

Multi-Sensor Data Fusion and Reliable Multi-Channel Computation: Unifying Concepts and Techniques

Behrooz Parhami

Department of Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106-9560, USA

Abstract

Multiple sensors are used for much the same reasons as redundant computation channels; viz to tolerate certain types of error, incompleteness, inaccuracy, and/or tardiness in the data supplied by individual system elements such as sensors or processors. Data fusion in multi-sensor systems is investigated by numerous signal processing researchers. The problem of "voting" or, more generally, integrating results from multiple identical or diverse computations has been studied by computer scientists/engineers interested in fault tolerance and distributed computing issues. With very few exceptions, these groups publish their results in disjoint conferences and journals. Thus, there is little interaction and cross-fertilization between the two efforts. We review common aspects of the two problems and show how a data-driven methodology may offer potentials for tackling both problems within a unified framework.

Keywords: Data diversity, Dependability, Redundancy, Reliable computation, Sensor fusion, Generalized voting.

1. Introduction

Sensor processing is needed in environmental monitoring, intelligent manufacturing, process control, military surveillance, medical imaging, robotics (e.g., handling of hazardous material), and remote sensing [IYEN95]. Sensors can range from very simple (e.g., metal or smoke detector, barometer) to highly complex (e.g., identifying obstacles in the path of a robot).

Data obtained from sensors may be erroneous, incomplete, inaccurate, or tardy for a variety of reasons. These can be classified roughly into four categories:

- (1) Inherent limitations of sensor capabilities,
- (2) Permanent or intermittent malfunctions,
- (3) Communications-related errors or delays, and
- (4) Anomalies or faults in storage/processing.

A good example of (1) is when one uses two sensors to obtain target position data: one with poor elevation and azimuth precision but with relatively accurate range (radar) and another with complementary characteristics (FLIR = forward looking infrared). Another example is when the coverage of a sensor (e.g., area sensed) is limited and multiple sensors are needed to provide the desired coverage.

Limitations are evaluated with respect to the level of certainty required. A particular sensor may be adequate for one application (uncritical situation or unsophisticated adversary) but deemed to be limited for another (battlefield, advanced jamming methods). As examples of (2), sensor unavailability may result from physical faults, jamming, or overload in the case of non-dedicated sensors. Both of the aspects (1) and (2) are usually handled through sensor replication while (3) and (4) are dealt with by means of information coding and redundant computations.

Sensor replication gives rise to the problem of data integration or fusion. Traditionally, researchers dealing with multi-sensor data fusion have assumed that the communication, storage, and processing subsystems are highly reliable and have focused only on algorithms for integrating data from homogeneous or heterogeneous collections of potentially faulty/inaccurate sensors.

On the other hand, researchers dealing with redundant or replicated computations have, for the most part, assumed that the input data is *perfect* and that the only sources of errors or inaccuracies are faults in the data communication, storage, and processing subsystems and perhaps the numerical characteristic of the algorithms being used (e.g., with regard to the accumulation of roundoff errors).

In this paper, following brief overviews of the two disciplines, we discuss the similarities between multi-sensor data fusion and multi-channel computation, point to the need for closer interaction between the two disciplines, and suggest that a data-centered or data-driven methodology can be used as a framework for dealing with both problems. The methodology fosters a unified treatment of data errors, inaccuracies, and tardiness regardless of the underlying cause (data generation/collection, transmission, storage, manipulation, or interpretation) and explicates the optimal allocation of resources for dealing with various error sources. It also allows the designer to view varied redundancy features or techniques, from data and design diversity to re-try or replication, as instances of a general data reliability enhancement process, thus facilitating comparisons and tradeoff studies.

2. Multi-Sensor Data Fusion

The sensor data fusion community, which has grown tremendously since the emergence of early rudimentary systems in the late 1970s, is fragmented and in search of a unified approach [HALL92]. One segment of this community is concerned with military applications and communicates primarily through the Tri-Service Data Fusion Conference and meetings on automated target recognition. Within this segment, the need to interrelate and understand diverse activities carried out under data fusion led to the formation of the Data Fusion Sub-Panel of the Joint Directors of Laboratories Technical Panel for C³ in 1984 [WALT90]. The non-military segment, on the other hand, has a more theoretical bent and communicates via journals on control or robotics and the SPIE Conference on Sensor Fusion, among others.

