International Journal of Combinatorial Optimization Problems and Informatics, 10(3), Sep-Dec 2019, 21-32. ISSN: 2007-1558.
www.editada.org

Data fusion, a multidisciplinary technique

Michelle Baeza-Regalado [1], Patricio Reyes-León [1], Juan-Carlos Rojo-Mendez [1], Jazmín Rodríguez-Flores [1], Eucario-Gonzalo Perez-Perez [1]
[1] Posgrado en Tecnologías de Información y Comunicaciones, Universidad Politécnica de Pachuca
michellebaeza@micorreo.upp.edu.mx
patricioreyesleon@micorreo.upp.edu.mx
jcrojo@micorreo.upp.edu.mx
jrodriguez@upp.edu.mx
egonzalo@upp.edu.mx

Abstract. In this paper, show the general concepts of data fusion, and makes a collection of some outstanding multidisciplinary works related to the existing techniques of data fusion, its objective is to make known the methods and the multiplex ways to use this technique that can be adapted for the requested requirements.

Article Info
Received Mar 10, 2019
Accepted Mar 11, 2019

1. Introduction

Data fusion is a multidisciplinary technique based on the integration of data obtained by sensors, this technique was born as an alternative for obtaining information reducing costs and increasing the speed to obtain an approach with perspectives of accuracy.

The combination of data obtained from different sources allows data fusion to use the most outstanding information available for statistical estimation of missing data and reconstruct the unknown information. In general terms, it is the statistical estimation of the missing data. By means of data-mining it is possible to merge information to obtain missing data and later to be analyzed for the desired purposes.

2. Background

This technique was developed during the 90s, at present, numerous methodologies are based on multiple algorithms for information processing, the data fusion technique is used for the processing of information in various disciplines, thus verifying the adaptability of the technique for the requested requirements.

Data fusion is a technique for information processing, Tapiador and Casanova [1] state Data fusion can be defined as the set of techniques consisting in transmuting data sets of different characteristics and procedures in one new, in order to enhance the capabilities they enjoyed separately and prevent them from being handed over, improving with them their utilization capabilities".

Data fusion (DF) or multisensory data fusion (MSDF) is the process of combining or integrating measured or preprocessed data or information originating from different active or passive sensors or sources to produce a more specific, comprehensive and unified dataset or world model about an entity or event of interest that has been observed.

Multisensory data fusion seeks to combine information from multiple sources (including sensors, human reports and data from the internet) to achieve to inferences that cannot be obtained from a single sensor of source, or whose quality exceeds that of an interface drawn from any single source.

3. Work was done on data fusion

Below are some works related to the fusion of data recently made.

3.1 Combining mind infrared spectroscopy and paper spray mass spectrometry in data fusion model to predict the composition of coffee blends

Assis, and others [2] described a robust multivariate model for quantifying and characterizing blends of Robusta and Arabica coffees. At different degrees of roasting, 120 ground coffee blends (0.0-33.0%) were formulated. Spectra were obtained by two different techniques, attenuated total reflectance Fourier transform infrared (ATR-FTIR) spectroscopy and paper spray mass spectrometry (PS-MS). Partial least squares (PLS) models were built individually with the two types of spectra. Nevertheless, better predictions were obtained by low and medium-level data fusion, taking advantage of the synergy between these two data sets. Data fusion models were improved by variable selection, using genetic algorithms (GA) and ordered predictors selection (OPS). The smallest prediction errors were provided by OPS low-level data fusion model. The number of variables used for regression was reduced from 2145 (full spectra) to 230. The model interpretation was performed by assigning some of the selected variables to specific coffee components, such as trigonelline and chlorogenic acids.

3.2 Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research direction

Nweke et al. [3] revised was to provide an in-depth and comprehensive analysis of data fusion and multiple classifier systems techniques for human activity recognition with emphasis on mobile and wearable devices. First, data fusion methods and modalities were presented and also feature fusion, including deep learning fusion for human activity recognition were critically analyzed, and their applications, strengths and issues were identified. Furthermore, the review presents different multiple classifier system designs and fusion methods that were recently proposed in the literature. Finally, open research problems that require further research and improvements are identified and discussed.

3.3 Real-time detection of moving objects in a video sequence by using a data fusion algorithm

Tang et al. [4] proposed a novel data fusion approach to attack this problem, which combines an entropy-based Canny (EC) operator with the local and global optical flow (LGOF) method, namely EC-LGOF. Its operation contains four steps. The EC operator firstly computes the contour of moving objects in a video sequence, and the LGOF method then establishes the motion vector field. Thirdly, the minimum error threshold selection (METS) method is employed to distinguish the moving object from the background. Finally, edge information fuses temporal information concerning the optic flow to label the moving objects. Experiments are conducted and the results are given to show the feasibility and effectiveness of the proposed method.

