Research Article

Special Object Recognition Based on Sparse Representation in Multisource Data Fusion Samples

Changjun Zha

1College of Advanced Manufacturing Engineering, Hefei University, Hefei 230601, China
2Key Laboratory of Intelligent Computing & Signal Processing, Ministry of Education, Anhui University, Hefei 230039, China

Correspondence should be addressed to Changjun Zha; 11586292@qq.com

Received 16 December 2019; Accepted 15 May 2020; Published 28 May 2020

Academic Editor: Sebastian Anita

Copyright © 2020 Changjun Zha. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Wireless sensor networks (WSNs) suffer from limited power and large amounts of redundant data. This paper describes a multisource data fusion method for WSNs that can be combined with the characteristics of a profile detection system. First, principal component analysis is used to extract sample features and eliminate redundant information. Feature samples from different sources are then fused using a method of superposition to reduce the amount of data transmitted by the network. Finally, a mathematical model is proposed. On the basis of this model, a novel method of special object recognition based on sparse representation is developed for multisource data fusion samples according to the distribution of nonzero coefficients under an overcomplete dictionary. The experimental results from numerical simulations show that the proposed recognition method can effectively identify special objects in the fusion samples, and the overall performance is better than that of traditional methods.

1. Introduction

In general, objects moving across borders or uninhabited regions could be either humans or animals. As there are significant differences between human and animal profiles, it is feasible to use certain features to recognize and monitor special objects (humans) [1–3]. However, previous studies are largely concerned with the acquisition of object profile samples and recognition using traditional methods such as k-nearest neighbors (KNN) and support vector machines (SVMs) [4–6]. There has been no detailed discussion on the fusion of data or reducing the amount of data transmitted over the network. In this paper, we discuss these issues in detail based on previous work [7].

The main contribution of this paper is to construct a wireless monitoring network with a profile detection system as a network node. To reduce the amount of data transmitted in the network as much as possible, the network uses a sink node to fuse the feature samples sent by each network node; the fused data are then sent to the terminal for identification, allowing the determination of whether a special object (human) is passing through the monitoring area.

This paper describes the fusion of multisource feature samples through a superposition approach, meaning that traditional recognition methods such as KNN, SVM, and sparse representation classification (SRC) [8–10] cannot accurately identify whether there is one or more class objects in the fused sample. To solve this problem, we propose a novel object recognition method based on sparse representation. Different from classical SRC, the method described in this paper uses the distribution of nonzero coefficients in the sparse representation to identify special objects in the fusion sample. The experimental results verify the effectiveness of this method.

2. Related Work

In our previous work, we designed the profile detection system shown in Figure 1. In the system, the signal sensing unit is composed of photoelectric sensors installed at even intervals on a vertical fixed bracket with a length of 2 m. When a moving object passes through the sensor’s field of view, the profile feature information of the object is captured by the sensing unit [1, 2, 7].
This paper mainly studies the profile classification of three kinds of moving objects: humans, squatting humans, and animals (namely, dogs). In practice, when objects enter the sensor’s field of view, their position, speed, attitude, and other conditions are different. To obtain a better recognition effect, samples of various angles, speeds, and attitudes entering the field of view are collected for training, as shown in Figure 2.

3. Sample Feature Extraction and Data Fusion

3.1. Feature Extraction. According to Zha et al. [7], the profile detection system of the network monitoring node in the wireless monitoring network expresses an object profile sample $P$ as an $S \times L$ matrix. The number of columns of each sample matrix is different, and there is a large amount of redundant information. Sending the sample matrix signal directly through the network wastes network resources and is inconvenient for data processing because of the different dimensions of the matrix formed by $P$. To solve these problems, we use principal component analysis (PCA) [11–13] to pretreat the sample data. The specific process is as follows:

1. **Input**: sample matrix $P \in \mathbb{R}^{S \times L}$.

2. Calculate the covariance:

   $$C_{XX} = E[(P^T - E(P^T))(P - E(P))].$$  

3. Determine eigenvalues and eigenvectors:

   $$C_{XX} = U\Lambda U^T,$$

   namely, the eigenvalues $\Lambda = [\lambda_1 \lambda_2 \cdots \lambda_L]$ and the corresponding eigenvectors $U = [\mu_1 \mu_2 \cdots \mu_L]$.

