Sensors **2014**, *14*, 19669-19686; doi:10.3390/s141019669

OPEN ACCESS

sensors

ISSN 1424-8220

www.mdpi.com/journal/sensors

Article

An Integrated Model for Robust Multisensor Data Fusion

Bo Shen *, Yun Liu and Jun-Song Fu

School of Electronic and Information Engineering, Key Laboratory of Communication and Information Systems, Beijing Municipal Commission of Education, Beijing Jiaotong University, Beijing 100044, China; E-Mails: liuyun@bjtu.edu.cn (Y.L.); 12120067@bjtu.edu.cn (J.-S.F.)

* Author to whom correspondence should be addressed; E-Mail: bshen@bjtu.edu.cn; Tel./Fax: +86-10-5168-4227.

External Editor: Leonhard M. Reindl

Received: 3 September 2014; in revised form: 4 October 2014 / Accepted: 17 October 2014 / Published: 22 October 2014

Abstract: This paper presents an integrated model aimed at obtaining robust and reliable results in decision level multisensor data fusion applications. The proposed model is based on the connection of Dempster-Shafer evidence theory and an extreme learning machine. It includes three main improvement aspects: a mass constructing algorithm to build reasonable basic belief assignments (BBAs); an evidence synthesis method to get a comprehensive BBA for an information source from several mass functions or experts; and a new way to make high-precision decisions based on an extreme learning machine (ELM). Compared to some universal classification methods, the proposed one can be directly applied in multisensor data fusion applications, but not only for conventional classifications. Experimental results demonstrate that the proposed model is able to yield robust and reliable results in multisensor data fusion problems. In addition, this paper also draws some meaningful conclusions, which have significant implications for future studies.

Keywords: multisensors; data fusion; Dempster-Shafer theory; extreme learning machine

---

1. Introduction

Multisensor data fusion is a technology to enable combining information from several sensors into a unified result [1]. In multisensor data fusion, the information to be handled is always random, vague, imprecise and heterogeneous. The developed data fusion framework needs to be able to eliminate the

redundancy, uncertainty, and fuzziness of the data sources to achieve robust and accurate fusion results. Besides, the framework is required to be able to obtain unified results from multisensor sources, even when conflicts exist among them. Yet data fusion has proved to be valuable in many applications, like pattern recognition and classification [2], image processing [3] and expert systems [4]. Many theories have been applied in multisensor data fusion, such as the Bayesian approach [5], evidential theory [6], fuzzy set theory [7], and rough set theory [8].

The Dempster Shafer evidence theory (DSET), also known as the evidential theory, is a flexible method in multisensor data fusion [9–11]. Its framework is able to deal with information uncertainty, imprecision, randomness, conflicts and heterogeneity. In addition, Dempster's combination rule allows for combining several information sources into a unified one, which makes it popular in multisensor data fusion applications. However, when using DSET in practice, the accuracy of fusion results is always far from satisfactory for two reasons. First, it is difficult to get reasonable and accurate basic belief assignment (BBA); second, it is a thorny problem to make decision with the unified BBA.

A good and effective BBA function, also called mass function, is the prerequisite for applying evidential theory with an aim to combine pieces of evidences into a unified one. There is no general solution for mass function in DSET, and BBAs are usually calculated in heuristic ways. The most typical way is to define the membership functions to mapping data into masses. These methods are easy to implement, such as the fuzzy C-mean algorithm [12], automatic thresholding method [13]. Another way is to develop a transducer mechanism to obtain BBAs by using artificial neural network (ANN) [14–16] and it is has been proved in image data fusion applications. In wireless sensor networks (WSNs), the calculation capacity of sensor nodes is limited, thus implementing the ANN based BBA constructing with high computation complexity is not the most appropriate solution. In DSET, the sources to be combined with Dempster's combinational rule are set to be independent from each other. In reality, it is difficult to meet the condition. Many endeavors have been devoted to solving this problem, and the Belief Transfer Model (TBM) [17] presented by Semts is widely accepted. In TBM, two levels of the beliefs are assumed: the Credal Level (CL) where compound classes are allowed to exist and the Pignistic Level (PL) where beliefs are transferred into probability to make decisions [18]. However, the final pignistic probability is always updated by carving up the intersection classes to the singleton class with a certain proportion. The final decisions are actually determined by the belief of the singleton classes, but have nothing to do with the compound classes [19,20], making the compound classes useless in final decisions. As a result, how to make decisions according to the unified BBA becomes a thorny problem.

To solve the mass constructing problem, a BBA function converting distance to mass is developed in this paper. Distance is a direct and effective reflection measuring the similarity or difference of different classes. It is not affected by data dimension and there are several distance definitions we can choose according to the reality. Besides, distance has a wide range of application, but not only for some specific situations. The key point of the mass function is that the transducer mechanism between distance and BBAs must be reasonable, effective and flexible. A radial basis function can be used to mapping distance to BBAs of different subsets, and the obtained BBAs should be able to adjust according to different parameters. As a result, the mass function will be able to adaptively generate reasonable BBAs according to the given sample sets.

Another endeavor to solve the mass constructing problem is presenting an algorithm to synthetize BBAs from different BBAs obtained by different functions or experts. The reason is that different BBA

functions or experts maintain their own positions and we've found that the synthetic BBA is always more reliable than independent BBAs. However, how to build reasonable BBA synthetizing algorithm becomes another problem. The Jousselme distance [21] is a widely accepted way of measuring distance between two evidences bodies and it is able to reflect the conflict degree of evidences properly. Hence we develop the synthetizing algorithm by utilizing Jousselme distance to get synthetic BBAs, which represents a comprehensive knowledge of the information source by combining the views of different BBA functions or experts.

The decision making mechanism is also vital to obtain high accurate fusion results. In the perspective of dimensionality reduction, the unified BBA can be regarded as the comprehensive presentation of the original multisensor data. It's much more easily to make decision with the unified BBA rather than the original data. Traditionally, the Belief Transfer Model (TBM) is always used to convert the final BBA to Pignistic probability. However, the transferred Pignistic probability essentially depends on the singleton classes while the compound classes have no decisive influence on the final decision. In some conditions when the belief assignments of the compound classes are larger than the singleton classes, the accuracy of TBM is suspicious. Thus a decision making mechanism based on Extreme learning machine (ELM) is presented to solve the decision making problem. ELM [22] is a fast and easy implementing ANN without iterated operation, and to our knowledge, the accuracy of ELM is no worse compared to any other ANNs. Thus a decision making mechanism based on ELM is presented to solve the decision making problem.

A systematic multisensor data fusion model is built up in the basis of the above three main improvements. The framework will be illustrated in detailed, which includes four steps: BBA construction, BBA synthesis, combination of evidences and decision making. Experimental results and analysis on the IRIS data set and Diabetes Diagnosis data set will be illustrated to show the performance and result accuracy of the proposed algorithm.

