Multi-Attribute Fusion Algorithm Based on Improved Evidence Theory and Clustering

Wenqing Wang 1, Yuan Yan 1,*, Rundong Zhang 2, Zhen Wang 1, Yongqing Fan 1 and Chunjie Yang 1

1 School of Automation, Xi’an University of Posts and Telecommunications, Xi’an 710121, China; wwq@xupt.edu.cn (W.W.); wangzhen9623@163.com (Z.W.); fanyongqing@xupt.edu.cn (Y.F.); 13772180915@163.com (C.Y.)
2 School of Automation, Central South University, Changsha 410083, China; zrd1018@csu.edu.cn

* Correspondence: 18700907589@163.com

Received: 8 August 2019; Accepted: 23 September 2019; Published: 25 September 2019

Abstract: In most of the application scenarios of industrial control systems, the switching threshold of a device, such as a street light system, is typically set to a fixed value. To meet the requirements for a smart city, it is necessary to set a threshold that is adaptive to different conditions by fusing the multi-attribute observations of the sensors. This paper proposes a multi-attribute fusion algorithm based on fuzzy clustering and improved evidence theory. All of the observations are clustered by fuzzy clustering, where a proper clustering method is chosen, and the improved evidence theory is used to fuse the observations. In the experiments, two-dimensional observations for the street light illumination and for the ambient illumination are used in a campus-intelligent lighting system based on a narrowband Internet of things, and the results demonstrate the effectiveness of the proposed fusion algorithm. The proposed algorithm can be applied to a variety of multi-attribute fusion scenarios.

Keywords: data fusion; fuzzy clustering; evidence theory

1. Introduction

Multi-attribute fusion is a fusion method involving data from multiple attributes of many sensors, in order to obtain more accurate and reliable conclusions. Many scholars have studied this issue, and have even applied this algorithm to numerous aspects of life. An opinion of multi-attribute information fusion was proposed in order to improve the accuracy of sensor networks. An important part of the unmanned parking-lot system designed by the authors of [1] is license plate recognition. A background fill light is installed under the camera, which can enhance the capture effect of the camera in a dim environment. The main role of the background fill light is to improve the brightness of the capture environment and to make the picture clearer. If the illuminance meter can be used to measure the light source illuminance and the ambient illuminance of the background fill light many times, so as to fuse a switch value, it is possible to accurately determine whether a background fill light needs to be turned on and captured at a certain time. In the literature [2–4], the intelligent road-mounting system detected and recognized the license plate, model, color, facial image of the people in the vehicle, and the fill light technology of the camera during capture. The outstanding effect of the LED (Light Emitting Diode) fill light in the work of the authors of [5] showed the growth and development of high-quality vegetable seedlings in low light conditions, light intensities, and photoperiods. It is shown in the literature [6] that an intelligent lighting system for an exhibition hall network, which may be a scene illuminance sensor installed according to the moment of sunlight illumination change, sets a different direct light brightness for exhibits when visitors approach, and also sets up the illumination according
to the peak and trough time of the audience entering. Based on multi-dimensional data fusion, a power transformer fault diagnosis system showed the algorithm efficiency in the study [7]. The fuzzy clustering algorithm of the authors of [8] simulated the hydrological process without the relevant data, and better reflected the process of the runoff variation within the error range. The research by the authors of [9] was also prominent. The bias correction term adjusted the influence of initialization on the fuzzy clustering algorithm.

The above algorithms [1–8] were based on real-datasets and information fusion, and these algorithms assumed, directly or indirectly, that all observations are subject to normal distributions with the same parameter. All of the observations were located near a true value with the same weight. However, because of the inevitable system error of the sensors, the measuring instruments, human factors, and environmental interference, the observations of the sensors are located around several approximate true values; these approximations are considered to be the clustering centers, and different sensors obey normal distributions in a small range around an approximate true value. All of the observations can be clustered according to the deviation of the observations from the true values. Each category can be assigned a different weight. Evidence combination is the indispensable procedure in the improved evidence theory data fusion algorithm. This method can improve the accuracy and credibility of fusion, and this is the role of data fusion.

The switching threshold of traditional control devices is generally set to a fixed value, which is simple and easy to imply. However, by fusing the data information of multiple attributes to obtain a more precise switching threshold, it is beneficial by saving energy, reducing emissions, and refine the automatic control of the device, promoting the process of a smart city. To this end, based on the idea of fuzzy clustering and evidence theory, this paper proposes a new multi-attribute observation fusion algorithm. The DS evidence theory has steps of grouping the observations before the data fusion. As the fuzzy c-means (FCM) algorithm needs to specify the number of cluster centers in advance, we set the number of cluster centers as two, three, and four, and the data sets were divided into two categories, three categories, and four categories, respectively. The DS evidence theory was used on every clustering result in order to obtain the most accurate data for the fusion results. This data can be considered as the switching threshold of the device for a period of time. This threshold can provide more abundant information than single one-dimensional data, so as to achieve the purpose of improving the fusion precision. The algorithm does not need any prior information or the historical data of the sensors, nor does it need to assume that sensor observations obey the normal distribution of the same parameter [10,11]. This algorithm has the characteristics of a wide application range, simplicity, and a high reliability of the fusion results.