Research in data fusion is multidisciplinary and uses techniques from signal processing, statistics, pattern recognition, and information theory, among others. Hence, "data fusion is not a discipline in the same sense as ... signal processing or numerical methods. ... Well-defined techniques, terminology, and a professional community do not yet exist" [HALL92]. A variety of results/techniques from decision or detection theory (e.g., Bayes' method), estimation theory (least-squares, max-likelihood), association or correlation, and uncertainty management (evidence/belief theory, Shafer-Dempster reasoning, fuzzy calculus) provide the theoretical bases for system implementations [WALT90]. In view of the above, it is quite difficult to characterize multi-sensor data fusion in order to focus on the common underlying properties.

Fig. 1 represents an attempt to capture the main elements of a multi-sensor data fusion system. The multiple sensors used may be similar/competitive (e.g., the two eyes in a human) or diverse/complementary (e.g., visual, tactile, and auditory sensors). If the sensor data have no common feature, fusion is clearly impossible. At the other extreme, if all features of the multiple sensors are common, then fusion reduces to filtering out the variations in measurement [ROTH91]. In general, multi-sensor data fusion may be done in a hierarchical manner, with fused data from one level forming the raw data for the next level. However, in this paper, we will focus only on one level of integration as depicted in Fig. 1. Each of the elements in Fig. 1 is briefly described below.

Sensors have been classified based on their decision strategies into *hard-* and *soft-decision* sensors. The former process their incident signal data and utilize decision rules to declare target identity while the latter may provide partial evidence for identity upon detection of signals. Soft-decision sensors "accumulate and integrate evidence,

reporting partial evidence and associated uncertainty (via probabilities, fuzzy membership functions, confidence factors, or evidential intervals)" [HALL92]. Data provided by a sensor comes in varying forms (waveform, integer designating a class, numerical vector, image), and usually includes information about the sensor itself; current state (e.g., pointing angle), configuration, health, etc.

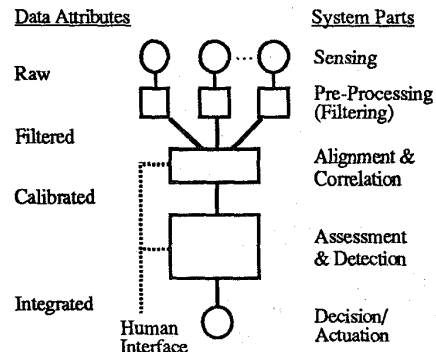


Fig. 1. Simplified view of multi-sensor data fusion.

Pre-processing or filtering is often needed to reduce the volume of data which may otherwise overwhelm the data processing part, but its presence is optional. The low-level processing stage, labeled "alignment & correlation", consolidates the data by using spatial and temporal references, unit conversions, pairwise association of observations (a quadratic-time process), and position/identity determination. The high-level processing stage, labeled "assessment & detection" in Fig. 2, interprets the output of the previous stage and makes inferences in the context of system structure and goals (e.g., threat assessment in a battlefield).

Finally, the output of a multi-sensor data fusion system is presented in the form of a decision which may be further processed externally, used to directly control some system components (e.g., adjust or reorient the sensors), or tied to other systems (e.g., warning or response units).

3. Multi-Channel Computation

Multi-channel computation is an established method of ensuring reliable results despite the occurrence of failures in computational elements. The method has its roots in an early paper by von Neumann [VONN56], with many extensions, enhancements, and implementation strategies reported since (see, e.g., [PARH94] for a survey of fault-tolerant computing and [PARH94a] for background on replication and voting). Results in this and other aspects of reliable computation are reported in the annual Symposia on Fault-Tolerant Computing [FTCSyr], now in the 25th year, and several other computer conferences. Of particular relevance to multi-sensor data fusion are the

symposia on (reliable) distributed systems in which topics such as tolerating site and link failures, and reaching consistent decisions in the presence of such failures, are prominently featured.

Research in multi-channel computation is highly focused, with the underlying theoretical framework and the needed terminology firmly in place. Early research in this area assumed identical, replicated computation channels, and was aimed primarily at the tolerance of operational hardware faults. This limited the role of the data combining unit (the *voter*) to bitwise comparison of data and the resolution of minor timing differences. More recently, with the higher emphasis placed on software-based redundancy methods and the introduction of diversity in multiple computation channels, in an attempt to deal with design flaws as well as operational faults, use of complex voters has become quite common. Such voters deal with higher-level data objects and thus require attention to differences in accuracy, partial errors, and incompleteness of data. The problems faced by such complex voters are almost identical to those in data fusion.