The remainder of this paper is organized as follows: Section 2 introduces the preliminaries of thee Dempster Shafer evidence theory. Section 3 illustrates the proposed method in detailed. The experiments along with the observations are provided in Section 4. Conclusions and discussions are finally presented in Section 5.

2. Preliminaries of Dempster Shafer Evidence Theory

Dempster Shafer evidence theory (DSET) is an extension of the classical probability theory. It is a flexible evidential reasoning approach for dealing with the uncertainty in multisensor data fusion. Let $\Omega = (\omega_1, \cdots, \omega_c)$ be a finite non-empty set and Ω is mutually exclusive and exhaustive. Ω is called the frame of discernment, the corresponding power set is 2^{Ω}, which is composed by all possible subsets of Ω. The mass function of 2^{Ω} is defined as a function m: $2^{\Omega} \rightarrow [0, 1]$ and it satisfies the following property:

$$\sum_{A \subseteq \Omega} m(A) = 1 \text{ and } m(\varnothing) = 0 \tag{1}$$

where $\emptyset$ denotes the null set, $m(A)$ is called the basic belief assignment (BBA) of subset A. The numerical value of $m(A)$ can be interpreted as the support degree of proportion A belongs to Ω with all

relevant and available evidences. A focal element is a subset A with non-zero mass assignment and we call $(A, m(A))$ a piece of evidence.

For a given element A of Ω, the belief function and plausibility function of A are denoted by $Bel(A)$ and $Pls(A)$, respectively. $Bel(A)$ is the total mass of elements belonging to A and $Pls(A)$ is the maximum total mass of elements that may be distributed in A. Therefore, they are defined as the following expressions:

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{2}$$

$$Pls(A) = \sum_{A \cap B \neq \varnothing} m(B) \tag{3}$$

where $Bel(A)$ and $Pls(A)$ are the lower and upper belief of hypothesis A, respectively. $Bel(A)$ and $Pls(A)$ satisfy the following relation:

$$Pls(A) = 1 - Bel(\overline{A}) \tag{4}$$

$$Pls(A) \geq Bel(A) \tag{5}$$

where $\bar{A}$ is the complementary set of A. When $Pls(A) = Bel(A)$, A must be a singleton class. For an arbitrary focal element in Ω, its BBA distributes in an explicit measure of a belief interval $[Pls(A), Bel(A)]$.

In DSET, The Dempster combinational rule can be used to fuse all independent evidences into one. It is expressed as:

$$\oplus_{i=1}^{n} m_i(A) = \frac{1}{1-K} \sum_{\cap_{i=1}^{n} A_n = A} \prod_{i=1}^{n} m_i(A^i) \tag{6}$$

$$\text{with} \quad K = \sum_{\cap_{i=1}^{n} A_n = \varnothing} \prod_{i=1}^{n} m_i(A^i) \tag{7}$$

where $\oplus$ denotes the combinational operator. A_i designates the focal element regarding to data source i. K indicates the conflict among the sources to be combined. After combining, a Pignistic probability can be obtained by using the Transfer Belief Model (TBM) and a typical transfer formula is defined as [17]:

$$\text{Bet}P(A_i) = \sum_{A_i \subseteq A_M} \left(\frac{1}{|A_M|}\right) m(A_M) \tag{8}$$

where $\text{Bet}P(A_i)$ is the transferred Pignistic probability regarding to A_i. At last, a decision can be made by choosing the class with maximum Pignistic probability as the result of the multisensor data fusion process.

3. The DSET-E Multisensor Data Fusion Model

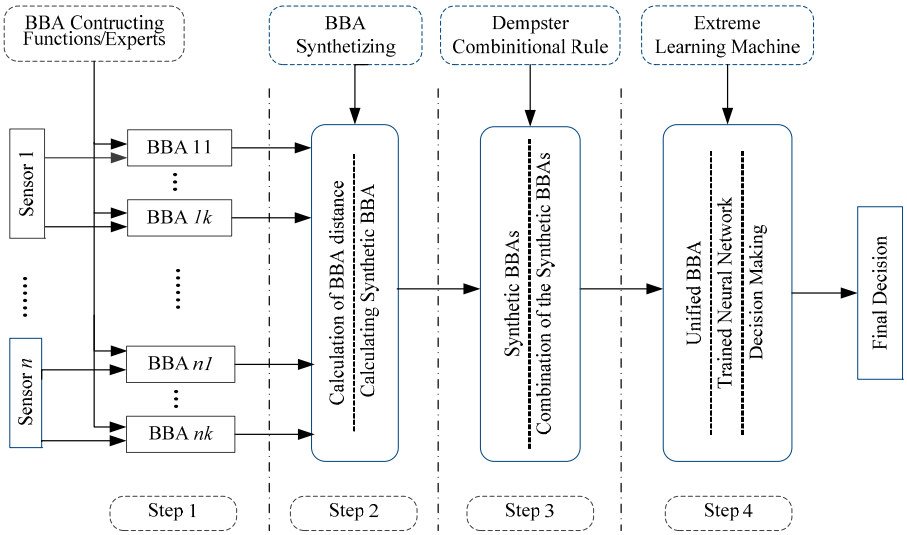
3.1. Overview

The main goal of the framework is to guarantee accurate and robust decision-level fusion results, even in situations with high complex and nonlinear data sources. Thus, robustness is the principle of the

algorithms in the entire process and it is guaranteed by three aforementioned reliable and robust measures: BBA constructing, BBA synthetizing and decision making. The process of the data fusion model is divided into four steps, as shown in Figure 1.

Let $\Omega = (\omega_1, \cdots, \omega_c)$ be the frame of discernment containing c elements, $F = \{F_1, \cdots, F_k\}$ are the BBA functions or experts, $\mathbf{S} = \{s_1, \cdots, s_n\}$ are a set of sensors. In step 1, there are k functions or experts and n sensors. Every BBA constructing function/expert will generate a BBA regarding to a sensor, hence there are $n \times k$ BBAs obtained in this step 1. Step 2 is the BBA synthesis process involving a developed synthetizing algorithm, based on which the $n \times k$ BBAs will be decreased into n synthetic BBAs. The combination calculation is conducted in step 3. Using the Dempster's combinational rule, the n synthetic BBAs will be combined into one unified BBA. Step 4 is the decision making process. With a trained ELM, the unified BBA will be transferred into the final output, which is easy for making decisions.

Figure 1. Process of the proposed evidential multisensor data fusion algorithm.