4. Construct transformation matrix: select the eigenvectors corresponding to the $k (k \leq L)$ largest eigenvalues to construct the transformation matrix, namely,

   $$U_k = [\mu_1, \mu_2, \cdots, \mu_k], \text{ i.e., } U_k \in \mathbb{R}^{L \times k}.$$  

5. Dimension reduction:

   $$\hat{P} = U_k^T P^T.$$  

6. **Output**: matrix vectorization:

   $$y = \text{vec}(\hat{P}) \in \mathbb{R}^c \quad (c = k \times S).$$  

After pretreatment, the vector $y$ is taken as the feature sample of the object and sent to the sink node for data fusion.

3.2. Multisource Data Fusion Model. Suppose that the wireless sensor network (WSN) contains $N$ network monitoring nodes and that the $j$th network node acquires the feature sample of the object as $y_j \in \mathbb{R}^c (1 \leq j \leq N)$. The sink node receives the feature sample of each network node and performs data fusion. The network topology is shown in Figure 3. To reduce the computational complexity of the data fusion process, we adopt the method of superposition to achieve fusion. Thus, the fusion sample can be expressed as follows:

$$r = y_1 + y_2 + \cdots + y_N = \sum_{j=1}^{N} y_j.$$  

Considering the signal transmission power, channel fading, noise interference, and other factors associated with each network node, equation (6) can be rewritten as follows:

$$r = (\sqrt{p_1} \cdot \alpha_1 \cdot y_1 + \sqrt{p_2} \cdot \alpha_2 \cdot y_2 + \cdots + \sqrt{p_N} \cdot \alpha_N \cdot y_N) + n$$  

$$= \sum_{j=1}^{N} \sqrt{p_j} \alpha_j y_j + n,$$  

where $r \in \mathbb{R}^c$ and the parameters $p_j$ and $\alpha_j (1 \leq j \leq N)$ are the signal transmission power and channel link gain of the $j$th network node, respectively and $n$ is the additive white Gaussian noise. When the dimension of the feature sample of each network node is large and the amount of data to be transmitted needs to be further reduced, the feature sample can be projected into a low-dimensional space through the projection matrix $\Psi \in \mathbb{R}^{d \times c} (d < c)$. The signal model is shown in Figure 4. The mathematical model of the fusion samples can be written as follows:

$$\hat{r} = \sum_{j=1}^{N} \sqrt{p_j} \alpha_j \hat{y}_j + n = \sum_{j=1}^{N} \sqrt{p_j} \alpha_j \hat{y}_j + n,$$  

Figure 1: Profile detection system.
where $\hat{y}_j = \Psi y_j \in \mathbb{R}^d$.

4. Sparsity Analysis and Construction of Overcomplete Dictionary

To analyze the sparsity of the fusion samples, it is assumed that $r$ is composed of three different feature samples (i.e., $y_1$, $y_2$, and $y_3$), namely,

$$ r = y_1 + y_2 + y_3. \quad (9) $$

Furthermore, suppose that $x_i (1 \leq i \leq 3)$ are sparse representation coefficients of eigenvector $y_i$ under the overcomplete dictionary $A_i$. Then,

$$\begin{align*}
y_1 &= A_1 x_1, \\
y_2 &= A_2 x_2, \\
y_3 &= A_3 x_3.
\end{align*} \quad (10)$$

Combined with equation (10), equation (9) can be rewritten as follows:
\[
\mathbf{r} = \frac{1}{3} \mathbf{y}_j = \frac{1}{3} \mathbf{A}_i \mathbf{x}_i = \left[ \mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3 \right] \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix}
\]

(11)

where \( \mathbf{A}_i = [\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3] \) and \( \mathbf{x}_i = [\mathbf{x}_1^T, \mathbf{x}_2^T, \mathbf{x}_3^T]^T \). It can be seen from equation (11) that the fusion sample \( \mathbf{r} \) can be sparsely represented under the dictionary \( \mathbf{A}_n \), and the sparse coefficients can be obtained by solving the following \( l_1 \)-minimization problem [14–16]:

\[
\hat{x}_n = \arg \min_{x_n} \|x_n\|_1 \quad \text{s.t.} \|\mathbf{r} - \mathbf{A}_n x_n\|_2^2 \leq \epsilon,
\]