Fig. 2 captures the main elements of a reliable multi-channel computation scheme. The multiple processing channels used may be similar or diverse, with consistency checking and voting becoming more complex if they are diverse. Acceptance tests are optionally applied to computation results in order to filter out clearly incorrect results. This testing, if present, may then be followed by data compression to obtain a *signature* for each channel in an attempt to reduce the complexity of the subsequent phases. Consistency checking is again optional but is frequently found in modern voting schemes in an attempt to remove non-conforming results prior to combining them by methods such as averaging, e.g., in which a single input that is "way off" may greatly influence the output. The voting block is the heart of this system and is the element whose internal design and operation most distinguishes the particular scheme used. Thus, the study of voting methods has become a prime focus in this area.

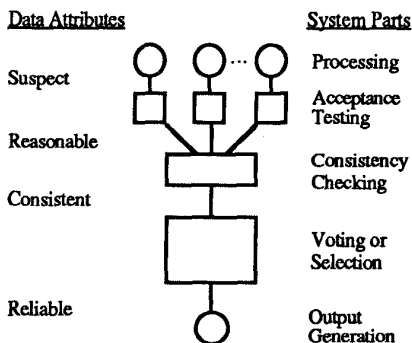


Fig. 2. Elements of multi-channel computation.

The output of a multi-channel computation scheme is frequently used as input for another computation phase. Thus, to prevent the consistency checker and voter from becoming the weak links whose failure would interrupt correct system operation, replication of these parts has been suggested. Such a system would yield the results in multi-channel format and would thus allow the cascading of several computations without the danger of catastrophic propagation of single-point errors.

4. A Unified View

Researchers in both multi-sensor data fusion and reliable multi-channel computation are becoming increasingly aware of the need for unifying theories. For example, it has been observed that "while the term data fusion is widely used, its meaning is subject to varying interpretations", but "data fusion has a common basis in theory, which is independent of application" [WALT90]. On the other side of the fence, the need for a generalized formulation of voting to cover the wide variety of methods in current use has been duly noted [LORC89]. It is our contention that we should go beyond separate unification on each side and aim instead for unifying theories that would remove the fence. A superficial glance at Figs. 1 and 2 (which are admittedly drawn to highlight the similarities) should be sufficient to convince the reader that in-depth comparison of such systems is worth the effort and may lead to advances in both areas. Fig. 3 is a merged version of Figs. 1 and 2, using more neutral terminology.

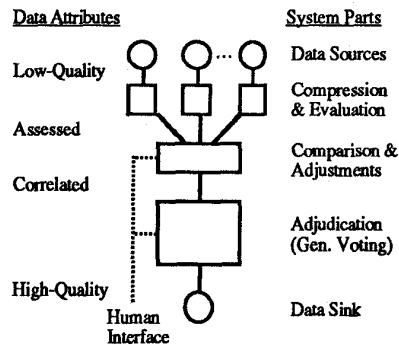


Fig. 3. Unified model of multi-sensor data fusion and reliable multi-channel computation.

Perhaps the most important system component in Fig. 3 from the point of view of unifying the two areas is the box labeled "adjudication". This term was coined by some researchers in fault-tolerant computing in order to avoid the restrictive meaning often associated with "voting". Others have preferred to use "generalized voting". The human interface, which is absent in Fig. 2 has been added to account for the fact that advanced voting techniques often involve adjustable parameters.

5. Example Applications

In this section, we take two example applications from the multi-sensor data fusion literature and formulate them in terms of generalized voting. Our goal is to show that some techniques from one area can be useful in the other. In addition, we discuss an example application where both techniques are applied in a complementary manner.

The first example is from p. 93 of [WALT90]. Two sensors produce ambiguity sets following their attempt to recognize the class of a target. Sensor 1 supplies the ambiguity set {4, 5, 12, 18} while the second sensor provides {12, 21, 32, 33}. Combined, the two sensors unambiguously identify the target as being in class 12. This type of fusion is actually a special case of approval voting [PARH92], [PARH94a]. Each of the multiple computation channels provides as output a set of approved values (e.g., system states that are deemed safe following a detected fault). The approved values, with each item or each set perhaps having an associated weight or confidence level, are combined through an approval voting algorithm to identify the best value or set of values.