As such problem settings and application scenarios are quite common, the method can also be applied to other scenarios that require multi-attribute data fusion, and two-dimensional attributes are extended to multidimensional attributes, such as traffic control, battlefield situation estimation, target classification and tracking, smart cities, and agriculture.

2. Algorithm Description

2.1. Fuzzy C-Means Clustering Algorithm

Research on clustering algorithms has achieved many results [12,13]. Comparing the K-means algorithm, hierarchical clustering algorithm, self-organizing maps (SOM) neural network clustering algorithm, and fuzzy c-means (FCM) clustering algorithm, it has been found that although the hierarchical clustering does not need to determine the category number, once a segmentation or merging was performed, it could not be corrected. The clustering quality was limited [14]. The SOM neural network has a longer processing time, because its mapping has a topology-like retention characteristic, similar to the brain nervous system. FCM is a clustering algorithm based on data division; this algorithm was thought to be an improvement to the hard c-means (HCM) algorithm [15]. The FCM algorithm obtains the membership of each sample to all of the category centers by optimizing
the objective function. It can determine generics of the samples so as to achieve the purpose of grouping
the observations automatically, and its accuracy is higher than the classical K-means clustering
algorithm. Therefore, the FCM algorithm was selected for the experiment.

In this paper, 100 sets of observations with two attributes were tested by measuring the
ambient illuminance and light illuminance of street lights, based on a smart campus lighting system.
These observations were clustered by an FCM algorithm. The idea of the algorithm has been described
in detail [16–20].

It divides 100 sets of observations \(x_n(n = 1, 2, \ldots, 100)\) into categories. Each sample is not
strictly divided into a certain category, but belongs to a category in a certain membership grade.
The division can be stated by the membership matrix, \(U = [u_{ik}]_{a \times 100}\). \(u_{ik}\) shows the membership
grade of \(i (i = 1, 2, \ldots, 100)\) to \(k (k = 1, 2, \ldots, a)\). \(v = (v_1, v_2, \ldots, v_a)\) is the center of the clustering,
and \(d_{ik} = \|x_i - v_k\|\) is the Euclid distance between the samples and each clustering center. FCM divides
the objective function as follows:

\[
\min J(U, v) = \sum_{i=1}^{100} \sum_{k=1}^{a} u_{ik}^h d_{ik}^2
\]

(1)

where the parameter \(h > 1\) is a weighted index to control the fuzzy degree of the membership matrix,
\(U\). \(h\) will be fuzzier when it gets bigger. When \(h = 1\), the fuzzy clustering will decline to the
HCM clustering.

For the FCM algorithm, the algorithm has to specify the number of categories in advance,
initialize \(a\) and \(h\) here, \(a = 2, 3, 4\) and \(h = 2\). \(\epsilon = 1 \times 10^{-4}\) is the accuracy of the cluster center. The initial
cluster centers were generated randomly. The membership matrix can be calculated according to
Equation (3). Secondly, we used Equation (2) to adjust the cluster centers and categories. Finally, in
the light of the termination condition, we used \(\epsilon = 1 \times 10^{-4}\) to determine whether the accuracy of the
cluster centers met the requirement. We recalculated the membership matrix and cluster centers if the
termination condition was not met by the iteration. The clustering effect of the different categories of
these 100 samples were different.

Using the Lagrange multiplier method, the necessary conditions to minimize Equation (1) were
as follows:

\[
u_{ik} = \left( \sum_{j=1}^{a} \left( \frac{\|x_j - v_k\|^2}{\|x_j - v_k\|^2 + \|x_i - v_k\|^2} \right)^{2 \frac{1}{h-1}} \right)^{-1}
\]

(2)

\[
v_i = \frac{\sum_{k=1}^{n} (u_{ik})^h x_k}{\sum_{k=1}^{n} (u_{ik})^h}
\]

(3)

The specific flow chart is shown in Figure 1.
2.2. Improved Dempster–Shafer Evidence Theory

The Dempster–Shafer evidence theory (DS evidence theory) was presented by Dempster, and Shafer perfected the theory based on the research of Dempster. The DS evidence theory can describe the uncertainty and incompleteness of the evaluation information well [21].

Identification framework. Given a finite nonempty set identification framework, it includes \( N \) mutually exclusive elements, \( \Theta = \{L_1, L_2, \ldots, L_N\} \). The set of all of the subsets in \( \Theta \) is called a power set of \( \Theta \), that is, \( 2^\Theta \).

Basic probability assignment. Suppose that in the identification framework of \( \Theta \), \( m(L) \) satisfies the mapping of \( 2^\Theta \rightarrow [0, 1] \), \( L \) is any subset of the identification framework. \( L \) makes \( m(L) > 0 \), which is known as the focal element, and \( m(L) \) is regarded as the basic probability assignment (BPA; also known as the mass function), which indicates the degree of support for proposition \( L \). \( L \) satisfies the following:

\[
\begin{align*}
\sum m(L) &= 1 \\
m(\emptyset) &= 0
\end{align*}
\] (4)

We used the obtained mass function to make a decision, then, we used the Euclidean distance to help find the data near the true value, and even to find the true value, eliminate the elements that can be removed, and fuse the evidence with a fusion rule based on the reliability factor.