3.4 A novel improved full vector spectrum algorithm and its application in multi-sensor data fusion for hydraulic pumps

Yu [5] described a novel approach called EWT-VCR based on Empirical Wavelet Transform (EWT) and Variance Contribution Rate (VCR) to improve the adaptability and accuracy of the fusion method. EWT is introduced as a signal preprocessing technique to decompose complex signals into variable frequency bands. Moreover, VCR is proposed to denoise, fuse EWT components at different frequency bands, and enhance useful harmonic components. The full vector spectrum technology is utilized to carry out the full vector information fusion of the improved multi-sensor signals for further spectrum analysis. The proposed methodology is applied to multi-channel vibration signal fusion for hydraulic pumps to detect specific frequencies related to pump's degradation process and a novel degradation feature named Full Vector Factor Entropy (FVFE) is extracted to describe hydraulic pump's degradation process during its life cycle. The effectiveness of the proposed methods is validated through two experimental cases.

3.5 Measurement Data Fusion Based on Optimized Weighted Least-Squares Algorithm for Multi-Target Tracking

He et al. [6] developed an improved weighted least-square algorithm based on an enhanced non-naive Bayesian classifier (ENNBC) method. According to the ENNBC method, the outliers in the measurement data are removed effectively, dataset density peaks are found quickly, and remaining effective measurements are accurately classified. The ENNBC method improved the traditional direct classification method and took the dependence among continuous density attributes into account. Four common indexes of classifiers are used to evaluate the performance of the nine methods, i.e., the normal naive Bayesian, flexible naive Bayesian (FNB), the homologous model of FNB (FNB ROT), support vector machine, k-means, fuzzy c-means (FCM), possibilistic c-means, possibilistic FCM, and our proposed ENNBC. The evaluation results show that ENNBC has the best performance based on the four indexes. Meanwhile, the multi-target tracking experimental results show that the proposed algorithm can reduce the root-mean-squared error of the position compared with the extended Kalman filter. In addition, the proposed algorithm has better robustness against large localization and tracking errors.

3.6 A Multi-Model Combined Filter with Dual Uncertainties for Data Fusion of MEMS Gyro Array

Shen Qiang et al. [7] proposed a multi-model combined filter with dual uncertainties is proposed to integrate the outputs from numerous gyroscopes. First, to avoid the limitations of the stochastic and set-membership approaches and to better utilize the potentials of both concepts, a dual-noise acceleration model was proposed to describe the angular rate. On this basis, a dual uncertainties model of gyro array was established. Then the multiple model theory was used to improve dynamic performance, and a multi-model combined filter with dual uncertainties was designed. This algorithm could simultaneously deal with stochastic uncertainties and set-membership uncertainties by calculating the Minkowski sum of multiple ellipsoidal sets. The experimental results proved the effectiveness of the proposed filter in improving gyroscope accuracy and adaptability to different kinds of uncertainties and different dynamic characteristics. Most of all, the method gave the boundary surrounding the true value, which is of great significance in attitude control and guidance applications.

3.7 Real-Time Weighted Data Fusion Algorithm for Temperature Detection Based on Small-Range Sensor Network

Zhang, Nan and Wang [8] established a heat transfer mechanism model. Second, we design a small-range sensor network for the pretreatment process and present a layered fusion structure of sharing sensors using a multi-connected fusion structure. Third, we introduce the idea of iterative operation in data processing. In addition, we use prior data for predicting state values twice in order to improve the effectiveness of extended Kalman filtering in one-time step. This study also proposes multi-fading factors on the basis of a weighted fading memory index to adjust the prediction error covariance. Finally, the state estimation accuracy of each sensor can be used as a weighting principle for the predictive confidence of each sensor by adding a weighting factor. In this study, the performance of the proposed method is verified by simulation and compared with the traditional single-sensor method. Actual industrial measurement data are processed by the proposed method for the equipment experiment. The performance index of the simulation and the experiment shows that the proposed method has a higher global accuracy than the traditional single-sensor method. Simulation results show that the accuracy of the proposed method has a 55% improvement upon that of the traditional single-sensor method, on average. In the equipment experiment, the accuracy of the industrial measurement improved by 37% when using the proposed method.

3.8 A Tilt Sensor Node Embedding a Data-Fusion Algorithm for Vibration-Based SHM

Testoni et al. [9] described a miniaturized sensor network based on low-power, light-weight and small footprint microelectromechanical (MEMS) sensor nodes capable to simultaneously measure tri-axial accelerations and tri-axial angular velocities. A real-time data fusion algorithm based on complementary filters is applied to extract tilt angles. The resulting device is designed to show competitive performance over the whole frequency range of the inertial units. Besides the capability to provide accurate measurements both in static and dynamic conditions, an optimization process has been designed to efficiently make the fusion procedure running on-sensor. An experimental campaign conducted on a pinned-pinned steel beam equipped with a network comprising several sensor nodes was used to evaluate the reliability of the developed architecture. Performance metrics revealed a satisfactory agreement to the physical model, thus making the network suitable for real-time tilt monitoring scenarios.

