d) Level 3: impact assessment (evaluation of the impact of the activities in previous level. It evaluates risk or threat and includes a prediction of the logical outcome).
   e) Level 4: process refinement (improvement of previous levels and resource management).

   Limitations of the JDL: how the uncertainty of previous or subsequent results could be employed to enhance fusion process.

   In general, the Dasarthy model provides a method for understanding relations between fusion tasks and employed data, whereas the JDL present a fusion perspective to design the system.

5) Classification based on Architecture. Designing fusion systems, create a question: where to the data fusion process will be performed?
   a) Centralized architecture (All the fusion processes take place in a central processor that uses raw data. In theory optimal, in practice requires large bandwidth to send raw data, data alignment and association aren't always performed correctly).
   b) Decentralized (consists of network of nodes, each one with its own process unit. Fusion is performed autonomously, with local information and from peer nodes. Main disadvantage the communication cost and lack of scalability).
   c) Distributed (measurements from each node are processed independently before the fusion. Fusion node is responsible for the information received from the other nodes).
   d) Hierarchical (combination of decentralized and distributed).

Decentralized data fusion is hard to implement and require huge computation power. But that does not mean that distributed systems are optimal. It depends on the requirements.

Data association techniques:

Goal is to establish a set of observations or measurements that are generated by the same target over time. The data association problem is NP complete (exhaustive search complexity grows exponentially as the number of targets grow), therefore various techniques are presented to solve the problem:

1) Nearest Neighbor and K means. NN is the most known clustering algorithm (nothing to analyze here). Bad performance in environments with frequent false measurements or with highly noisy environments. K means, a modification of the NN. The algorithm, does not always find the optimal solution for the cluster centers, the number of clusters must be known a priori and we have to assume that the number is the optimum and finally, the algorithm assumes that the covariance of the dataset is irrelevant or that it has been normalized. The positive is that there are tools to overcome those limitations.

2) Probabilistic data association (PDA). Also known as modified filter of all neighbors. Main disadvantages are loss of tracks (PDA ignores interference with other targets, wrong classification when targets are close or crossed), the suboptimal Bayesian approximation (when the source of information is uncertain, PDA is the suboptimal Bayesian approximation to the association problem), one target (PDA behaves incorrectly when there are multiple targets) and track management (assumes the track is already established). PDA is good for tracking targets that do not make abrupt changes in their movement patterns.

3) Joint probabilistic Data association. A suboptimal approach for multiple targets. Computationally forbidden and requires a mechanism for track initialization.

4) Multiple Hypothesis Test. Use more than two consecutive observations to make an association with better results. It is developed for multiple targets, therefore it combines data association and tracking into a unified framework. Again, huge computational cost.

5) Distributed Joint Probabilistic Data Association. Various limitations in practical applications

6) Distributed Multiple Hypothesis Test. Again, huge computational requirements.

7) Graphical Models. A formalism for representing reasoning with probabilities and independence. GM can solve the distributed data association problem in synchronized sensors with overlapped areas and where each sensor receives noisy

measurements. The data association problem is treated as an inference problem and solved by using the max-product algorithm.

State Estimation Methods (tracking):

Aims to obtain the state (usually the position) given the measurements. The estimation problem involves finding the values of the vector state that fits as much as possible, with the observed data.

1) Maximum Likelihood and Maximum Posterior. Both, aim to find the most likely value for the state. However, ML assumes that the state is a fixed but unknown point, whereas the MAP considers the state to be the output of a random variable. Both perform the same when no a priori information exists, meaning there are only observations.
2) Kalman Filter. The most famous technique. It estimates the state of a discrete time process, governed by a specific space-time model. Optimal estimations when system is linear and error can be modeled by Gaussian noise. For non-linear systems, some modifications are required.
3) Particle Filter. Recursive implementations of sequential Monte Carlo method. It is more flexible than the Kalman and can be cope with non-linear dependencies and non-Gaussian densities in the model and error respectively. However, a large number of particles needed to obtain a small variance in the estimator. The number of particles, also affect the computational cost.
4) Distributed Kalman Filter. It requires synchronization of all the sources. If the estimations are consistent and the cross covariance is known, then it is possible to use the DKF.
5) Distributed Particle Filter. Can be used to monitor an environment that could be captured by the Markovian state-space model, involving nonlinear dynamics and non-Gaussian noise. Better results compare to the simple Particle Filter, but requires huge amount of space to store the state of the particles instant, each time.
6) Covariance Consistency Methods (covariance union/intersection). They are fault tolerant for maintaining covariance means and estimation in a distributed network. The intersection method guarantees consistency and nondivergence for every sequence of mean and covariance-consistent estimations. Not suitable for non-consistent measurements. The covariance union was proposed to solve the problem of the inconsistent inputs. Big advantage of the CU algorithm is scalability.