2.2.1. Judging the Accuracy of the Observations

All of the observations are denoted by \( L_i = (x_i, y_i) \), the distance from observation \( L_i \) to observation \( L_j \) is set to \( \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \). \( d_i \) indicated the average value of the distance \( \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \) between the two samples. Moreover, \( \overline{d} \) representd the mean distance between all of the observations, and the formulas for \( d_i \) and \( \overline{d} \) are as follows:

\[
d_i = \sum_{i=1}^{N} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}
\] (5)

\[
\overline{d} = \frac{1}{N} \sum_{i=1}^{N} d_i/N
\] (6)
Because the observations from the sensors are different, the normal observations in all of the deviation ranges should be distributed near the true value, and the observations with large deviations are far from the normal observations.

**Definition 1.** $U_1$ is a small deviation set and $U_2 = \overline{U_1}$ is a large deviation set, where $U_1 \cup U_2 = \emptyset$. $U_1$ meets the following requirements:

\[
\begin{align*}
    &\exists d_i < \delta \left( \forall L_i \subseteq U_1 \right) \\
    &\exists d_i \geq \delta \left( \forall L_i \subseteq U_1 \right)
\end{align*}
\]

(7)

2.2.2. Observations Converted to Evidence

Converting observation $L_i$ to evidence $e_i$ is the core of evidence theory and the basis of data fusion.

**Definition 2.** If any of the observations of $L_i, i \in \mathbb{N}$ exist as $\Delta_i \geq 0$, such that the true value, $L_0$, is within the neighborhood, $L_i$ of $\delta_i$ (the circle center of $L_i$ and the radius of $\Delta_i$), $\delta_i$ is the scattering interval of $L_i$, and $\Delta_i$ is called the scattering radius. The size of $\Delta_i$ is determined by the deviation of the observation $L_i$ from the true value $L_0$.

If $L_i \subseteq U_1$, $L_i$ is a small deviation observation, the scattering radius is relatively small, and $L_i$ is located in the $\delta_i$ circle (in the center of $L_i$, with a radius of $\overline{\delta_i}$). There are $K$ observations; $X_K \subseteq \Theta(K = 1, 2, \ldots, K)$ are small deviation observations, and it is considered that the $K$ observations are close to the true value, $L_0$, with the same probability, that is, the $K$ small deviation observations obtain a basic probability allocation of $1/K$, and the basic probability assignment of the remaining $N - K$ observations are 0. The mass function of the evidence, $e_i$, obtained by $L_i$, is as follows:

\[
m_i(X_K) = 1/K (\forall X_K \subseteq \delta_i).
\]

(8)

$L_i$ is a large deviation observation if $L_i \subseteq U_2$, which is far away from the true value, $L_0$, and if the scattering radius is large, $d_{\max}$ stands for the distance between the maximum and minimum values, and the following formula is obtained:

\[
d_{\max} = \max[L_i] - \min[L_i].
\]

(9)

Taking $d_{\max}$ as the scattering radius of $L_i$, because the true value, $L_0$, must lie in the intermediate of the maximum and the minimum observations, all the observations are included in $\delta_2$ (the circle centered $L_i$ and radius $d_{\max}$), obtaining the evidence’s ($e_i$) mass function, as follows:

\[
m_i(X_K) = 1/N (\forall L_i \subseteq \Theta).
\]

(10)

Then, the evidence, $e_i$, is converted from the large deviation of the observation, $L_i$, and each observation obtains a basic probability assignment of $1/N$. The process of generating $N$ initial evidence $e_i(i = 1, 2, \ldots, N)$ from $N$ observations have been completed [22, 23].

The above initial evidences may contain focal points of both $U_1$ and $U_2$, and the degree of deviation in each observation has not been considered. The initial evidence, $e_i$, is corrected as follows: for any $X_1, X_2 \in U_1$, and the ratio of basic probability assignment obtained by $X_1$ and $X_2$ is as follows:

\[
m_i(X_1)/m_i(X_2) = d_{X_2}/d_{X_1}.
\]

(11)

forall $X_K \in U_1$ and $L_i \subseteq U_2$, between $X_K$ and $L_i$, the rate is as follows:

\[
m_i(X_K)/m_i(L_i) = d_{\max}/d_K.
\]

(12)
The above two formulas produce a set of correction coefficients \( \{\omega_n\} (n = 1, 2, \ldots, N) \) for the normalized, weighted, and corrected evidence, with the following equation:

\[
M_i(L_i) = \frac{\omega_i \cdot m_i(L_i)}{\sum_{n=1}^{N} \omega_n \cdot m_i(L_n)}.
\]  

(13)

2.2.3. Combination of Evidence

There may be a high conflict between all of the evidence obtained by the above algorithm, and in order to avoid the unreasonable weight distribution, combined with the literature [10,11], the evidence fusion formula that assigns the probability of supporting evidence conflict to the observation is as follows:

\[
m(L_i) = \prod_{i=1}^{N} M(L_i) + c \cdot \overline{M}(L_i)
\]