The second example is somewhat more complex and has been the subject of extensive research in multi-sensor data fusion [MARZ90], [IYEN94], [IYER95]. Assume that multiple sensors provide real-valued scalar data. A non-faulty sensor S_i provides the real-valued output x_i . With knowledge of sensor accuracy, one can define an interval $[x_i - \Delta x_i, x_i + \Delta x_i]$ as containing the *correct* or intended sensor data. The objective is then to obtain a value or an interval of values that represent the best estimate of the sensed quantity. Again, the multiple intervals can be viewed as sets of values approved by the sensors, with simple or weighted voting used to fuse the data. Faulty sensors are properly handled as their intervals likely do not overlap with those of correct sensors (correct conclusion is reached with very high probability even if they do). It is interesting to note that many of the results published in the above references can be obtained directly and simply from the approval voting interpretation. Conversely, some of the bounds derived in these references on the width of the fused interval can be applied to analyze the precision and fault diagnosability in interval voting applications.

Consider now either of the above problems in a situation where the processing part of the sensor fusion system is also subject to unavailability or failure. This motivates distributed multi-channel processing of sensor data, with the outcome being multiple fusion results for use/interpretation by humans or another reliable multi-channel system. In this context, voting and fusion become indistinguishable, as it is difficult to differentiate between sensor failures/inaccuracies and computation errors.

6. A Data-Centered Approach

In this section, a data-centered or data-driven methodology is introduced that can be used as a framework for dealing with both multi-sensor data fusion and reliable multi-channel computation. The desirability of associating a confidence level with each data item has been noted in both communities [PARR91], [PARH91]. Even though obtaining or assigning confidence levels is a non-trivial problem, this difficulty should not discourage us from seeking appropriate methodologies.

Our proposed data-driven methodology places the focus on data correctness/accuracy rather than the reliable operation of the subsystems producing or handling the data. To obtain the correctness probability for a given data object at particular points along a given data path, a history of transformations and operations (with attendant reliabilities) leading from known input conditions to that object is needed. A convenient way to keep such histories in a readily accessible form is to associate a dependability tag (d-tag) with each data object as an indicator of its correctness probability and to manipulate such d-tags to reflect the changing histories. Thus, a data object D and its d-tag d will comprise a tagged object $\langle D, d \rangle$ which is transformed, in both the data and tag parts, as it is processed or moved within the system.

We can essentially classify operations performed on data objects into those that lower and those that raise the dependabilities. Normal manipulations are dependability-lowering in the sense that given dependability levels d' and d'' for two data objects D' and D'' , the dependability level d that can be associated with the result D of the binary operation $b(D', D'')$ is no more than $\min(d', d'')$ and is often lower. Dependability-raising operations are mechanisms built into the system in order to restore the gradually deteriorating dependability values to acceptable levels for external use or to enable computations to proceed beyond certain *dependability checkpoints* that are strategically placed along the computation path(s).

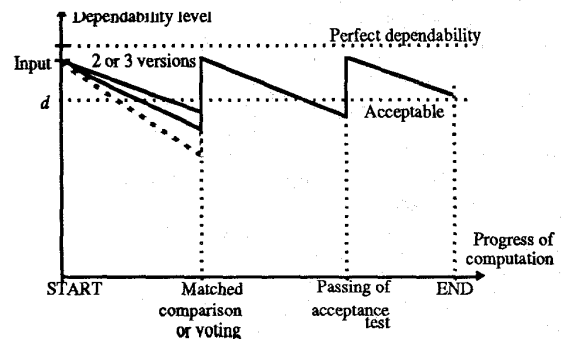


Fig. 4. Examples of dependability variation.

The challenge in such a dynamic approach is to judiciously intermix dependability-lowering and dependability-raising operations in a way that guarantees a specified level of dependability for computation results with minimal cost. The author has studied issues that are important in this optimization task and shown how such a data-driven methodology facilitates the formulation and solution of problems in multi-channel computation. It is our belief that the same method is applicable to data fusion, which can be viewed simply as a dependability-raising operation. Once the two problems are cast in the same framework, optimal resource allocation for data quality enhancement becomes possible [BALL89]. This in itself would be adequate motivation to pursue unifying methods.

7. Conclusion

We have reviewed multi-sensor data fusion and reliable multi-channel computation in order to point out their similarities and the benefits of a unified approach to their treatment. Clearly, this preliminary work can be extended in many directions (see, e.g., [PARH95], [PARH95a]).