3.2. BBA Constructing Function Model

For decision fusion, local classification or decision results are essential before fusing them into a unified one. In scenarios when there is a large amount of raw data, directly uploading them to the cluster node (or sink node) is very costly. However, uploading local decision results will greatly reduce the amount of data transmission, and greatly reduce energy consumption, which is significant for distributed sensor networks, especially wireless sensor networks (WSNs). Therefore, developing a BBA constructing algorithm is necessary to obtain the local classification results.

In expert systems, BBAs are constructed based on human decisions. This paper focuses on constructing a BBA from data sources. Distance is a widely used metric measure of the similarity of an object and a class. However, each kind of distance maintains its own views. Let S be the object to be classified. The training set is $\Gamma = \{(\phi_1, \omega_1), \cdots, (\phi_c, \omega_c)\}$ and ϕ_i $(i = 1, .., c)$ is the training class respect to ω_i. There are various kinds of distances, such as Euclidean distance, Mahalanobis distance, Manhattan distance, to name a few. Here, we set k distance expressed as $= \{dis_1, \ldots, dis_k\}$, where $dis_v (v = 1, \ldots, k)$ is the vth distance definition of D.

Firstly, the distance between an object and each sample class must be calculated. The general distance can be expressed by the following expression:

$$d_{\upsilon i} = \mathrm{dis}_{\upsilon}(\| s-\phi_i \|), \upsilon = 1,...,k, i = 1,...,c \tag{9}$$

where $d_{\upsilon i}$ is the distance of S and ϕ_i according to the υth distance definition in D. Sample set ϕ_i can be the whole given sample set or the nearest m objects around the object in the given sampe set. If an object belongs to a class ω_i, then the mass should be assigned to two subsets of Ω, and they are $\{\omega_i\}$ and Ω. Then the assigned BBA $m_\upsilon(\cdot\,|\phi_i)$ can be defined as:

$$m_\upsilon(\omega_i \,|\, \phi_i) = \alpha f_\upsilon(d_{\upsilon i}), 0 < \alpha < 1 \tag{10}$$

$$m_\upsilon(\Omega \,|\, \phi_i) = 1 - \alpha f_\upsilon(d_{\upsilon i}) \tag{11}$$

$$m_\upsilon(B \,|\, \phi_i) = 0, \forall B \subset 2^\Omega \setminus \{\{\omega_i\}, \Omega\} \tag{12}$$

where m_υ is the assigned BBA of S according to the υth distance definition in D. $\alpha \in (0,1)$ is a positive constant and f_υ is a monotonically radial basic function decreasing form $f_\upsilon(0) = 1$ to $\lim_{d\to\infty} f_\upsilon(d) = 0$. It can be postulated as:

$$f_\upsilon(d_i) = \exp(-r_\upsilon(d_{\upsilon i})^2) \tag{13}$$

where γ_υ is a positive constant value. In [23], a method for optimizing parameters α and γ has been described by using of KNN method. Here we give a more exact form of γ and it is defined as:

$$\gamma_\upsilon = \frac{1}{\overline{d}_\upsilon^2} \ln 2 \tag{14}$$

$$\text{with} \quad \overline{d}_\upsilon = \frac{1}{c} \sum_{i=1}^{c} d_{\upsilon i} \tag{15}$$

where $\overline{d}_\upsilon$ is the mean value of distance between the object and each class in Ω. α can be changed to adjust the discrimination degree of the obtained BBA. After considering each pattern, the BBA function υ to be calculated as:

$$m_\upsilon = m_\upsilon(\cdot\,|\,\phi_1) \oplus \cdots \oplus m_\upsilon(\cdot\,|\,\phi_c) \tag{16}$$

where m_υ is the BBA with respect to function F_υ, $m_\upsilon(\cdot\,|\phi_c)$ is the BBA calculated by Equations (10)–(12). With the k different BBA functions, for a hypothesis A, the generated corresponding BBAs are $\{A|m_1(A), \ldots, m_k(A)\}$. Next step involves composing the synthetic BBA from the k BBAs generated by different functions.

3.3. BBA Synthetic Algorithm

Correct BBAs are the prerequisite for applying DSET in multi-source information fusion. In reality, there is no universal applicable BBA-constructing function. Different kinds of BBA functions are based on different theories, and each of them maintains its own views. For one information source, BBAs generated by different functions are always different. Dempster's combinational rule requires mutually independent evidences, and it cannot be applied to combining BBAs obtained from the same source. Hence, a new method for combining these BBAs into a comprehensive one that achieves a wide range of

aspects of uncertainty is significant for obtaining reasonable synthetic BBAs. According to reference [21], the distance of evidence can be defined as:

$$d_{BBA}(m_1, m_2) = \sqrt{\frac{1}{2}(\overrightarrow{m_1} - \overrightarrow{m_2})^T \underline{\underline{D}}(\overrightarrow{m_1} - \overrightarrow{m_2})} \tag{17}$$

where $\vec{m}_1$ and $\vec{m}_2$ are evidence vectors. D is a positively defined matrix and its coefficients can be obtained as:

$$D(A_1, A_2) = \frac{|A_1 \cap A_2|}{|A_1 \cup A_2|} \tag{18}$$

In order to transform the reliability of a function to metric, we define the credibility of a BBA as:

$$a_\upsilon = 1 - \frac{\sum_{j=1}^{k} d_{BBA}(m_\upsilon, m_j)}{\sum_{i=1}^{k}\sum_{j=l}^{k} d_{BBA}(m_i, m_j)} \tag{19}$$

where a_υ is the credibility of BBA with respect to mass function m_υ. The expression above means that a BBA with further distance will be assigned with a low credibility value and compared with other BBAs, and it is not reliable. The sum of credibility is:

$$sum(a_\upsilon) = \sum_{\upsilon=1}^{k}\left(1 - \frac{\sum_{j=1}^{k} d_{BBA}(m_\upsilon, m_j)}{\sum_{i=1}^{k}\sum_{j=l}^{k} d_{BBA}(m_i, m_j)}\right) = k - 1 \tag{20}$$

Thus, the normalized $\bar{a}_i$ can be calculated by:

$$\overline{a_\upsilon} = \frac{1}{k-1} a_\upsilon = \frac{1}{k-1}\left[1 - \frac{\sum_{j=1}^{k} d_{BBA}(m_\upsilon, m_j)}{\sum_{i=1}^{k}\sum_{j=l}^{k} d_{BBA}(m_i, m_j)}\right] \tag{21}$$

The sum of $\bar{a}_i$ is $\sum_{i=1}^{k} \bar{a}_i = 1$. Then, the synthetic BBA can be calculated by the following weighted sum method:

$$m(A) = \sum_{\upsilon=1}^{k} \overline{a_\upsilon} m_\upsilon(A), \upsilon = 1, \ldots, k \tag{22}$$

where $m(A)$ denotes the hybrid BBA with respect to hypothesis A, $\bar{a}_i$ is the weight of BBA with respect to function F_i. The value $\bar{a}_i$ shows the reliability of the corresponding function. A believable function will be assigned a larger credibility compared to other functions. The final obtained BBA will be regarded as the local decision and uploaded to the fusion center.