(14)

where \( c \) is the conflict factor and \( \overline{M}(L_i) \) is the average distribution of \( L_i \) in all of the evidence. The formulas are as follows:

\[
c = 1 - \sum_{i=1}^{N} \prod_{n=1}^{N} M_i(L_i)
\]

(15)

\[
\overline{M}(L_i) = \frac{1}{N} \sum_{i=1}^{N} M_i(L_i)
\]

(16)

The basic probability distribution \( m(L_i) \) is the weight obtained by \( L_i \), and the fusion result is as follows:

\[
L_0 = \sum_{i=1}^{N} L_i m(L_i)
\]

(17)

3. Experiment Analysis

In order to verify the effectiveness of this algorithm, 100 sets of two-dimensional datasets were used for the experiments. The datasets involved the ambient illumination and street light illumination of intelligent campus lighting systems. As a result of the measurement error of the sensors and other factors, such as ambient lighting, the ambient illumination and street lights of different streets were measured to be different at the same time, but the observations were always distributed near a certain value. The reference value of the switch threshold is 20.00 LX.

The 100 sets of data were collected within 5 min of nightfall. At dusk, the brightness of the environment changed rapidly, and the process from daylight to dark was more obvious. Therefore, the ambient brightness varied from bright to dark, as the \( x \) column of the two-dimensional data sets. When the ambient light level was dimmed from light to dark, the data of this column should be changed from large to small. As the brightening of the street light source is a process, when the street light turned on, the brightness of the light source is relatively small. After a period of time, the light source will be brightened to a normal brightness. Thus, the change in the brightness of the light was recorded within 5 min. The light is a process from dark to light, as seen in the \( y \) column of the two-dimensional datasets, and the data of this column should be changed from small to large. This is the data law of the 100 datasets.

By combining the two brightness characteristics of street lights, a street light switching threshold suitable for the specific environment will be obtained. The brightness of the single light should be different in different weather, locations, human flows, and so on. The threshold obtained by fusion is also different. This threshold helps the street light’s manager to manage and control the switch of the street lights precisely. The administrator can achieve precise control of the street light switch in different weather, environments, and seasons, according to the fusion threshold [24].
In this paper, the distance between the two observations was calculated by the Euclidean distance. These two-dimensional observations were clustered by MATLAB software programming. FCM helps to obtain clustering images and clustering conditions. Each clustering observation is regarded as the identification framework, and is then converted into evidence. By modifying and combining the evidence, we get the weight assignment of the observations, which is called the mass function, and the preliminary fusion value of the samples can be obtained. Finally, the data of all of the categories are weighted according to different weights. The fusion result is one-dimensional data. We compared the fusion results of two, three, and four categories, using the highest precision data as the final fusion result, which is the switching threshold of the street light in smart campus lighting system.

The observations are as follows:

\[
X = [24.22, 27.45; 23.9, 27.68; 23.72, 27.76; 23.55, 27.83; 23.38, 27.9; 23.22, 27.95; 23.05, 28.11; 23.28, 28.21; 22.93, 28.33; 22.89, 28.52; 22.85, 28.54; 22.66, 28.6; 22.6, 28.68; 22.55, 28.69; 22.5, 28.7; \ldots \ 18.3, 34.66; 18.21, 34.78; 18.17, 34.82; 18.1, 34.86; 17.98, 34.91; 17.93, 35.02; 17.88, 35.1; 17.8, 35.18; 17.66, 35.27; 17.59, 35.35; 17.51, 35.41; 17.38, 35.64; 17.21, 35.9; 16.9, 36.12; 16.66, 36.34].
\]

In the above three figures, the abscissa indicates the ambient illuminance, and the illuminance of the street lights is shown in the ordinate. The samples with two attributes get the corresponding cluster images under the constraints of different cluster centers, respectively. Different categories of samples are represented by different shapes and colors. The red “x” indicates the cluster center. The first category of data is depicted by black circles, the second category is represented by purple triangles, the third category is shown by green “*”, and the fourth category is presented by blue triangles. Figure 2 shows the location of the cluster center and which category all of the sample points belong to when the number of the cluster centers is two. Figure 3 shows the number of cluster centers set to three, the location of the cluster center and the category to which all of the samples belong. The number of cluster centers is four, and the location of the cluster center and the category to which all of the data belong are shown in Figure 4. For the different datasets, it cannot be decided whether to divide the dataset into several categories in order to achieve the best fusion effect. Therefore, the DS theory was used to fuse the three clustering cases, and then the fusion results were compared. The highest precision result was the threshold of the intelligent campus lighting system.