3.9 Data Fusion of Multivariate Time Series: Application to Noisy 12-Lead ECG Signals

Diao, Wang and Cai [10] described based on the idea of the local weighted linear prediction algorithm, a novel fusion data algorithm is proposed, which was applied in data fusion of the 12-lead ECG signals. In order to analyze the signal quality comprehensively, the quality characteristics should be adequately retained in the final fused result. In our algorithm, the values for the weighted coefficient of state points were closely related to the final fused result. Thus, two fuzzy inference systems were designed to calculate the weighted coefficients. For the sake of assessing the performance of our method, synthetic ECG signals and realistic ECG signals were applied in the experiments. Experimental results indicate that our method can fuse the 12-lead ECG signals effectively with the quality characteristics of original ECG signals inherited properly.

3.10 Quasi-real-time and continuous non-stationary strain estimation in bottom- fixed offshore structures by multimetric data fusion

Palanisamy et al. [11] investigated the use of virtual sensing of offshore structures. A Kalman filter based virtual sensing algorithm is developed to estimate responses at the location of interest. Further, this algorithm performs a multi-sensor data fusion to improve the estimation accuracy under non-stationary tidal loading. Numerical analysis and laboratory experiments are conducted to verify the performance of the virtual sensing strategy using a bottom-fixed offshore structural model. Numerical and experimental results show that the unmeasured responses can be reasonably recovered from the measured responses.

3.11 Total Variation and Signature-Based Regularizations on Coupled Nonnegative Matrix Factorization for Data Fusion

Yang et al. [12] formulated a well-posed fusion problem by incorporating total variation and signature-based regularizations for image smoothing and high-fidelity signature reconstruction. Then, the problem can be decoupled into two convex subproblems, which yield closed-form solutions separately by the alternating direction method of multipliers algorithms. Due to the large sizes of the problems, a few of constructed matrices and tensor operations are employed to simplify the expressions for reducing the computational complexities. Simulation and experimental results not only demonstrate that the performance of the proposed fusion algorithm is much better than that of state-of-the-art methods but also show that the total variation and signature-based regularizers are of paramount importance in yielding the high-spatial-resolution hyperspectral images.

3.12 Research on transformer fault diagnosis method and calculation model by using fuzzy data fusion in multi-sensor detection system

Zhang and Li [13] proposed a new transformer fault diagnosis method by using multi-band infrared image sensor and discharge circuit detection sensor to set up the eight sensors detection platform. Analyzes the transformer fuzzy fault characteristics, designs the detection platform of transformer fault diagnosis, establishes a transformer fuzzy fault diagnosis model with the exponential trust function of the photoelectric sensor, the average weighted consistency of each photoelectric sensor and the support degree of the photoelectric sensors, gives the data fusion method with the trust function of the sensor and the maximum model value. Through experiment and analysis, this fault diagnosis model can improve the fusion precision and diagnostic accuracy more effectively than the general fusion algorithm and arithmetic mean value algorithm, and can scientifically diagnose the transformer equipment of power grid from multi-angle. Thereby, the feasibility and effectiveness of the method are verified.

3.13 Data fusion based multi-rate Kalman filtering with unknown input for on-line estimation of dynamic displacements

Zheng et al. [14] proposed an improved technique for dynamic displacement estimation by fusing biased high-sampling rate acceleration and low-sampling rate displacement measurements. However, this technique can only take constant acceleration bias into account. In this paper, based on the algorithm of Kalman filter with unknown input recently developed by the authors, dynamic displacement is online estimated based on multi-rate data fusion of high-sampling rate acceleration with time-varying bias and low-sampling rate displacement measurements. The time history of time-varying acceleration bias is treated as "unknown input" in the algorithm of Kalman filter with unknown input to overcome the limitations of the previous technique. Some numerical examples with linear or polynomial acceleration bias are used to demonstrate the effectiveness of the proposed approach for online estimation of structural dynamic displacement.

3.14 A Temporal Adaptive Access Mechanism for Data Fusion in an IoT Environment

Xu et al. [15] investigated on data transmissions in IoT mainly focuses on networking without sufficiently considering the special requirements of the upper-layer applications, such as the data fusion process, that are consuming the transmitted data. In this paper, we tackle the problem of data transmission for data fusion in an IoT environment by proposing a distributed scheduling mechanism VD-CSMA in wireless sensor networks, which considers the values for data fusion, as well as the delay constraints of packets when determining their priority levels for transmission. Simulation results have shown that VD-CSMA may enhance both throughput and delay performance of data transmission as compared to the typical scheduling schemes used for data fusion in IoT.

3.15 Skeleton estimation and tracking by means of depth data fusion from depth camera networks

Carraro, Munaro and Menegatti [16] described an approach for estimation and tracking of the skeleton of the human body from camera networks exploiting only depth data. The algorithm takes advantage of multiple views by building and merging the 3D point clouds. The final skeleton is computed from a virtual depth image generated from this point cloud by means of back-projection to a reference camera image plane. Before the back -projection, the person point cloud is frontalized concerning the reference camera, so that the virtual depth image represents the person from a frontal viewpoint and the accuracy of the skeleton estimation algorithm is maximized. Our experiments show how the proposed approach boosts the performance concerning other state-of-the-art approaches. Moreover, the proposed algorithm requires low computational burden, thus running in real-time.