(12)

where the parameter \( \epsilon \) is the error tolerance. Similarly, when the fusion sample \( \mathbf{r} \) is composed of \( k \) feature samples, it can be sparsely represented, and the overcomplete dictionary is

\[
\mathbf{A}_n = [\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k].
\]

To represent the fusion sample sparsely, we use the training samples to construct an overcomplete dictionary directly [7, 17–19]. Assuming that there are \( T \) classes of training samples, the number of training samples in each class is \( N_1, N_2, \ldots, N_T \), and the \( i \)th \((1 \leq i \leq N_j)\) training sample in the \( j \)th \((1 \leq j \leq T)\) class is expressed as \( \mathbf{P}_{ij} \in \mathbb{R}^{S \times L_i} \), and the specific process of constructing the dictionary is as follows:

1. **Input**: training samples \( \mathbf{P}_{ij} \in \mathbb{R}^{S \times L_i} \).

2. **Pretreatment of training samples**:

   All training samples \( \mathbf{P}_{ij} \in \mathbb{R}^{S \times L_i} \) are pretreated, and the feature vector \( \mathbf{\varphi}_{ij} \in \mathbb{R}^c \) ( \( c = S \times k \) ) of the pretreated sample is used as a dictionary atom.

3. **Output**: construct overcomplete dictionary:

   \[
   \mathbf{A}_n = [\mathbf{\varphi}_{11}, \mathbf{\varphi}_{12}, \ldots, \mathbf{\varphi}_{TN}] \in \mathbb{R}^{c \times n} \quad \text{where,} \quad n = N_1 + N_2 + \cdots + N_T.
   \]

Through the above process, we can obtain an overcomplete dictionary.

5. Special Object Recognition Method

According to the above sparsity analysis, when there is one class of feature samples in the fusion sample, the main nonzero coefficients in the sparse coefficient vector obtained by \( l_1 \)-minimization are distributed on the corresponding class of atoms, whereas the coefficients for other classes of atoms are zero or very small. If the fusion sample contains multiple classes, the main nonzero coefficients in the sparse coefficient vector are distributed on these classes. Based on this feature, we propose a special object recognition method for multisource data fusion samples. The method is illustrated in Figure 5, and the specific steps are as follows:

1. **Input**: dictionary \( \mathbf{A}_n = [\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_T] \in \mathbb{R}^{c \times n} \) for \( T \) classes and fusion sample \( \mathbf{r} \in \mathbb{R}^r \).

2. **Sparse representation**:

\[
\hat{x}_n = \arg \min_{x_n} \|x_n\|_1 \quad \text{s.t.} \|\mathbf{r} - \mathbf{A}_n x_n\|_2^2 \leq \epsilon,
\]

where \( \epsilon \) is the error tolerance.

3. **Calculate coefficient** \( l_1 \)-norm of each class:

\[
s_j = \left\| \Delta \left( \hat{x}_n \right)_j \right\|_{1} \quad \text{for} \quad j = 1, \ldots, T,
\]

(14)

where \( \delta_j \) is the characteristic function that selects the coefficients associated with the \( j \)th class.

4. **Multiclass discriminant rule**:

\[
\text{identity}(\mathbf{r}) = \left\{ j \mid \frac{s_j}{\| \hat{x}_n \|_1} \geq r, \quad j = 1, \ldots, T \right\},
\]

(15)

where \( r \in [0, 1] \).

5. **Output**: based on the results of step 4, check whether the special object is included.

6. Experimental Simulation and Results Analysis

Experiments are conducted based on the profile detection system, which uses 16 ESF-R2NK photoelectric sensors to construct a signal sensing unit. The effective distance of the sensors is 2 m. In the actual environment, the profile detection system mainly collects three kinds of object profile samples for training and testing: humans, squatting humans, and animals (dogs). The numbers of training samples and test samples are listed in Table 1.

It is assumed that each network node can detect at most one object in a certain period of time. In the following experiments, the fusion samples are composed of humans and dogs, and we consider the humans as the special objects. Under different experimental conditions, the proposed method can judge whether there is a special object (human) in the fusion sample and considers this a correct recognition. For the convenience of analysis and comparison, it is assumed that the signal transmission power of each network node is the same, and the channel link gain remains constant at 1. The performance of the proposed method is compared with that of the traditional nearest-neighbor classifier (NN) [20] and the KNN, SVM, and SRC methods.