As can be seen from Tables 1–3, the data fusions are firstly performed by each group of data obtained by FCM clustering, and then the fusion results of these three groups are weighted and fused to obtain the final fusion result. When the number of clusters is three, the relative error is 2.003%, the relative error with two categories is 2.431%, and the relative error with four categories is 2.015%, so it is best to divide the dataset into three categories. Take the third group of the highest precision fusion results as an example—the fusion result is shown in Table 4.
As can be seen from Tables 1–3, the data fusions are firstly performed by each group of data obtained by FCM clustering, and then the fusion results of these three groups are weighted and fused to obtain the final fusion result. When the number of clusters is three, the relative error is 2.003%, the relative error with two categories is 2.431%, and the relative error with four categories is 2.015%, so it is best to divide the dataset into three categories. Take the third group of the highest precision fusion results as an example—the fusion result is shown in Table 4.

Table 1. Clustering results with two cluster centers.

| Group Data Numbe r | Observations Result |
|--------------------|---------------------|
| 1                  | X = [24.22, 27.45; 23.9, 27.68; 23.72, 27.76; 23.55, 27.83; 23.38, 27.9; … 19.78, 31.9; 19.76, 31.95; 19.75, 32; 19.71, 32.04; 19.7, 32.08] | 21.5649 |
| 2                  | X = [19.68, 32.12; 19.66, 32.25; 19.63, 32.34; 19.62, 32.48; 19.59, 32.55; … 17.51, 35.41; 17.38, 35.64; 17.21, 35.9; 16.9, 36.12; 16.66, 36.34] | 19.0030 |
| Fusion result      | 20.5658             |

Table 2. Clustering results with three cluster centers.

| Group Data Numbe r | Observations Result |
|--------------------|---------------------|
| 1                  | X = [24.22, 27.45; 23.9, 27.68; 23.72, 27.76; 23.55, 27.83; 23.38, 27.9; … 22.87] | 20.4006 |
| 2                  | X = [19.68, 32.12; 19.66, 32.25; 19.63, 32.34; 19.62, 32.48; 19.59, 32.55; … 17.51] | 19.0030 |
| Fusion result      | 20.4006             |

The final fusion result of the three groups of data shown in Table 2 is 20.4006, and the relative error is 2.003%.

The fixed threshold method is a global fixed threshold, which means that all of the datasets of a system are binarized with a unified threshold. The multi-attribute fusion algorithm proposed in this paper can fuse a threshold adaptively, and can dynamically change the threshold to reduce the error. Take the data sample of high brightness ambient illumination as an example. The reference value of the switch threshold is 30.85 LX.

Figures 5–7 show the cluster centers of different categories. From the following Tables 5–7, it can be seen that when the samples are clustered into four categories, the fusion result has the highest precision, and the fusion result is 31.4590. The fixed threshold method aims to set all of the street lights to the same switch threshold. If the switch threshold is still set to 20.00 LX in the place where the ambient brightness is high, a large error will occur, resulting in a waste of energy.
Table 1. Clustering results with two cluster centers.

| Group | Data Number | Observations | Result |
|-------|-------------|--------------|--------|
| 1     | 61          | \(X = [24.22,27.45;23.9,27.68;23.72,27.76;23.55,27.83;23.38,27.9;\ldots;19.78,31.9;19.76,31.95;19.75,32;19.71,32.04;19.7,32.08]\) | 21.5649 |
| 2     | 39          | \(X = [19.68,32.12;19.66,32.25;19.63,32.34;19.62,32.48;19.59,32.55;\ldots;17.51,35.41;17.48,35.64;17.21,35.9;16.9,36.12;16.66,36.34]\) | 19.0030 |
| Fusion result | | | 20.5658 |

Table 2. Clustering results with three cluster centers.

| Group | Data Number | Observations | Result |
|-------|-------------|--------------|--------|
| 1     | 25          | \(X = [24.22,27.45;23.9,27.68;23.72,27.76;23.55,27.83;23.38,27.9;\ldots;21.85,29.1;21.72,29.25;21.69,29.42;21.55,29.5;21.47,29.65]\) | 22.8750 |
| 2     | 54          | \(X = [21.12,30.57;21,30.6;20.93,30.62;20.89,30.65;20.87,30.66;20.86,30.68;\ldots;19.33,32.92;19.28,32.98;19.2,33.15;19.17,33.22;19.13,33.38]\) | 20.1831 |
| 3     | 21          | \(X = [18.65,34.25;18.6,34.28;18.58,34.34;18.46,34.38;18.42,34.47;\ldots;17.51,35.41;17.38,35.64;17.21,35.9;16.9,36.12;16.66,36.34]\) | 18.1422 |
| Fusion result | | | 20.4006 |
Table 3. Clustering results with four cluster centers.