3.4. Combination of the Synthetic BBAs

The Dempster's combinational rule has been widely accepted as a method for combining various evidences into a unified one. With the obtained *n* evidences, the unified evidence can be defined as:

$$(m_1 \oplus \cdots \oplus m_n)(A) = \frac{1}{1-K} \sum_{A_1 \cap \cdots \cap A_n = A} m_1(A_1) \cdots m_n(A_n) \tag{23}$$

$$\text{with} \quad K = \sum_{A_1 \cap \cdots \cap A_n = \varnothing} m(A_1) \cdots m(A_n) \tag{24}$$

where K denotes the conflict degree, final unified BBA contains the comprehensive knowledge of all information sources. Note that other combinational rules, including the method proposed in Section 3.3., can be applied instead of the Dempster's combinational rule.

3.5. ELM Based Decision Making Model

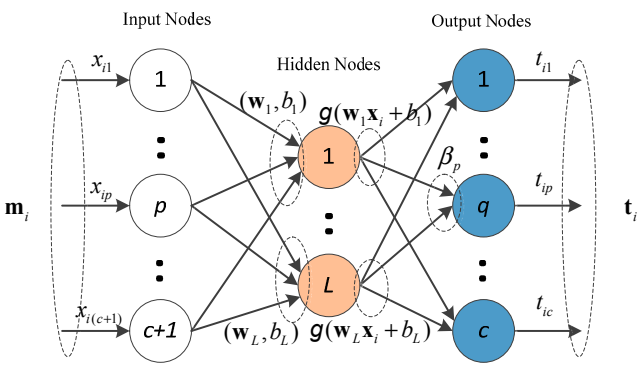
Unlike the traditional TBM transfer mechanism, this paper uses a decision-making algorithm based on ELM. Huang and his colleagues [24] proposed extreme Learning Machine (ELM). This single feed-forward neural network (SLFN) with a fast learning speed has both universal approximation and classification capabilities [25]. In ELM, the input weights of the hidden neurons can be generated randomly, and they are independent from applications. Thus, ELM requires no iterative calculation to determine the input weights, which is a big advantage of the traditional artificial neural networks.

As shown in Figure 2, ELM can be directly applied to conduct the decision making process. For N arbitrary distinct samples $(\mathbf{m}_1, \mathbf{t}_1)(i = 1, \ldots, N)$, where input data $\mathbf{m}_i = (m_i(\omega_1), \ldots, m_i(\omega_c), m_i(\omega_\Omega)) \in R^{c+1}$ and output data $\mathbf{t}_i = (t_{i1}, \ldots, t_{ic}) \in R^c$, the output of the network with L hidden neural nodes can be expressed as:

$$\mathbf{t}_i = \sum_{j=1}^{L} \beta_j g(\mathbf{w}_j \mathbf{m}_i + b_j), i = 1, \ldots, N \tag{25}$$

where $\mathbf{t}_i$ is the network output of $\mathbf{X}_i$, $\mathbf{w}_j = [w_{j1}, \ldots, w_{j(c+1)}]^T$ is the input weight matrix between the input neural node and the jth hidden neural nodes, $\beta_j = [\beta_{j1}, \ldots, \beta_{jc}]^T$ is the output weight matrix between the ith hidden nodes and output nodes, b_j is the bias threshold respect to the ith hidden node. $\mathbf{w}_j$ and b_j are generated randomly and are independent from any specific applications.

Figure 2. Decision making network of ELM.



Let $g(x)$ be the activation function. The above N equations can be written as:

$$\mathbf{H}\beta = \mathbf{T} \tag{26}$$

with:

$$\mathbf{H}=\begin{bmatrix} g(\mathbf{w}_1\mathbf{x}_1+b_1) & \cdots & g(\mathbf{w}_L\mathbf{x}_1+b_L) \\ \vdots & \cdots & \vdots \\ g(\mathbf{w}_1\mathbf{x}_N+b_1) & \cdots & g(\mathbf{w}_L\mathbf{x}_N+b_L) \end{bmatrix}_{N\times L}, \beta=\begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_2^T \end{bmatrix}_{L\times m} \text{ and } \mathbf{T}=\begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_N^T \end{bmatrix}_{N\times m} \tag{27}$$

where $\mathbf{H}$ is hidden neural output matrix. The output matrix β can be calculate by:

$$\beta=\mathbf{H}^{\dagger}\mathbf{T} \tag{28}$$

where $\mathbf{H}^{\dagger}$ is the Moore–Penrose generalized inverse of matrix $\mathbf{H}$. To improve the robustness of the generalization performance, the above expression can modified as [22]:

$$\beta=(\frac{\mathbf{I}}{C}+\mathbf{H}^T H)^{-1}\mathbf{H}^T\mathbf{T} \tag{29}$$

The ELM algorithm can be designed in 3 steps: (1) Assign the input weight matrix $\mathbf{W}$ and bias b randomly; (2) Calculate the output matrix $\mathbf{H}$ of the hidden neural nodes; (3) Calculate the output weight matrix β. After training the ELM, it is able to perform specific functions, like approximation and regression.

Given N_o new observed unified BBAs with masses $\mathbf{m}_o=\{m_{o1},\dots,m_{oNo}\}$ and the respective m-th one is $\mathbf{m}_{om}, m=1,\dots,N_o$, the output of the ELM is:

$$\mathbf{t}_o=g(\mathbf{w}\cdot\mathbf{m}_o+b)\beta \tag{30}$$

where $\mathbf{t}_\mathrm{o}$ is the output matrix and $\mathbf{t}_\mathrm{o}=[\mathbf{t}_{\mathrm{o}1},\dots,\mathbf{t}_{\mathrm{o}2}]$. Then the decision can be made by:

$$\omega_{Dg}=\max(\mathbf{t}_{og}), g=1,\dots,c \tag{31}$$

where ω_{Dg} is the final decision with respect to g-th column in $\mathbf{t}_\mathrm{o}$. Note that other ANN algorithms, like BP and RBF, also can be applied in this step, the decision making policy is the same as ELM.