The experiments are conducted on a Windows 7 PC running MATLAB (R2014a) with a 3.0 GHz CPU and 8 GB memory.

6.1. Experiment 1: Verifying the Sparsity of Fusion Samples

It is assumed that, within a certain period of time, two network nodes in the monitoring area detect objects passing by, among which one node is a single person and one node is a dog. After each node extracts the sample feature information, the sink node performs data fusion. The sparsity of the fused feature sample is analyzed, and the distribution of sparse representation coefficients is shown in Figure 6.
As can be seen from Figure 6, the coefficients of the fusion samples are still sparse and distributed on the atomic terms of the two classes of human and animal. The fusion samples do not contain images that include squatting humans, so the coefficients for the linear representation of a squatting human are close to zero.

These experimental results show that the following: (1) multisource data fusion samples can be sparsely represented under the overcomplete dictionary; (2) according to the distribution of the main nonzero coefficients in the coefficient vector, the combination of different classes in the fusion sample can be distinguished, and the special object (human) can be identified.

6.2. Experiment 2: Relationship between Number of Objects and Recognition Rate for Fixed Signal-to-Noise Ratio (SNR)

6.2.1. Relationship between the Number of Animals and Recognition Rate. Assuming that there is one human, scenarios including 1–5 animals are considered (see Table 2). The experimental simulation results are shown in Figure 7.

As can be seen from Figure 7, the correct recognition rates achieved by the traditional NN, KNN, SVM, and SRC are relatively low. As the number of animals increases, the traditional methods struggle to recognize the special object correctly. The proposed method outperforms these traditional methods and exhibits better anti-interference ability.

6.2.2. Relationship between Number of Humans and Recognition Rate. For the case of only one animal, experiments with 1–5 humans are considered (see Table 3). The experimental simulation results are shown in Figure 8.

From Figure 8, we can see that an increase in the number of humans enhances the recognition rate of various methods. However, the recognition rate of the method proposed in this paper is consistently 100%, which is obviously superior to that of traditional methods.

6.3. Experiment 3: Verification of Relationship between SNR and Recognition Rate.

For WSNs, reducing the transmission power of the signal can effectively extend the network lifetime. In this experiment, the number of humans and animals is known, and we analyze the impact of different SNRs on the recognition performance. The experimental condition settings are listed in Table 4, and the simulation results are shown in Figure 9.

It can be seen from Figure 9 that changes in the SNR have little effect on the recognition rates of the various methods.
The experimental results show that the proposed method outperforms the traditional methods, and we can extend the service life of the network by reducing the transmission power of the network nodes without affecting the recognition rate.

6.4. Experiment 4: Verification of Relationship between Number of Animals and False Alarm Rate.

In border regions or special monitoring areas, there may be long periods...
without humans passing through, during which only animals will be detected. Therefore, it is very important to correctly judge that no humans are present, thus reducing the false alarm rate of the system. The experimental conditions are listed in Table 5, and the simulation results are shown in Figure 10.

As can be seen from Figure 10, when the SNR is 15 dB or 20 dB, the false alarm rate of the proposed method is slightly higher than that of traditional methods. However, as the number of animals increases, the animal class information in the fusion sample increases, which reduces the possibility of false alarms. The experimental results show that an increase in the number of animals will not increase the false alarm rate, but will actually help to reduce the false alarm rate.

In summary, the traditional recognition methods mainly focus on the test samples of a single class. When the samples are composed of multiple classes, the other classes in the fusion sample produce interference that affects their correct recognition performance. The proposed method is based on sparse representation theory. When the samples are composed of multiple classes, we can effectively separate the fusion samples in the sparse domain. According to the coefficient distribution, we can then judge the class combination in the fusion sample. The experimental results verify the validity of this method from many aspects of recognition tasks.