| Group | Data Number | Observations                                                                 | Result    |
|-------|-------------|-------------------------------------------------------------------------------|-----------|
| 1     | 23          | \(X = [24.22, 27.45; 23.9, 27.68; 23.72, 27.76; 23.55, 27.83; 23.38, 27.9;\) | 22.9340   |
|       |             | \(\ldots\)                                                                   |           |
|       |             | 21.97, 29; 21.9, 29.06; 21.85, 29.1; 21.72, 29.25; 21.69, 29.42]               |           |
| 2     | 30          | \(X = [21.55, 29.5; 21.47, 29.65; 21.12, 30.57; 21.30, 6; 20.93, 30.62;\)  | 20.7256   |
|       |             | \(\ldots\)                                                                   |           |
|       |             | 20.09, 31.53; 20.06, 31.55; 20.01, 31.6; 19.98, 31.65; 19.9, 31.75]          |           |
| 3     | 26          | \(X = [19.87, 31.8; 19.83, 31.81; 19.8, 31.88; 19.78, 31.9; 19.76, 31.95;\) | 19.6177   |
|       |             | \(\ldots\)                                                                   |           |
|       |             | 19.33, 32.92; 19.28, 32.98; 19.2, 33.15; 19.17, 33.22; 19.13, 33.38]         |           |
| 4     | 21          | \(X = [18.65, 34.25; 18.6, 34.28; 18.58, 34.34; 18.46, 34.38; 18.42, 34.47;\) | 18.1422   |
|       |             | \(\ldots\)                                                                   |           |
|       |             | 17.51, 35.41; 17.38, 35.64; 17.21, 35.9; 16.9, 36.12; 16.66, 36.34]          |           |
|       | Fusion result |                                                                                  | 20.4030   |
Table 4. Group 3 evidence of mass function.

| Evidence | \( m_3(L_{80}) \) | \( m_3(L_{81}) \) | \( m_3(L_{82}) \) | \( m_3(L_{83}) \) | \( m_3(L_{84}) \) | \( m_3(L_{96}) \) | \( m_3(L_{97}) \) | \( m_3(L_{98}) \) | \( m_3(L_{99}) \) | \( m_3(L_{100}) \) |
|----------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| e_{80}   | 0.0088            | 0.0261            | 0.0610            | 0.1092            | 0.1757            | ...               | 0                 | 0                 | 0                 | 0                 |
| e_{81}   | 0.0088            | 0.0261            | 0.0610            | 0.1092            | 0.1757            | ...               | 0                 | 0                 | 0                 | 0                 |
| e_{82}   | 0.0060            | 0.0177            | 0.0414            | 0.0740            | 0.1192            | ...               | 0                 | 0                 | 0                 | 0                 |
| e_{83}   | 0.0031            | 0.0093            | 0.0217            | 0.0388            | 0.0624            | ...               | 0                 | 0                 | 0                 | 0                 |
| e_{84}   | 0.0024            | 0.0093            | 0.0217            | 0.0388            | 0.0624            | ...               | 0                 | 0                 | 0                 | 0                 |
| ...      | ...               | ...               | ...               | ...               | ...               | ...               | ...               | ...               | ...               | ...               |
| e_{96}   | 0.0017            | 0.0051            | 0.0119            | 0.0212            | 0.0342            | ...               | 0.0070            | 0.0061            | 0.0053            | 0.0046            |
| e_{97}   | 0.0017            | 0.0051            | 0.0119            | 0.0212            | 0.0342            | ...               | 0.0070            | 0.0061            | 0.0053            | 0.0046            |
| e_{98}   | 0.0017            | 0.0051            | 0.0119            | 0.0212            | 0.0342            | ...               | 0.0070            | 0.0061            | 0.0053            | 0.0046            |
| e_{99}   | 0.0017            | 0.0051            | 0.0119            | 0.0212            | 0.0342            | ...               | 0.0070            | 0.0061            | 0.0053            | 0.0046            |
| e_{100}  | 0.0017            | 0.0051            | 0.0119            | 0.0212            | 0.0342            | ...               | 0.0070            | 0.0061            | 0.0053            | 0.0046            |
| Synthetic evidence | 0.0023            | 0.0070            | 0.0169            | 0.0327            | 0.0522            | ...               | 0.0030            | 0.0026            | 0.0023            | 0.0020            |
| Fusion result | 18.1422           | 18.1422           | 18.1422           | 18.1422           | 18.1422           | ...               | ...               | ...               | ...               | ...               |

Table 5. Clustering results with two cluster centers based on the fixed threshold method.

| Group | Data Number | Observations | Result |
|-------|-------------|--------------|--------|
| 1     | 52          | \( X = [37.6,26.88;37.26,26.99;36.98,27.33;36.75,27.49;36.54,27.85; \ldots; 31.58,33.78;31.56,33.88;31.49,33.99;28.13,35.79;31.38,34.26] \) | 34.8223 |
| 2     | 48          | \( X = [31.19,34.27;31.07,34.31;31.03,34.38;30.88,34.45;30.76,34.56; \ldots; 21.56,36.43;21.23,36.47;20.7,36.52;20.3,36.52;19.95,36.57] \) | 29.2341 |
| Fusion result | | | 32.1400 |
Table 6. Clustering results with three cluster centers based on the fixed threshold method.

| Group | Data number | Observations                                      | Result |
|-------|-------------|---------------------------------------------------|--------|
| 1     | 50          | $X = [37.6,26.88;37.26,26.99;36.98,27.33;36.75,27.49;36.54,27.85;\ldots;31.76,33.64;31.65,33.69;31.58,33.78;31.56,33.88;31.49,33.99]$ | 34.9136 |
| 2     | 19          | $X = [28.13,35.79;31.38,34.26;31.19,34.27;31.07,34.31;31.03,34.38;\ldots;30.12,35.19;30.09,35.21;30.01,35.26;29.89,35.42;29.83,35.49]$ | 30.5002 |
| 3     | 31          | $X = [29.8,35.41;29.75,35.45;29.68,35.48;29.62,35.52;29.58,35.54;\ldots;21.56,36.43;21.23,36.47;20.7,36.5;20.3,36.52;19.95,36.57]$ | 27.2545 |
|       |             | Fusion result                                     | 31.7007 |