3.16 Method of speed data fusion based on Bayesian combination algorithm and high-order multi-variable Markov model

Zhang et al. [17] studied a data fusion method based on Bayesian fusion rules is proposed to merge traffic speed from different data sources according to their prior probability that can be inferred from a high-order multivariable Markov model by considering the relations of multiple traffic factors in a systemic perspective. Case studies based on freeway data, such as loop data, INRIX data, and data from the National Performance Management and Research Data Set, are performed to validate the effectiveness of the proposed speed fusion method.

3.17 Multi-sensor data fusion using NIHS transform and decomposition algorithms

Ankarao, Sowmya and Soman [18] proposed three methods using intensity, hue, saturation (IHS) and nonlinear IHS (NIHS) transform along with the Dynamic Mode Decomposition (DMD) and 2D-Empirical Mode Decomposition (2D-EMD or IEMD). An intensity plane is calculated from the NIHS transform. The modes are constructed using DMD by considering the variations between the intensity plane computed using NIHS transforms of a low-resolution multi-spectral image and a panchromatic image. Similarly, 2D-EMD is also used for image fusion. Modes are subjected to weighted fusion rule to get an intensity plane with spatial and edge information. Finally, the calculated intensity plane is concatenated along with the hue and saturation plane of the low-resolution multi-spectral image and transformed into RGB colour space. Thus, the fused images have high spatial and edge information on spectral bands. The experiments and its quality assessment assure that proposed methods perform better than the existing methods.

3.18 Trunk detection based on laser radar and vision data fusion

Xue et al. [19] developed a novel algorithm based on data fusion with a vision camera and a 2D laser scanner to detect tree trunks accurately. The transformation was built from a laser coordinate system to an image coordinate system, and the model of a rectangle calibration plate with two inward concave regions was established to implement data alignment between two sensors data. Then, data fusion and decision with Dempster-Shafer theory were achieved through integration of decision level after designing and determining basic probability assignments of regions of interesting (RoIs) for laser and vision data respectively. Tree trunk width was calculated by using laser data to determine basic probability assignments of RoIs of laser data. And a stripping segmentation algorithm was presented to determine basic probability assignments of RoIs of vision data, by calculating the matching level of RoIs like tree trunks. A robot platform was used to acquire data from sensors and to perform the developed tree trunk detection algorithm. Combined calibration tests were conducted to calculate a conversion matrix transforming from the laser coordinate system to the image coordinate system, and then field experiments were carried out in a real pear orchard under sunny and cloudy conditions, with trunk width measurement of 120 trees and 40 images processed by the presented stripping segmentation algorithm. Results showed that the algorithm was successful in detecting tree trunks and data fusion improved the ability for tree trunk detection. This algorithm could provide a new method for tree trunk detection and accurate production and management in orchards.

3.19 Data fusion combined with echo state network for multivariate time series prediction in the complex electromechanical system

Liang et al. [20] combined data fusion with ESN to improve the prediction accuracy of multivariate time series, which focuses on both the processing of the input information of predictive model and the optimizing of predictive model. First, multi-sources data fusion is presented to obtain the new input for prediction model before predicting; phase space reconstruction and self-adaptive weighted fusion algorithm are adapted to fuse multivariate time series and obtain more complete information. Then, leaky rectifier liner units are used to replace the original activation function, and locality preserving projection is employed to optimize the state matrix of the reservoir. Finally, the effectiveness of the proposed method is verified by an analysis of one case study of real compressor groups data sets in chemical production system. The results and a comparison with the traditional method show that the proposed method can greatly enhance the prediction accuracy of multivariate time series and the one-step-ahead prediction accuracy is improved by three orders of magnitude as well as a better generalization ability is obtained in the multi-steps ahead prediction.

3.20 Data Fusion Using Improved Support Degree Function in Aquaculture Wireless Sensor Networks

Shi et al. [21] proposed a data fusion method using a novel function that is Dynamic Time Warping time series strategy improved support degree (DTWS-ISD) for enhancing data quality, which employs a Dynamic Time Warping (DTW) time series segmentation strategy to the improved support degree (ISD) function. We use the DTW distance to replace Euclidean distance, which can explore the continuity and fuzziness of data streams, and the time series segmentation strategy is adopted to reduce the computation dimension of DTW algorithm. Unlike Gauss support function, ISD function obtains mutual support degree of sensors without the exponent calculation. Several experiments were finished to evaluate the accuracy and efficiency of DTWS-ISD with different performance metrics. The experimental results demonstrated that DTWS-ISD achieved better fusion precision than three existing functions in a real-world WSN water quality monitoring application.