4. Experimental Results

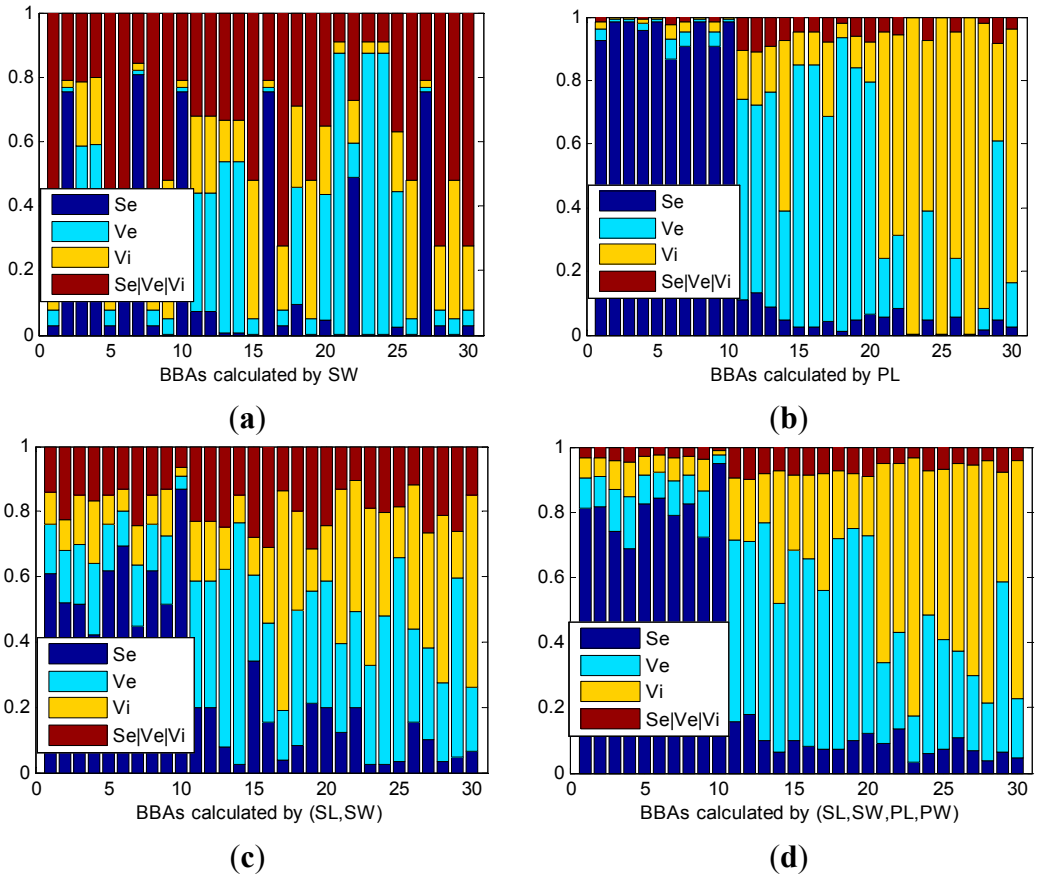
This section reviews the experiments that are performed to test the performances of the proposed data fusion algorithm. In the foregoing experiments, we simulate our model in three steps: First, we use the IRIS dataset to illustrate the performance of the proposed mass construction algorithm. Second, we use the Diabetes Database dataset to train our model, and then, we collect data from the people whose age range from 40 to 60 by human body sensors, and predict their health condition. In addition, the last experiment applied the proposed framework in vehicle type classification. Introductions about these tests and their corresponding results will be described in the following sections.

4.1. Experiment on IRIS Data Set

In this experiment, we use the IRIS data collected by statistician Fisher [26] to simulate the algorithm. There are three species of Iris in this data set: Setosa (Se), Versicolor (Ve), and Virginica (Vi), and each plant includes four indices: sepal length (SL), sepal width (SW), petal length (PL), and petal width (PW). The total number of data is 600 (150×4), or 50 patterns for each plant. To classify the plants, each of the four features represents an information source. Among the 50 patterns of each plant, 30 patterns are randomly selected as the training sets, the remaining 20 patterns of each plant are the testing sets.

Four kinds of distances are used in the experiment: Euclidean Distance (Eu), Mahalanobis Distance (Mahal), Chebyshev distance (Ch), and Manhattan Distance (Ma). Different functions generate BBAs with different accuracies. The test data sets include seven patterns, which are SL, SW, PL, PW, (SL, SW), (PL, PW), and (SL, SW, PL, PW). The BBAs are calculated using Eu, as shown in Figure 3.

Figure 3. Partial BBAs calculate by the Eu-function includes (**a**): training set SW; (**b**): training set PL; (**c**) training set (SL, SW); (**d**) training set (SL, SW, PL, PW).



After obtaining all BBAs from different dimensional data set and BBA functions, the synthesized BBAs of the four distance BBA functions can be obtained. To examine the rationality and validity of the proposed functions, the accuracies are calculated in every step of the BBA constructing process. The accuracy degree of a BBA is calculated using the following expression:

$$r(\omega_i) = \frac{N_y(m(\omega_i) = \max\{m(\omega_1),...,m(\omega_c)\})}{N} \times 100\%, i = 1,...,c \tag{32}$$

where $r(\omega_i)$ denotes the accuracy rate with respect to class ω_i, N is the total BBA number of the test sets. N_y is the number of accurate BBAs. If the object to be classified belongs to ω_i, $m(\omega_i)$ should be larger compared to any other singleton classes. The accuracies of the results are calculated by the 7 source data set. The accuracies of the BBAs for different sources and BBA functions are shown in Table 1.

Table 1. Accuracies of the BBAs obtained for the IRIS data (%).

	SL	SW	PL	PW	SL, SW	PL, PW	All
Eu-BBA	73.33	58.33	96.67	96.67	83.33	96.67	98.33
Mahal-BBA	80.00	58.33	96.67	96.67	86.67	96.67	98.33
Ch-BBA	73.33	58.33	96.67	96.67	83.33	96.67	95.00
Ma-BBA	73.33	58.33	96.67	96.67	81.67	95.00	96.67
Syn-BBA	73.33	58.33	96.67	96.67	86.67	96.67	96.67

Figure 3 shows partial BBAs calculated using the Euclidean Distance. The horizontal axis denotes the number of test objects and vertical axis represents the mass assignments of each object. The sum of each mass equals 1. The first 10 objects are Se, next 10 objects are Ve, and the last 10 objects are Vi. In Table 1, the syn-BBA denotes the BBAs synthesized from the four BBAs with different distance definitions. We make the following observations:

(1) The proposed mass construction method is able to build BBAs from observed data and information accurately and effectively. With a distinguishable data set, the mass of the compound classes will be much lower compared to the sum of the singleton classes, and the belief assignment with respect to the class to which the object belongs will always be much larger compared to other classes. As shown in (b), (c), (d) and Table 1, the accuracies of PL, (SL, SW), and (SL, SW, PL, PW) calculated by the Eu are 96.67%, 83.33%, and 98.33%, respectively. While given an ambiguous data set, as shown in (a), the masses of each object will likely to be confusing, significantly decreasing the accuracy rate and belief assignments of Ω (Se|Ve|Vi), with the accuracy being only 58.33%.