Table 7. Clustering results with four cluster centers based on the fixed threshold method.

| Group | Data Number | Observations                                      | Result |
|-------|-------------|---------------------------------------------------|--------|
| 1     | 34          | $X = [37.6,26.88;37.26,26.99;36.98,27.33;36.75,27.49;36.54,27.85;\ldots;33.82,32.33;33.76,32.41;33.65,32.55;33.53,32.67;33.38,32.72]$ | 35.6463 |
| 2     | 17          | $X = [33.25,32.85;33.21,32.96;33.16,33.21;33.05,33.26;32.51,33.3;\ldots;31.65,33.69;31.58,33.78;31.56,33.88;31.49,33.99;31.42,34.4]$ | 32.3908 |
| 3     | 28          | $X = [31.38,34.26;31.19,34.27;31.07,34.31;31.03,34.38;30.88,34.45;\ldots;29.51,35.61;29.4,35.66;28.65,35.7;28.4,35.73;28.13,35.79]$ | 30.3405 |
| 4     | 21          | $X = [27.85,35.85;27.6,35.91;27.38,35.96;27.06,36.01;26.42,36.04;\ldots;21.56,36.43;21.23,36.47;20.7,36.5;20.3,36.52;19.95,36.57]$ | 25.4167 |
|       |             | Fusion result                                     | 31.4590 |
The literature [25] shows us the remediation of failed identification in product multi-source information fusion based on DS evidence theory. This work is aimed at the wear and tear, corrosion, and pollution of two-dimensional data matrix symbols lead by the complexity of the discrete manufacturing enterprise, production environment, and production process. The method firstly established a remedy technical framework. Secondly, it calculated the similarity measurement of the multi-attribute data identification of invalid distinguish goods. Finally, it recognized the data matrix code based on the multi-attribute fusion of DS theory. The experiment analysis demonstrated a processing statistic for 120 failure data matrix codes, and matching the correct rate can be up to 96%. Compared with this algorithm, the correct rate of the algorithm proposed in this paper is above 97.997%,

**Figure 5.** FCM cluster image based on the fixed threshold method (two cluster centers).

**Figure 6.** FCM cluster image based on the fixed threshold method (three cluster centers).

**Figure 7.** FCM cluster image based on the fixed threshold method (four cluster centers).

| Group | Data Number | Observations | Result |
|-------|-------------|--------------|--------|
| 1     | 52          | X = [37.6, 26.88; 37.26, 26.99; 36.98, 27.33; 36.75, 27.49; 36.54, 27.85; … 31.58, 33.78; 31.56, 33.88; 31.49, 33.99; 28.13, 35.79; 31.38, 34.26] | 34.8223 |
| 2     | 48          | X = [31.19, 34.27; 31.07, 34.31; 31.03, 34.38; 30.88, 34.45; 30.76, 34.56; … 29.2341 |
which has the advantage of a higher accuracy. This opinion has more robustness to switching threshold of the smart street light system. This method can better integrate the street light switching threshold to be suitable for a specific environment, which provides effective support for other two-dimensional information fusion application scenarios.

4. Conclusions

Based on the FCM clustering algorithm and improved evidence theory, a novel multi-attribute fusion algorithm is proposed. The algorithm can combine the light source illumination and ambient illumination of the smart campus street light system effectively in order to obtain the switching threshold. To the best of our knowledge, it is the first time that the FCM clustering algorithm is being combined with evidence theory for multi-attribute fusion. The effectiveness of the algorithm is also tested by the real-life datasets based on the smart campus street lights system. The effect of the algorithm is proven by the change of the cluster centers, which lays a foundation for future multi-attribute fusion.

This algorithm also involves some future issues. First of all, the clustering method cannot be fully applied to the data of all of the application scenarios. When the number of samples is large, the FCM algorithm needs to define the cluster center, and it easily falls into the local optimal solution. Secondly, is how to distinguish useful information from noise in the process of classifying samples. The next step is to try other algorithms (such as the Bayesian estimation) to cluster the initial observations more accurately. For different application problems, how to make the algorithm adapt to the cluster center intelligently also has a lot of space for improvement.

Author Contributions: Conceptualization and methodology, W.W. and Y.Y.; software, D.Z.; validation, Z.W. and Y.F.; data curation, C.Y.; writing (original draft preparation), Y.Y.; writing (review and editing), Y.F.; project administration, W.W.