3.21 Modified Bayesian data fusion model for travel time estimation considering spurious data and traffic conditions

Mil and Piantanakulchai [22] presented a framework for the development of the travel time estimation model using multiple sources of data with consideration of spurious data and traffic conditions. A modified Bayesian data fusion approach, combined with the Gaussian mixture model, is used to fuse the travel time data, which are estimated from different types of sensors to improve accuracy, precision, as well as completeness of data, in terms of spatial and temporal distribution. Two additional features are added into existing models including the difference of traffic conditions classified by the Gaussian mixture model and the bias estimation from individual sensor by introducing a non-zero mean Gaussian distribution which learned from the training dataset. The methodology and computational procedure are presented. The Gaussian mixture model is used to classify states of traffic into predefined number of traffic regimes. Once a traffic condition is classified, the modified Bayesian data fusion approach is used to estimate travel time. The proposed model provides explicit advantages over the basic Bayesian approach, such as being robust to noisy data, reducing biases of an individual estimation, and producing a more precise estimation of travel time. Two different real-world datasets and one simulated dataset are used to evaluate the performance of the proposed model under three different traffic regimes: free flow, transitional flow and congested flow regimes. The results when compared with the results from benchmark models show significant improvement in the accuracy of travel time estimation in terms of mean absolute percentage errors (MAPE) in the range of 3.46% to 16.3%.

3.22 Framework for Location Data Fusion and Pose Estimation of Excavators Using Stereo Vision

Soltani, Zhu and Hammad [23] investigated the opportunities to fuse CV-based methods and real-time location systems (RTLSs) and apply stereo vision methods to formulate a comprehensive framework for estimating the three-dimensional (3D) poses of excavators as some of the most widely used equipment on construction sites. Instead of using specialized tools, such as off-the-shelf stereo cameras or markers, this study evaluates the applicability of using surveillance cameras on construction sites as stereo cameras. Moreover, RTLS data and two or more cameras' data are fused by synchronizing the time and coordinate systems of the cameras and RTLS to investigate the potential of enhancing the accuracy of the pose estimation system and reducing the computational load. Finally, the performance of the proposed framework is evaluated by integrating the results of the excavator parts' detection, the backgrounds' subtraction, and the two-dimensional (2D) skeletons' extraction of the parts from each camera's view.

3.23 Data fusion of surface data sets of X-ray computed tomography measurements using locally determined surface quality values

Muller and Hausotte [24] proposed creating several measurements of the same object, which only differ in their orientation inside the ray path of the measurement system. These measurements are then processed to selectively correct faulty surface regions. To calculate the needed geometrical transformations between the different measurements with the goal of a congruent alignment in one coordinate system, an extension of the iterative closest point (ICP) algorithm is used. To quantitatively classify any surface point regarding its quality value to determine the individual need of correction for each point, the local quality value (LQV) method is used, which has been developed at the Institute of Manufacturing Metrology. Different data fusion algorithms are presented whose performances are tested and verified using nominal-actual comparisons.

3.24 Medical data fusion algorithm based on the Internet of things

Zhang et al. [17] studied the monitoring of medical data in the Internet of things on data fusion and related routing technology. According to the particularity of the data in the medical Internet of things, a data fusion cluster-tree construction algorithm based on event-driven (DFCTA) is proposed. The fusion delay problem in the network is analyzed, and the minimum fusion delay method is proposed by calculation of the fusion waiting time of the nodes. Finally, the intelligent health management data fusion system in the medical Internet of things is designed. Aiming at the characteristics of multilevel integration of multisource heterogeneous data fusion for intelligent health management, the data fusion architecture of fusion tree composed of fusion nodes is proposed. The experiment shows that the DFCTA algorithm has good fusion performance. Based on the above findings, it is concluded that the algorithm is a fast and reliable method, which has important practical significance.

3.25 Occupancy prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings

Wang, Chen and Hong [25] compared three popular machine learning algorithms, including k-nearest neighbors (kNN), support vector machine (SVM), and artificial neural network (ANN), combined with three data sources, including environmental data, Wi-Fi data, and fused data, to optimize the occupancy models' performance in various scenarios. Three error measurement metrics, the mean average error (MAE), mean average percentage error (MAPE), and root mean squared error (RMSE), have been employed to compare the models' accuracies. Examined with an on-site experiment, the results suggest that the ANN-based model with fused data has the best performance, while the SVM model is more suitable with Wi-Fi data. The results also indicate that, compared with independent data sources, the fused data set does not necessarily improve model accuracy but shows a better robustness for occupancy prediction.

3.26 Multi-sensors data fusion through fuzzy clustering and predictive tools

Majumder and Pratihar [26] developed a multi-sensors data fusion technique by using fuzzy clustering and predictive tools. The data were first clustered based on their similarity using an entropy-based fuzzy C-means clustering technique and the obtained clusters were utilized to develop a fuzzy reasoning based predictive tool. The novelty of this study lies with the application of a clustering algorithm, which can ensure both compactness and distinctness of the developed clusters, and development of a reasoning tool utilizing the information of obtained clusters. Two types of multi-sensors data were used to test the performance of the proposed algorithm. Results were compared with those available in the literature, and the developed technique was found to perform better than the previous approaches on both the data sets. The better performance of the proposed algorithm could be due to its in-depth search of the data set through similarity-based fuzzy clustering followed by the development of fuzzy reasoning tool utilizing the information of obtained clusters. (C) 2018 Elsevier Ltd. All rights reserved.