(2) Higher dimensionality data enhances the accuracy of BBAs. With the same BBA function in (a), the BBAs obtained from SW have low accuracy (58.3%), while the accuracy is much higher in (c) (83.33), where the BBAs are calculated by (SL, SW). In (d), the BBAs' accuracy is 98.33% and the dimension is 4. In a more dimensional space, the boundaries of the plants can be classified more clearly. Generally, higher dimensional data brings more accurate BBAs, as the BBAs accuracies of (SL, SW, PL, PW) are higher compared to others, except for the Ch-BBA function.

(3) The synthesized BBA is able to comprehensively illustrate the data source. When assigning belief for an information source, different methods hold their own views and their results may different, too. Thus, the method of synthesizing different BBAs into a unified comprehensive BBA is able to get that the result reflecting the views of the majority.

4.2. Experiment on Diabetes Data Set

The Pima Indians Diabetes Database (available at [27]) was developed at the Applied Physics Laboratory, Johns Hopkins University. The eight indices in the data represent the diagnostic signs of diabetes according to World Health Organization criteria. The database comprises the data from 768 women over the age of 21 residing in Phoenix (Arizona, USA). All examples belongs to either positive (denotes by 1) or negative (denotes by 0) class. All input values are within [0, 1]. To test the

proposed method, 75% (576) and 25% (192) samples are chosen randomly for training and testing at each trial, respectively.

4.2.1. Experimental Results with Changing α

To get a better understanding of the proposed algorithm, additional experiments are conducted. In another trial, we set different value of α to find out its influences on BBA constructing and final result accuracy. The object is selected randomly from the test data set. Three classes of the power set are $\{\text{diagnosis}, \text{Not Diagnosis}, \text{All}\}$, where 'All' denotes the compound set. Parameter α is set monotonically, increasing from 0 to 1. The values are used to calculate the corresponding BBAs. The BBA obtained with different α is shown in Figure 4. The accuracies of final unified BBA and final decision are shown in Figure 5.

Figure 4. BBAs obtained with different α.

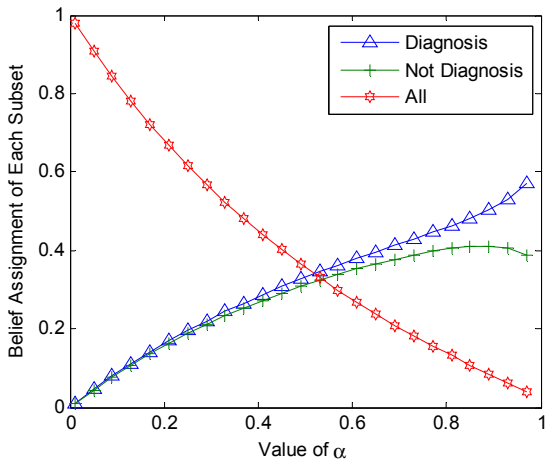
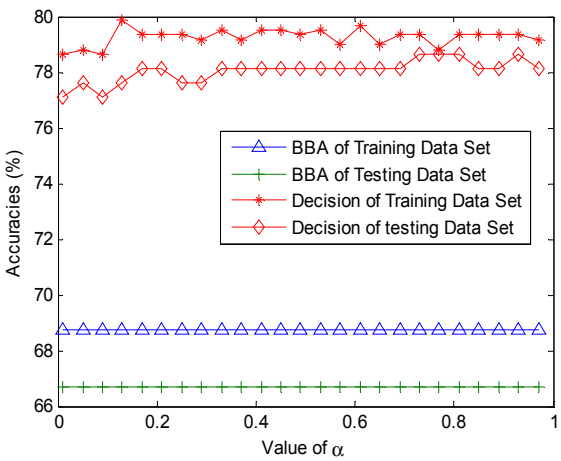


Figure 5. Accuracies with different α.



In Figure 4, the BBA is obtained from the same object belonging to the diabetes diagnosis class. In Figure 5, the accuracy rate of BBA and decisions are both determined based on the training data set and testing data. From Figures 4 and 5, we can make the following conclusions:

(1) Parameter α will change the belief assignments of each subset in 2^{Ω}. When α is closer to 0, $m(\Omega) \approx 1$ and the belief assignments of singleton class are close to 0. With an increasing α, $m(\Omega)$ decreases to a very low level while the belief assignments of singleton class increases to high levels. The gap between them will gradually diminish. However, it is strongly advised to set $\alpha > 0.7$ to get a high differentiation degree for the BBA.

(2) Parameter α has no influence on the average accuracy of the BBA and decisions. In Figure 4, the BBA is larger for 'Diagnosis' compared to 'Not Diagnosis'', regardless of the value of α. In Figure 5, the accuracies of unified BBAs in training data set and testing data set are 68.7 and 66.7, respectively. The accuracies of final decision results in training data set and testing data set are about 79% and 78%, respectively. Note that the decision accuracy fluctuation is caused by the instability of ELM. The stable accuracy rates illustrate that α has no influence on the accuracy of the BBA and decision accuracy.

4.2.2. Experimental Results of Accuracies

Many algorithms and methods, such as BP neural network [28], Support Vector Machine (SVM) [29], ELM [22–24], and others, can use the database to get the classification results. To obtain a clear picture of the performance, we compare different algorithms, including BPNN, SVM, ELM, evidential data fusion with Pignistic transfer method (DSET-P) and the proposed DSET-E.

In this test, the parameter *C* of SVM algorithm is set at 10, and its accuracy results are obtained with 317 support vectors in average. All hidden nodes of the BPNN and ELM are 20. The DSET-P and DSET-ELM use the same process of calculating the unified BBA, thus their final BBAs are the same. α is set at 1 and the BBA function uses only the Mahalanobis distance because we have found that it has a high accuracy in constructing masses. In DSET -P, the unified BBA is converted to probability using the Pignistic transferring method in [19]. While in DSET-E, unified BBA is the input of a trained ELM, which is used to transfer the BBA to results and make decisions.

As shown in Table 2, all accuracies are calculated by the average accuracy results of repeating 100 times. From Table 2, we have the following observations:

(1) The proposed DSET-E algorithm performed well in classifying problems. Compared with other algorithms, the proposed DSET-E algorithm obtains a testing rate of 78.14%, outperforming other methods, though the improvement is not sufficiently distinctive. Compared with the DSET-P method, the accuracy increases from 66.67% to 78.14%, which is sufficient to prove that the whole algorithm is reasonable and effective.