Funding: This work is supported by the Shaanxi Provincial Department of Science and Technology key project in the field of industry (2018ZDXM-GY-039).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, H.J. Design of Unmanned Toll Parking System Based on License Plate Recognition. Zhejiang Constr. 2018, 35, 45–48.
2. Gao, D.D. Research on Anti-glare and License Plate Anti-overexposure of Intelligent Traffic Filling System. Doctoral Thesis, Zhejiang University of Technology, Hangzhou, China, 13 December 2016.
3. Shou, Z.Z.; Ye, J.Y.; Wei, P. Illumination compensation of electronic bayonet system. China Public Secur. Acad. Ed. 2011, 22, 86–90.
4. Gao, W.; Wu, S.L. From “bright eyes” to “no flash”—“ecological” cameras help solve road junction light pollution. In Proceedings of the 4th Shenzhen International Drone Exhibition, Shenzhen, China, 20–22 June 2019.
5. Ma, N.; Tian, J.; Lu, S.W. Advances in research on effects of LED fill light on growth of seedling vegetables under low light conditions. Anhui Agric. Sci. 2018, 46, 1–3, 6.
6. Litxp. Design of Yancheng planning exhibition hall network intelligent lighting system. Digit. Community Smart Home 2008, 1, 94–96.
7. Xu, Y.; Xie, T.X. Research of practical expert system for transformer fault diagnosis based on multidimensional data fusion. In Proceedings of the Asia-Pacific Power and Energy Engineering Conference (APPEEC), Xi’an, China, 25–28 December 2016.
8. Wu, G.Q.; Gan, S.W.; Liu, J.T. Simulation of hydrological processes in dataless watershed based on fuzzy clustering algorithm. Water Resour. Power 2018, 36, 13–16.
9. Yang, M.S.; Cheng, T.Y. Bias-correction fuzzy clustering algorithms. Inf. Sci. 2015, 309, 138–162. [CrossRef]
10. Xiong, Y.M.; Yang, Z.P. Sensor data fusion algorithm inspired by evidence theory. J. Huazhong Univ. Sci. Technol. (Nat. Sci. Ed.) 2011, 39, 50–54.
11. Wang, W.Q.; Yang, Y.L.; Yang, C.J. A data fusion algorithm based on evidence theory. *Control Decis.* 2013, 28, 1427–1430.

12. Hira, M.; Laghari, N.U.Z.; Sangrasi, D.M.; Dayo, Z.A. Comparative Analysis of Hybrid Clustering Algorithm on Different Dataset. In Proceedings of the 8th IEEE International Conference on Electronics Information and Emergency Communication, ICEIEC, Beijing, China, 15–17 June 2018.

13. Shashi, M.; Shruti, K. Comparative analysis of K-Means with other clustering algorithms to improve search result. In Proceedings of the 1st International Conference on Green Computing and Internet of Things, ICGCIOIT, Noida, India, 8–10 October 2015.

14. Choe, H.; Jordan, J.B. On the optimal choice of parameters in a fuzzy c-means algorithm. In Proceedings of the IEEE International Conference on Fuzzy Systems, San Diego, CA, USA, 8–12 March 1992.

15. Zhang, W.; Ji, X.L. Data fusion method based on improved D-S evidence theory. In Proceedings of the 2018 IEEE International Conference on Big Data and Smart Computing, Big Comp, Shanghai, China, 15–17 January 2018.

16. Wu, X.; Zhou, J. Novel possiblistic fuzzy c-means clustering. *Acta Electron. Sin.* 2008, 36, 1996–2000.

17. XU, R.; Wunsch, D. Survey of clustering algorithm. *IEEE Trans. Neural Netw.* 2005, 16, 645–678. [CrossRef] [PubMed]

18. Yang, X. Research of key techniques in cluster analysis. Doctoral Thesis, Zhejiang University, Hangzhou, China, May 2005.

19. Wu, J.; Feng, G.R. Intrusion detection based on simulated annealing and fuzzy c-means clustering. In Proceedings of the 2009 International Conference on Multimedia Information Networking and Security, Hubei, China, 18–20 November 2009; pp. 382–385.

20. Khalilia, M.A.; Bezdek, J.; Popescu, M.; Keller, J.M. Improvements to the relational fuzzy c-means clustering algorithm. *Pattern Recognit.* 2014, 47, 3920–3930. [CrossRef]

21. Cao, M.L.; D, Y.Y. An improved D-S evidence theory in tire pressure monitoring system. *IOP Conf. Ser. Mater. Sci. Eng.* 2018, 394, 032112. [CrossRef]

22. Pan, G.; Wu, L.L. Information fusion based on improved D-S evidence theory. *Appl. Mech. Mater.* 2013, 411–414, 49–52. [CrossRef]

23. Zhu, Y.; Zhang, W. Trust model regarding QoS using FCMs and D-S evidence theory. In Proceedings of the 2008 IEEE International Conference on Service Operations and Logistics, and Informatics, IEEE/SOLI 2008, Beijing, China, 12–15 October 2008.

24. Shu, X.; Xu, F. Hangzhou street light automatically opens and closes according to the environment. *China Illuminating Eng. J.* 2018, 29, 91–94.

25. Wang, J.; He, W.-P.; Wang, W.; Li, X.S.; Guo, G.F. Remediation of failed identification in product multi-information fusion based on D-S evidence theory. *Comput. Integr. Manuf. Syst.* 2015, 21, 1142–1149.