3.27 Gas Leak Location Detection Based on Data Fusion with Time Difference of Arrival and Energy Decay Using an Ultrasonic Sensor Array

Wang, Wang, and Hong [27] proposed an ultrasonic leak location approach based on multi-algorithm data fusion. With the help of a planar ultrasonic sensor array, the eigenvectors of two individual algorithms, i.e., the arrival distance difference, as determined from the time difference of arrival (TDOA) location algorithm, and the ratio of arrival distances from the energy decay (ED) location algorithm, are extracted and fused to calculate the three-dimensional coordinates of leak holes. The fusion is based on an extended Kalman filter, in which the results of the individual algorithms are seen as observation values. The final system state matrix is composed of distances between the measured leak hole and the sensors. Our experiments show that, under the condition in which the pressure in the measured container is 100 kPa, and the leak hole-sensor distance is 800 mm, the maximum error of the calculated results based on the data fusion location algorithm is less than 20 mm, and the combined accuracy is better than those of the individual location algorithms.

3.28 Data Fusion Based on Adaptive Interacting Multiple Model for GPS/INS Integrated Navigation System

Zhang [17] proposed the adaptive interacting multiple model (AIMM) filter method to enhance navigation performance. The soft-switching characteristic, which is provided by interacting multiple model algorithm, permits process noise to be converted between upper and lower limits, and the measurement covariance is regulated by Sage adaptive filtering online Moreover, since the pseudo-range and Doppler observations need to be updated, an updating policy for classified measurement is considered. Finally, the performance of the GPS/INS integration method by AIMM is evaluated by a real ship, and comparison results demonstrate that AIMM could achieve a more position accuracy.

3.29 Data fusion for integrated planar and cylindrical tomographic flame sensing

Liu et al. [28] studied an alternative sensing strategy integrating a circular planar electrode array with a cylindrical electrode assembly is proposed. Analyses of the 3D sensitivity maps show complementary properties of the new sensor, as by positioning the cylindrical sensor the combined sensitivity map can be more uniform and originally weak portion be enhanced. Meanwhile, a data fusion algorithm based on the Newton-Raphson method and expansion by Taylor series is derived to facilitate the integrated advantage of the two sensor assemblies. Simulations are performed and the new algorithm produces better images with the new sensor. Also, data analysis shows that the error of the reconstructed image can be 50-60% smaller by the new method than a traditional sensor of the equal number of the electrodes. Experimental visualization of the flames is carried out and the data fusion algorithm applied to verify the new sensing strategy. The results demonstrate the effectiveness of such a strategy, as the whole length of the flames over the entire visualization zone, particularly up to 170 mm, is well presented by the 3D images using the derived data fusion algorithm, which is in sharp contrast to the conventional 3D image methods without data fusion.

3.30 Research on data fusion algorithm and anti-collision algorithm based on the internet of things

Cui et al. [29] introduced the hierarchical architecture of the Internet of things and its key technologies, and introduces the sensing and recognition technology of the internet of things, and expatriates the data fusion model and routing algorithm of the sensor network and the working principle and the anti-collision problem of the RFID system. The improved clustering routing protocol Leach based on distributed data fusion effectively prolongs the lifetime of the sensor network. A data fusion algorithm based on Gauss membership function has introduced, and the fusion technology and routing technology have combined to process the data, which improves the efficiency and accuracy of data acquisition in Sensor Networks. An efficient binary search anti-collision algorithm called AEBS, with unknown tag number estimation, which has proposed to recognize unknown tags in the range of reader's function. Theoretical analysis and experimental simulation results show that the AEBS algorithm improves the efficiency and stability of the system identification and reduces the recognition time of the whole system.

3.31 Utilization of Multisensor Data Fusion for Magnetic Nondestructive Evaluation of Defects in Steel Elements under Various Operation Strategies

Psuj [30] applied a magnetic sensor-array-based nondestructive system to inspect defects inside circular-shaped steel elements. The experiments were carried out for various sensor network strategies, followed by the fusion of multisensor data for each case. In order to combine the measurements, first data registration and then four algorithms based on spatial and transformed representations of sensor signals were applied. In the case of spatial representation, the data were combined using an algorithm operating directly on input signals, allowing pooling of information. To build the transformed representation, a multiresolution analysis based on the Laplacian pyramid was used. Finally, the quality of the obtained results was assessed. The details of algorithms are given and the results are presented and discussed. It is shown that the application of data fusion rules for magnetic multisensor inspection systems can result in the growth of reliability of proper identification and classification of defects in steel elements depending on the utilized configuration of the sensor network.

3.32 Hairline breakage detection in X-ray images using data fusion

Harrieta and Wiselin [31] dialled with identification of hairline breakage in the X-Ray images. The crack in the X-Ray images can be missed due to absence of sharp edges and the intensity inhomogeneity. This work is carried out in two phases. In the first phase, the preprocessing step is done using anisotropic diffusion filter and wavelet to preserve the edges and fine details. In the second phase Expectation Maximization (EM) algorithm is used for segmenting the image. The mask produced from the EM algorithm separates the bone region. The intensity variation calculation is performed over the selected region to detect the cracks. The performance of the proposed work is calculated using the parameters sensitivity and accuracy. This new approach is experimented with ten patient's data and validated by Radiologists. The performance of the proposed work is compared with recent works. This work greatly improves the accuracy of the segmentation on medical images and the overall accuracy is about 98%.