(2) The traditional Pignistic transferring method is not a desirable algorithm in evidential data fusion, especially in situations with high complex and nonlinear data sources. The accuracy of DSET-P is 66.7%, which is much lower compared to the accuracies of other methods. Actually, in this problem, the highly complex source data are difficult to distinguish, and the final unified BBA has a low accuracy when calculated by Equation (30). The final decision made by Pignistic probability transferring model has the same accuracy as the unified BBA, which is 66.67%. With the same final unified BBA, the DSET-E decision accuracy rate of the test objects increases by 11.74%.

(3) The belief assignments of the compound classes are also important in decision-making. In traditional Pignistic probability transferring model, the belief assignments of a compound class are carved up by proportion, which makes no difference in decision-making. Apparently, it is not suitable for all conditions. The belief assignments of a compound class show uncertainty, making it difficult to decide to which class it should belong. It should be allocated to other singleton classes according to the reality situations.

Table 2. Accuracy comparison for DIABETES data set (average, %).

Algorithms	Training Accuracies	Testing Accuracies
DSET-E	79.39	78.14
DSET-P	-	66.67
ELM	78.68	77.57
SVM	78.76	77.31
BPNN	86.63	74.73

4.3. Experiment on Vehicle Type Classification Data Set

In this experiment, a data set for vehicle type classification (the data set can be downloaded at [30]) is used to test the proposed algorithm. In the test, 23 wireless distributed sensor nodes are used to classify the types of the vehicles. When a vehicle is passing by, the nearby sensor nodes are able to record the signals in three modalities: acoustic, seismic and infrared. We use the recorded acoustic and seismic signals to classify two possible vehicles: Assault Amphibian Vehicle (AAV) and DragonWagon (DW). Before classification, feature vectors must be extracted from the raw signals. A detailed introduction of the feature extraction method can be found in [31].

The experiment includes two parts: part one is the classification based on the whole data set. In this scenario, universal classification algorithms can be directly used and their classification results will be presented; part two is the classification conducted in a distributed multisensor data fusion way. Experimental results of local classification and data fusion will also be presented in the following sections.

In the first test, five classification methods are used, including k-NN, ELM, SVM, DSET-P and DSET-E. The sample set is consisted by 535 feature vectors, which are randomly selected from the provided whole feature data set. Among the sample set, 277 feature vectors belong to vehicle AAV, the rest are DW. The valid data set has 236 feature vectors, which are also randomly selected from the whole feature data set. The classification results are given in the following Table 3.

Table 3. Classification results of different methods.

Algorithm	AAV	DW	Average
k-NN	70.11	70.57	70.32
ELM	73.90	78.28	76.06
SVM	62.51	**81.55**	71.64
DSET-P	73.54	76.22	74.93
DSET-E	**76.01**	76.79	**76.62**

As shown in Table 1, five classification algorithms are used for local classification. The parameter k in k-NN method, hidden nodes number of ELM and parameter C of SVM are set as 15, 100 and 1, respectively. The mass construction used in DSET-P and DSET-E is the method proposed in Section 3.2 and parameter α is set as 0.85. From Table 3, we can conclude that the proposed DSET-E has a more reliable result than other methods, which is consistent with the results of Table 2.

Then we conducted the task in a multisensor data fusion scenario, in which each sensor node has its own sample data set collected by itself. Since the energy and bandwidth of wireless sensors are strictly limited, uploading the raw data to the sink node is unpractical. Therefore, a local classification in the sensor node needs to be conducted and then these local results are uploaded to the fusion center for final decision by data fusion algorithms. Except for DSET-P and DSET-E, the algorithms used in Table 3 cannot be used for classification. The local classification accuracies the final data fusion accuracies are shown in Tables 2 and 3, respectively.

As shown in Tables 4 and 5, there are 11 sensor nodes used for the collaborative data fusion task. Both test set have 1177 vector samples, in which 615 vectors belong to AAV, the other 562 vectors are DW. From Tables 3 to Table 5, it can be concluded that the proposed method has good performance of in multisensor data fusion applications, because it is able to get reliable and robust results. In Tables 2 and 3, the accuracies of DSET-E are always higher than the accuracies of other classification methods. In Table 4, the average classification accuracies of AAV and DW are 68.13% and 59.83%, respectively. However, the fusion results are greatly improved by both the DSET-P and DSET-E methods. The final average accuracy of DSET-E is 75.32%, which increased by 2.52% compared to DSET-P. And also, the average accuracy of DSET-E is close to the results of DSET-E in Table 3, which equals to 76.62%. These results also show that the proposed BBA constructing function is reasonable and effective for the DSET based data fusion framework.

Table 4. Local classification accuracy of each sensor node (%).

Node	1	2	3	4	5	6	7	8	9	10	11	Average
AAV	56.42	53.04	85.35	65.31	75.29	85.43	62.52	63.52	70.24	66.57	65.73	68.13
DW	72.12	43.37	45.67	62.81	48.70	48.57	75.80	64.54	63.38	62.55	70.64	59.83

Table 5. Final data fusion accuracies (average, %).

	AAV	DW	Average
DSET-P	79.84	65.50	72.80
DSET-E	82.41	67.57	75.32

To conclude the experimental section, the three tests prove that the proposed mass constructing and decision making method is reasonable and practical. The classification results with other universal classification methods (*i.e.*, k-NN, BPNN, ELM and SVM) prove that the proposed method is able to obtain robust and reliable results. The experiment on vehicle type classification demonstrates that the proposed method has high performance in multisensor data fusion applications, but not only in practice for conventional problems. Therefore, the proposed model is robust and reliable in mutisensor data fusion.

5. Discussion and Conclusions

In conclusion, this paper proposed a systematic multisensor data fusion model to obtain robust and high-precision fusion results. DSET is used to provide a flexible way to combine multiple information sources into a unified one, and ELM is applied to make decisions. The combination of the two theories achieves a greater capacity in multisensor data fusion. Compared with the existing methods, the proposed framework gives more flexibility and rationality in constructing reasonable BBAs from data sources. Additionally, the framework is able to make decisions according to the actual situation. Moreover, it is stable and easy to implement. With adequate training samples, the algorithm is able to reason and learn and make decisions in a coherent process. The drawback is that it needs to be trained and the computation complexity is greater than that of traditional DSET-P, though ELM is 'extremely' fast in ANNs. However, it should be clear that the proposed method is not intended to achieve great improvement over other classification algorithms, but rather it is aimed at building a robust and reliable data fusion model practical for multisensor applications. Thus, the accuracy improvement is not the key point of the proposed model. Though the improvement of training accuracy and testing accuracy is not significant, the results still prove that the proposed is robust and reliable.