7. Conclusions

According to the characteristics of the profile detection system, combined with WSN and sparse representation theory, we have proposed a method of multisource data processing and an associated mathematical model. Based on this model, a novel special object recognition method has been developed. Compared with traditional methods, the method described in this paper achieves better performance in many aspects of recognition. Moreover, in practical applications, the system parameters (such as the transmission power of the network nodes) can be adjusted to satisfy the actual requirements.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Key Projects of Natural Science Research in Universities in Anhui (no. KJ2017A531), the Opening Foundation of Key Laboratory of Intelligent Computing & Signal Processing (Anhui University), Ministry of Education, and Key Discipline Construction Project of Hefei University (no. 2018xk03). We thank Stuart Jenkinson, PhD, from Liwen Bianji, Edanz Group China, for editing the English text of a draft of this manuscript.

References

[1] J. Russomanno David, C. Srikan, E. Kenny et al., “Testing and evaluation of profiling sensors for perimeter security,” *ITEA*, vol. 31, no. 1, pp. 121–130, 2010.
[2] B. Sartain Ronald, A. Keith, T. Alexander et al., “Long-wave infrared profile feature extractor (PFx) sensor,” in *Proceedings of SPIE 2009, Unattended Ground, Sea, and Air Sensor Technologies and Applications XI*, vol. 7333, pp. 733311–733317, Orlando, FL, USA, May 2009.

[3] D. J. Russomanno, S. Chari, E. L. Jacobs, and C. Halford, “Near-IR sparse detector sensor for intelligent electronic fence applications,” *IEEE Sensors Journal*, vol. 10, no. 6, pp. 1106–1107, 2010.

[4] C. Srikant, H. Carl, E. Jacobs et al., “Classification of humans and animals using an infrared profiling sensor,” in *Proceedings of SPIE 2009, Unattended Ground, Sea, and Air Sensor Technologies and Applications XI*, Orlando, FL, USA, May 2009.

[5] D. Russomanno, S. Chari, and C. Halford, “Sparse detector imaging sensor with two-class silhouette classification,” *Sensors*, vol. 8, no. 12, pp. 7996–8015, 2008.

[6] R. K. Reynolds, S. Chari, and D. J. Russomanno, “Embedded real-time classifier for profiling sensors and custom detector configuration,” in *Proceedings of the SPIE, Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR II*, Orlando, FL, USA, May 2011.

[7] C.-J. Zha, N. Sun, and C. Zhang, “Special object recognition based on sparse representation,” *Journal of Jilin University (Engineering and Technology Edition)*, vol. 43, no. 1, pp. 256–260, 2013.

[8] W. Wei and X. Lihong, “A modified sparse representation method for facial expression recognition,” *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 5687602, 12 pages, 2016.

[9] E. R. Oliveros, G. Coelho, P. Marrero-Fernández et al., “Evaluation of K-SVD method in Facial expression recognition based on sparse representation problems,” *International Conference on Articulated Motion & Deformable Objects*, Springer, Berlin, Germany, 2016.

[10] J. Wright, Y. Yang Allen, G. Arvind et al., “Robust face recognition via sparse representation,” *IEEE Transactions, PAMI*, vol. 31, no. 2, pp. 1–18, 2009.

[11] D. Granato, J. S. Santos, B. L. Ferreira, G. B. Escher, and R. M. Maggio, “Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: a critical perspective,” *Trends in Food Science & Technology*, vol. 72, pp. 83–90, 2018.

[12] D. Russomanno, M. Yeasin, E. Jacobs et al., “Sparse detector sensor: profiling experiments for broad-scale classification,” in *Proceedings of SPIE 2008, Unattended Ground, Sea, and Air Sensor Technologies and Applications XI*, Orlando, FL, USA, May 2008.

[13] J. Wang, C. Luo, H. Huang, H. Zhao, and S. Wang, “Transferring pre-trained deep CNNs for remote scene classification with general features learned from linear PCA network,” *Remote Sensing*, vol. 9, no. 3, p. 225, 2017.

[14] H. Gao, H. Zhang, Z. Li et al., “Optimality analysis on partial l1-minimization recovery,” *Journal of Global Optimization*, vol. 70, no. 1, pp. 1–12, 2017.

[15] S. U. H. Qazi, L. X. Shi, L. M. Tao, and S. Q. Yang, “A l1-minimization based approach for hyperspectral data classification,” *Key Engineering Materials*, vol. 500, pp. 675–681, 2012.

[16] D. L. Donoho, M. Elad, and V. N. Temlyakov, “Stable recovery of sparse overcomplete representations in the presence of noise,” *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 6–18, 2006.