3.34 Measurement data fusion with cascaded Kalman and complementary filter in the flight parameter indicator for hang-glider and paraglider

Brzozowki et al. [32] investigated an assistive instrument for hang-gliders and paragliders resulted in the three prototypes of the integrated device for sensing, processing and displaying data to the pilot. The data processing contains innovative data fusion and filtration algorithms with cascaded filter build with Kalman and complementary fillers. The developed prototypes have been successfully tested in the laboratories and during flight tests.

3.35 Research on estimation of optical fiber probe gas holdup based on the adaptive weighted data fusion algorithm

Li and liao [33] presented an evaluation method to measure the gas holdup of fiber optic probe based on data fusion. Firstly, the method applies the Grubbs criterion to eliminate the omission errors in the gas holdup data collected by a single fiber optic probe in a period of time. Secondly, it acquires the optimal value of the fiber optic probe in a certain time quantum by estimating the data collected by the fiber optic probe in batches. After obtaining the optimal value of the data collected by the four fiber optic probes in the downhole optical fiber probe array, the author adopts the self-adaptive weighted fusion algorithm, fuses data according to the optimization principle of the total minimum mean squared error and calculates the accurate value of the gas holdup of oil-gas-water three-phase flow in the wellbore in a certain time quantum. The experimental results show that this kind of method can improve the measurement accuracy of gas holdup and achieve more accurate estimate in oil-gas-water three-phase flow.

4. Conclusions

The fusion of data is a useful technique for the processing of existing information that allows an adequate analysis, the comparative advantages of data fusion allow the analysis of information, the generation of data, objectivity and study of existing data to cover the needs required.

As you can read, the data fusion technique can be adapted for studies of various disciplines, such as Chemistry, Medicine, Information Technology, among many others.