It is necessary to emphasize that in Section 3.3, 'minority is subordinate to the majority' underlies the synthetizing algorithm, which is useful when the performance of the adopted BBA functions or experts are unknown to us. It can be modified as 'outstanding is preferred', which means the weights of the BBA functions or experts are assigned according to their own accuracies. In Section 3.5, the activation function is the 'radbas' function. Other function types are also feasible. In Section 4.2, the experimental results indicate that a decision-making method exists after the combinational step, although it is not the Pignistic transferring method. If we could improve the conventional Pignistic transfer method, DSET could be greatly promoted in real applications. Future work may involve the following: (1) discovering a new decision-making method to get rid of the low accuracy limitation of the existing Pignistic methods; (2) developing a DSET-embedded ELM that is able to deal with pattern recognition or classification problems; and (3) exploring more inherent laws in the transducer mechanism and probability.

Acknowledgements

This research is supported by National Natural Science Foundation under Grant 61371071, Beijing Natural Science Foundation under Grant 4132057, Beijing Science and Technology Program under Grant Z121100007612003, Academic Discipline and Postgraduate Education Project of Beijing Municipal Commission of Education.

Author Contributions

In this paper, the idea and primary algorithm were proposed by Bo Shen. Bo Shen, Yun Liu and Jun-Song Fu conducted the simulation and analysis of the paper together.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Khaleghi, B.; Khamis, A.; Karray, F.O.; Razavi, S.N. Multisensor data fusion: A review of the state-of-the-art. *Inf. Fusion* **2011**, *14*, 28–44.
2. Al Momani, B.; Morrow, P.; McClean, S. Fusion of Elevation Data into Satellite Image Classification Using Refined Production Rules. In *Image Analysis and Recognition*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 211–220.
3. Lelandais, B.; Gardin, I.; Mouchard, L.; Vera, P.; Ruan, S. Using Belief Function Theory to Deal with Uncertainties and Imprecisions in Image Processing. In *Belief Functions: Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 197–204.
4. Deng, Y.; Chan, F.T.S. A new fuzzy Dempster MCDM method and its application in supplier selection. *Expert Syst. Appl.* **2011**, *38*, 9854–9861.
5. Guerriero, M.; Svensson, L.; Willett, P. Bayesian data fusion for distributed target detection in sensor networks. *IEEE Trans. Signal Process.* **2010**, *58*, 3417–3421.
6. Du, P.; Liu, S.; Xia, J.; Zhao, Y. Information fusion techniques for change detection from multi-temporal remote sensing images. *Inf. Fusion* **2013**, *14*, 19–27.
7. Zimmermann, H J. *Fuzzy Set Theory—And Its Applications*; Springer: Berlin/Heidelberg, Germany, 2001.
8. Qian, Y.; Li, S.; Liang, J.; Shi, Z.; Wang, F. Pessimistic rough set based decisions: A multigranulation fusion strategy. *Inf. Sci.* **2014**, *264*, 196–210.
9. Shafer, G. *A Mathematical Theory of Evidence*; Princeton University Press: Priceton, NJ, USA, 1976.
10. Yager, R.; Fedrizzi, M.; Kacprzyk, J. *Advances in the Dempster-Shafer Theory of Evidence*; John Wiley & Sons, Inc.: New York, NY, USA, 1994.
11. Zhang, Z.; Liu, T.; Zhang, W. Novel Paradigm for Constructing Masses in Dempster-Shafer Evidence Theory for Wireless Sensor Network's Multisource Data Fusion. *Sensors* **2014**, *14*, 7049–7065.
12. Chakeri, A.; Nekooimehr, I.; Hall, L.O. Dempster-Shafer Theory of Evidence in Single Pass Fuzzy C Means. In Proceedings of the 2013 IEEE International Conference on IEEE Fuzzy Systems, Hyderabad, Indian, 7–10 July 2013; pp. 1–5.
13. Ben Chaabane, S.; Fnaiech, F.; Sayadi, M.; Brassart, E. Estimation of the Mass Function in the Dempster-Shafer's Evidence Theory Using Automatic Thresholding for Color Image Segmentation. In Proceedings of the 2nd International Conference on IEEE Signals, Circuits and Systems, Philadelphia, PA, USA, 3–5 November 2008; pp. 1–5.
14. Zhu, H.; Basir, O. A Scheme for Constructing Evidence Structures in Dempster-Shafer Evidence Theory for Data Fusion. In Proceedings of the 2003 IEEE International Symposium on IEEE Computational Intelligence in Robotics and Automation, Kobe, Japan, 16–20 July 2003; Volume 2, pp. 960–965.
15. Basir, O.; Karray, F.; Zhu, H. Connectionist-based Dempster-Shafer evidential reasoning for data fusion. *IEEE Trans. Neural Netw.* **2005**, *16*, 1513–1530.
16. Zhu, H.; Basir, O. A novel fuzzy evidential reasoning paradigm for data fusion with applications in image processing. *Soft Comput.* **2006**, *10*, 1169–1180.

17. Smets, P.; Kennes, R. The transferable belief model. *Artif. Intell.* **1994**, *66*, 191–234.
18. Smets, P. Decision making in the TBM: The necessity of the pignistic transformation. *Int. J. Approx. Reason.* **2005**, *38*, 133–147.
19. Sudano, J. Pignistic Probability Transforms for Mixes of Low- and High-Probability Events. Available online: http://www.incose.org/delvalley/FUSION2001.pdf (accessed on 21 October 2014).
20. Chen, S.; Deng, Y.; Wu, J. Fuzzy sensor fusion based on evidence theory and its application. *Appl. Artif. Intell.* **2013**, *27*, 235–248.
21. Jousselme, A.-L.; Grenier, D.; Bossé, E. A new distance between two bodies of evidence. *Inf. Fusion* **2001**, *2*, 91–101.
22. Huang, G.B.; Zhou, H.; Ding, X.; Zhang, R. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2012**, *42*, 513–529.
23. Zouhal, L.M.; Denœux, T. An evidence-theoretic k-NN rule with parameter optimization. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **1998**, *28*, 263–271.
24. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501.
25. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. In Proceedings of the IEEE International Joint Conference on Neural Networks, Budapest, Hungary, 25–29 July 2004; pp. 985–990.
26. Fish, R.A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **1936**, *7*, 179–188.
27. Pima Indians Diabetes Data Set. Available online: http://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes (accessed on 10 September 2014).
28. Gupta, C.N.; Palaniappan, R.; Swaminathan, S.; Krishnan, S.M. Neural network classification of homomorphic segmented heart sounds. *Appl. Soft Comput.* **2007**, *7*, 286–297.
29. Bazi, Y.; Melgani, F. Toward an optimal SVM classification system for hyperspectral remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3374–3385.
30. Matlab Toolboxes. Available online: http://www.ecs.umass.edu/~mduarte/Software.html (accessed on 15 September 2014).
31. Duarte, M.F.; Hen, H.Y. Vehicle classification in distributed sensor networks. *J. Parallel Distrib. Comput.* **2004**, *64*, 826–838.