References

- [BALL89] Ballou, D.P. and G. Kumar-Tayi, "Methodology for allocating Resources for Data Quality Enhancement", *Commun. ACM*, Vol. 32, pp. 320-329, Mar. 1989.
- [FREU95] Freund, Y., "Boosting a Weak Learning Algorithm by Majority", *Info. & Computation*, Vol. 121, pp. 256-285, Sep. 1995.
- [FTCSyr] *Proc. Int'l Symp. Fault-Tolerant Computing*, held annually, usually in June.
- [HACK90] Hackett, J.K. and M. Shah, "Multi-Sensor Fusion: A Perspective", *Proc. IEEE Int'l Conf. on Robotics and Automation*, Cincinnati, OH, May 1990, pp. 1324-1330.
- [HALL92] Hall, D.L., *Mathematical Techniques in Multi-sensor Data Fusion*, Artech House, 1992.
- [IYEN91] Iyengar, S.S., R.L. Kashyap, and R.N. Madan (Guest Eds.), Special Section on Distributed Sensor Networks, *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 21, No. 5, pp. 1027-1031, Sep/Oct. 1991.
- [IYEN94] Iyengar, S.S., D.N. Jayasimha, and D. Nadig, "A Versatile Architecture for the Distributed Sensor Integration Problem", *IEEE Trans. Computers*, Vol. 43, No. 2, pp. 175-185, Feb. 1994.
- [IYEN95] Iyengar, S.S., L. Prasad, and H. Min, *Advances in Distributed Sensor Integration*, Prentice-Hall, 1995.
- [LIBB91] Libby, V. (Ed.), *Data Structures and Target Classification*, SPIE Proc. Vol. 1470, Apr. 1991.
- [LINN91] Linn, R.J., D.L. Hall, and J. Llinas, "A Survey of Multi-Sensor Data Fusion Systems", *Proc. Data Structures and Target Classification* (SPIE Vol. 1470), Orlando, FL, Apr. 1991, pp. 13-29.
- [LORC89] Lorzak, P.R., A.K. Caglayan, and D.E. Eckhardt, "A Theoretical Investigation of Generalized Voters for Redundant Systems", in [FTCS89], pp. 444-451.
- [LUOR89] Luo, R.C. and M.G. Kay, "Multisensor Integration and Fusion in Intelligent Systems", *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 19, No. 5, pp. 901-927, Sep. 1989.
- [MARZ90] Marzullo, K., "Tolerating Failures of Continuous-Valued Sensors", *ACM Trans. Computer Systems*, Vol. 8, pp. 284-304, Nov. 1990.
- [PARH89] Parhami, B., "A Framework for the Study of Computer System Dependability", *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Oct./Nov. 1989, pp. 1017-1021.
- [PARH91] Parhami, B., "A Data-Driven Dependability Assurance Scheme with Applications to Data and Design Diversity", in *Dependable Computing for Critical Applications*, 1991, pp. 257-282.
- [PARH92] Parhami, B., "Optimal Algorithms for Exact, Inexact, and Approval Voting", in [FTCS92], pp. 404-411.
- [PARH94] Parhami, B., "A Multi-Level View of Dependable Computing", *Computers & Electrical Eng'g*, Vol. 20, No. 4, pp. 347-368, 1994.
- [PARH94a] Parhami, B., "Voting Algorithms", *IEEE Trans. Reliability*, Vol. 43, No. 4, pp. 617-629, Dec. 1994.
- [PARH95] Parhami, B., "A Taxonomy of Voting Schemes for Data Fusion and Dependable Computation", Submitted for publication.
- [PARH95a] Parhami, B., "Design of Reliable Software via General Combination of N-Version Programming and Acceptance Testing", Submitted for publication.
- [PARR91] Parra-Loera, R., W.E. Thompson, and A.P. Salvi, "Adaptive Selection of Sensors Based on Individual Performances in a Multisensor Environment", in [LIBB91], pp. 30-36.
- [PRAS91] Prasad, L., S.S. Iyengar, R.L. Kashyap, and R.N. Madan, "Functional Characterization of Fault Tolerant Integration in Distributed Sensor Networks", *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 21, pp. 1082-1087, Sep. 1991.
- [ROTH91] Rothman, P.L. and R.W. Denton, "Fusion or Confusion: Knowledge or Nonsense?", in [LIBB91], pp. 2-12.
- [STOT93] Stotts, L. and C. Stewart (Chairs), "Sensor Fusion", Session MP7 at the Asilomar Conf. on Signals, Systems, and Computers, Nov. 1993, Vol. 1, pp. 258-304.
- [TANN92] Tanner, R. and N.K. Loh, "A Taxonomy of Multisensor Fusion", *J. Manufacturing Systems*, Vol. 11, No. 5, pp. 314-325, 1992.
- [VONN56] von Neumann, J., "Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components", in *Annals of Math. Studies*, No. 34, Princeton Univ. Press, 1956, pp. 43-98.
- [WALT90] Waltz, E. and J. Llinas, *Multi-Sensor Data Fusion*, Artech House, 1990.