References

1. Tapiador, F. J., & Casanova, J. L. (2001). Aplicaciones de la Fusión de Datos en la Ordenación Territorial. Teledetección, Medio Ambiente y Cambio Global, 508-511.
2. Assis, C., Pereira, H.V., Amador, V.S., Augusti, R., de Oliveira, L.S., & Sena, M.M. (2019). Combining mid infrared spectroscopy and paper spray mass spectrometry in a data fusion model to predict the composition of coffee blends. *Food chemistry*, *281*, 71-77.
3. Nweke, H.F., Teh, Y.W., Mujtaba, G., & Al-Garadi, M.A. (2019). Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions. *Information Fusion*, *46*, 147-170.
4. Tang, C., Hu, H., Zhang, M., Wang, W.J., Wang, X.F., Cao, F., and Li, W. (2018). Detección en tiempo real de objetos en movimiento en una secuencia de video mediante el uso del algoritmo de fusión de datos. *Transacciones del Instituto de Medición y Control, 0142331218773550.*
5. Yu, H., Li, H., Li, Y., y Li, Y. (2019). Un nuevo algoritmo mejorado de espectro vectorial completo y su aplicación en la fusión de datos de sensores múltiples para bombas hidráulicas. *Medida*, *133*, 145-161.
6. He, X., Wang, T., Liu, W., & Luo, T. (2019). Measurement Data Fusion Based on Optimized Weighted Least-Squares Algorithm for Multi-Target Tracking. *IEEE Access*.
7. Shen, Q., Liu, J., Zhou, X., & Wang, L. (2019). A Multi-Model Combined Filter with Dual Uncertainties for Data Fusion of MEMS Gyro Array. *Sensors*, *19*(1), 85.
8. Zhang, Z., Nan, X., & Wang, C. (2019). Real-Time Weighted Data Fusion Algorithm for Temperature Detection Based on Small-Range Sensor Network. *Sensors*, *19*(1), 64.
9. Testoni, N., Zonzini, F., Marzani, A., Scarponi, V., & De Marchi, L. (2019). A Tilt Sensor Node Embedding a Data-Fusion Algorithm for Vibration-Based SHM. *Electronics*, *8*(1), 45.
10. Diao, C., Wang, B., & Cai, N. (2019). Data fusion of multivariate time series: Application to noisy 12-lead ECG signals. *Applied Sciences*, *9*(1), 105.
11. Palanisamy, R. P., Jung, B. J., Sim, S. H., & Yi, J. H. (2019). Quasi real-time and continuous non-stationary strain estimation in bottom-fixed offshore structures by multimetric data fusion. *SMART STRUCTURES AND SYSTEMS*, *23*(1), 61-69.
12. Yang, F., Ma, F., Ping, Z., & Xu, G. (2019). Total Variation and Signature-Based Regularizations on Coupled Nonnegative Matrix Factorization for Data Fusion. *IEEE Access*, *7*, 2695-2706.
13. Zhang, C., Guo, C., & Zhang, D. (2018). Data Fusion Based on Adaptive Interacting Multiple Model for GPS/INS Integrated Navigation System. *Applied Sciences*, *8*(9), 1682.
14. Zheng, Z., Qiu, H., Wang, Z., Luo, S., & Lei, Y. (2019). Data fusion based multi-rate Kalman filtering with unknown input for on-line estimation of dynamic displacements. *Measurement*, *131*, 211-218.
15. Xu, J., Liu, S., Lu, X., Li, L., Liang, H., Duan, Q., & Liu, R. (2018). A Temporal Adaptive Access Mechanism for Data Fusion in an IoT Environment. *Sensors*, *18*(12), 4205.
16. Carraro, M., Munaro, M., & Menegatti, E. (2018). Skeleton estimation and tracking by means of depth data fusion from depth camera networks. *Robotics and Autonomous Systems*, *110*, 151-159.
17. Zhang, W., Qi, Y., Zhou, Z., Biancardo, S. A., & Wang, Y. (2018). Method of speed data fusion based on Bayesian combination algorithm and high-order multi-variable Markov model. *IET Intelligent Transport Systems*, *12*(10), 1312-1321.
18. Ankarao, V., Sowmya, V., & Soman, K. P. (2018). Multi-sensor data fusion using NIHS transform and decomposition algorithms. *Multimedia Tools and Applications*, *77*(23), 30381-30402.
19. Xue, J., Fan, B., Yan, J., Dong, S., & Ding, Q. (2018). Trunk detection based on laser radar and vision data fusion. *International Journal of Agricultural and Biological Engineering, 11*(6), 20-26.
20. Liang, Y., Gao, Z., Gao, J., Wang, R., & Zhao, H. (2018). Data fusion combined with echo state network for multivariate time series prediction in complex electromechanical system. *Computational and Applied Mathematics*, *37*(5), 5920-5934.
21. Shi, P., Li, G., Yuan, Y., & Kuang, L. (2018). Data Fusion Using Improved Support Degree Function in Aquaculture Wireless Sensor Networks. *Sensors*, *18*(11), 3851.
22. Mil, S., & Piantanakulchai, M. (2018). Modified Bayesian data fusion model for travel time estimation considering spurious data and traffic conditions. *Applied Soft Computing*, *72*, 65-78.
23. Soltani, M. M., Zhu, Z., & Hammad, A. (2018). Framework for Location Data Fusion and Pose Estimation of Excavators Using Stereo Vision. *Journal of Computing in Civil Engineering*, *32*(6), 04018045.
24. Müller, A. M., & Hausotte, T. (2018). Data fusion of surface data sets of X-ray computed tomography measurements using locally determined surface quality values. *Journal of Sensors and Sensor Systems*, *7*(2), 551-557.
25. Wang, W., Chen, J., and Hong, T. (2018). Occupancy prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings. *Automation in Construction, 94*, 233-243.
26. Majumder, S., & Pratihar, D.K. (2018). Multi-sensors data fusion through fuzzy clustering and predictive tools. *Expert Systems with Applications*, *107*, 165-172.
27. Wang, T., Wang, X., & Hong, M. (2018). Gas Leak Location Detection Based on Data Fusion with Time Difference of Arrival and Energy Decay Using an Ultrasonic Sensor Array. *Sensors*, *18*(9), 2985.
28. Liu, J., Liu, S., Sun, S., Wang, M., Schlaberg, I. H. I., & Lu, G. (2018). Data fusion for integrated planar and cylindrical tomographic flame sensing. *Flow Measurement and Instrumentation*, *62*, 213-222.
29. Cui, Y., Ma, Y., Zhao, Z., Li, Y., Liu, W., & Shu, W. (2018). Research on data fusion algorithm and anti-collision algorithm based on internet of things. *Future Generation Computer Systems*, *85*, 107-115.

30. Psuj, G. (2018). Utilization of Multisensor Data Fusion for Magnetic Nondestructive Evaluation of Defects in Steel Elements under Various Operation Strategies. *Sensors*, *18*(7), 2091.
31. Linda, C. H., & Jiji, G. W. (2018). Hairline breakage detection in X-ray images using data fusion. *Multimedia Tools and Applications*, *77*(13), 17207-17222.
32. Brzozowski, B., Rochala, Z., Wojtowicz, K., & Wieczorek, P. (2018). Measurement data fusion with cascaded Kalman and complementary filter in the flight parameter indicator for hang-glider and paraglider. *Measurement*, *123*, 94-101.
33. Li, R., & Liao, R. (2018). Research on estimation of optical fiber probe gas holdup based on the adaptive weighted data fusion algorithm. *International Journal of Heat and Technology*, *36*(2), 543-547.
34. Joyanes Aguilar, L. (2003). Fundamentos de programación: algoritmos y estructura de datos y objetos.
35. Raol, J. R. (2009). *Multi-sensor data fusion with MATLAB®*. CRC press.
36. Liggins II, M., Hall, D., & Llinas, J. (Eds.). (2017). *Handbook of multisensor data fusion: theory and practice*. CRC press.
37. Xuewei, Z., & Hanshan, L. (2019). Research on transformer fault diagnosis method and calculation model by using fuzzy data fusion in multi-sensor detection system. *Optik*, *176*, 716-723.
38. Zhang, W., Yang, J., Su, H., Kumar, M., & Mao, Y. (2018). Medical data fusion algorithm based on internet of things. *Personal and Ubiquitous Computing*, *22*(5-6), 895-902.