Decision Fusion Methods:

1) Bayesian Methods. Bayesian inference is based on the Bayes rule. Problems that occur are the difficulty in establishing the value of a priori probabilities, the complexity when there are multiple hypothesis, the fact that hypothesis should be mutually exclusive and the difficulty describing the uncertainty of the decisions.
2) Dempster-Shafer Inference. A generalization of Bayesian theory, provides a formalism that could be used to represent incomplete knowledge. No a priori probabilities are required.

3) Abductive Reasoning or inferring the best explanation, is a reasoning method in which a hypothesis is chosen under the assumption it is true and explains the observation more accurate than the others. It finds the posterior ML of the system variables given some observed variables. It is more of a reasoning pattern, rather than fusion technique, but therefore, different inference methods can be employed (Nearest neighbor).
4) Semantic Methods. They employ semantic data from different sources as input. More accurate than single source systems. Can be used to translate raw data into formal language, then the resulting languages gets compared to similar languages that are stored in a database. That method provides saving in terms of transmission cost, because no raw data gets transmitted.

Conclusions:

Various techniques can be optimal under specific conditions.

Classification based on Abstraction level.	
Low level fusion	Lower signal to noise ratio, more accurate data than individual sources
Medium level	Characteristics or features are fused to obtain features that could be employed for other tasks
High level	Takes symbolic representations as sources and combines them to obtain decision
Multiple level	Combines some or all of the above

JDL data fusion classification	
Level 0	Reduces the amount of data and maintains useful information for the higher levels
Level 1	Input information turns into consistent data structures
Level 2	Aims to perform high level inference and identify patterns. Output a set of high-level inference
Level 3	It evaluates risk or threat and includes a prediction of the logical outcome
Level 4	Improvement of previous levels and resource management

Classification based on Architecture	
Centralized architecture	In theory optimal, in practice requires large bandwidth to send raw data, data alignment and association aren't always performed correctly
Decentralized	Fusion is performed autonomously, with local information and from peer nodes. Main disadvantage the communication cost and lack of scalability. Decentralized data fusion is hard to implement and require huge computation power
Distributed	Measurements from each node are processed independently before the fusion. Fusion node is responsible for the information received from the other nodes.
Hierarchical	Combination of decentralized and distributed

Limitations of the JDL: how the uncertainty of previous or subsequent results could be employed to enhance fusion process. In general, the Dasarthy model provides a method for understanding relations between fusion tasks and employed data, whereas the JDL present a fusion perspective to design the system.

Data Association techniques.	
Nearest Neighbor and K means.	Bad performance in environments with frequent false measurements or with highly noisy environments.
Probabilistic data association (PDA).	Ignores interference with other targets, wrong classification when targets are close or crossed. Behaves incorrectly when there are multiple targets. Assumes the track is already established. Good for not abrupt changes in patterns.
Joint probabilistic Data association.	Good approach for multiple targets. Huge computational cost.

Multiple Hypothesis Test.	Combines data association and tracking into a unified framework. Huge computational cost.
Distributed Joint Probabilistic Data Association.	Various limitations in practical applications.
Distributed Multiple Hypothesis Test.	Huge computational cost.
Graphical Models	Can solve the distributed data association problem in synchronized sensors with overlapped areas and where each sensor receives noisy measurements.

State Estimation Methods(tracking)	
ML and MAP.	Equivalent, when no a priori info provided.
Kalman Filter.	Optimal only when system is linear and noise Gaussian.
Particle Filter.	More flexible than Kalman. Suitable for nonlinear, and non-Gaussian noise, systems A large number of particles needed to obtain a small variance in the estimator. More particles, mean more computational cost.
Distributed Kalman Filter	Requires synchronization of all the sources. If the estimations are consistent and the cross covariance is known, then it is possible to use the DKF.
Distributed Particle Filter.	Better results compare to the simple Particle Filter, but requires huge amount of space to store the state of the particles instant, each time
Covariance Consistency Methods	They are fault tolerant for maintaining covariance means and estimation in a distributed network. Check details above.

Decision Fusion Methods	
Bayesian Methods	Difficulty in establishing the value of a priori probabilities, the complexity when there are multiple hypothesis, the fact that hypothesis should be mutually exclusive and the difficulty describing the uncertainty of the decisions.
Dempster-Shafer Inference	Could be used to represent incomplete knowledge. No a priori information needed.
Abductive Reasoning	It is more of a reasoning pattern, rather than fusion technique but, different inference methods can be employed.
Semantic Methods	Provides saving in terms of transmission cost, because no raw data gets transmitted.