

Research on a Method of Data Aggregation of Precision Agriculture Information

Bao-yuan Chen¹, Ya-qiong Lan¹, Di Ma¹, Jing-yang Liu¹, Zhi-chao Huo¹

¹ The higher educational key laboratory for Measuring & Control Technology and Instrumentations of Heilongjiang Province Harbin University of Science and Technology, Harbin 150080, China
Chenbaoyuan@126.com

Abstract

As modern agriculture develops, the demand to the agricultural information acquirement technology grows. This paper applies the wireless sensing data acquisition technology to the agricultural production. A data aggregation method based on matching pursuit algorithm is proposed according to the features and function of WSN data aggregation. By the theoretical analysis and experimental simulation, it's proved that the data aggregation of wireless sensor network can be achieved with the application of matching pursuit algorithm. Thus both the amount of data transmission and energy consumption of the network are reduced effectively, the life of the network extends. And the high data reconstruction accuracy helps the large-scale application.

Keywords: Precision agriculture, Matching pursuit, Data aggregation

1. Introduction

Precision agriculture is an important form of modern agriculture production. It developed on the basis of a series of the latest high-tech achievements, such as modern information technology, biotechnology and engineering technology. [1, 2] While the entire core of precision agriculture is to catch the farmland environmental information, supporting the other subsystems. [3] WSN network is the most widely used application in the field of access to farmland environmental information. The WSN network is characterized by comprehensive and rapid information collection, information transmission reliability and intelligent information processing. [4] If it was applied into the monitor of large-scale agriculture production, the system will be more unified more automation and more information. It is effective to promote the productivity and reduce the system cost [5].

In this paper, the WSN network technology is applied into precision agriculture. We can take the measures such as precision irrigation, ventilation and sun shading by the surveillance of the temperature, humidity and light intensity and the requirements for the humidity that varies with the crops. [6] The energy can be saved and the crops growth rate will increase in this way. At the same time, considering the characteristics of farmland environment that the monitoring area is large, transmission distance is long, environmental information changes slowly, and the real-time is less demanded, [7] thus, this paper focus on the study of the information aggregation and transmission to achieve large-scale data collection. This paper analyzes the features and functions of wireless sensor network data aggregation, an aggregation algorithm based on matching pursuit data is proposed in this paper. The data collected by wireless sensor will be spared and the data aggregation transfer realizes. The network cost decreases and the life of network extends. The simulation proves that the spared

data that based on the matching pursuit algorithm has the high precision reconstruction and is adaptable to the large-scale collection data supervision.

2. Transmission Model of WSN using Distributed Nodes and the Features and Functions of Data Aggregation

2.1. Environmental monitoring model of Wireless network with distributed node

As we have mentioned above, WSN network can be applied into the field of precision agriculture environmental monitoring, including the parameters measuring and monitoring such as temperature, humidity, light intensity, soil composition, and the concentration of fertilizer. [8]As agricultural production requires large amounts of land, it is difficult to install sensor nodes in the complicated terrain, and the high requirements for production process automation and autonomy are needed. So things can be done by cooperation of multi-sensor nodes to complete the measurements and information processing.[9] It will improve the performance and energy effectiveness of the Internet of Things.

The WSN undertakes the data acquisition in monitoring area. Sensor nodes installing with temperature, humidity and other sensors formed MANET through Zigbee module. The sensor nodes not only collect and process the local information, but also transmit the data from other nodes to complete the monitoring together. And the collected data are sent to base station nodes through the node routing data according to the upload period. This system which is shown in Figure 1 is designed as three-tier.

a. The bottom-layer consists of the wireless sensor nodes arranged in the farmland monitoring areas. The wireless communication and the system bottom-layer access network is build up via the sensor wireless routing nodes. The goal of the system bottom-layer access network is to catch the farmland environmental parameters, such as the soil moisture, temperature and light intensity. Meanwhile, it also executes environmental information query instructions from the upper system. The character of wireless sensor network that it costs low power consumption is greatly played.

b.The middle-layer is made up of all the bottom-layer sensor nodes and aggregation nodes. It mainly gathers the environmental information on the bottom-layer. At the same time, it communicates with on-site monitoring center and other aggregation nodes to improve system scalability.

c. Upper-layer is the control layer which directly communicates with the aggregation nodes of middle-layer. It exchanges information with the remote background system through existing communication networks such as GPRS or Internet. And its main task is to memory and count the environmental information from the acquisition nodes of bottom-layer and the aggregation nodes of middle-layer. The information contains soil moisture, temperature and light intensity. The upper-layer sends the control instructions to the other two layers via the information it catches.

2.2. Environmental Monitoring Model of Wireless Network with Distributed Node

Data aggregation plays a very important role in wireless sensor network, which is the process of data aggregation is multiple of data collected by the sensor nodes for processing. After this process, we can reduce data redundancy, integrated more efficient and more accurate data. The aggregation mainly has the following features. The first one is that it can reduce the energy consumption, we can remove a large amount of

redundant information through the data aggregation. This method also can get more accurate information. Because the environmental factors based on reality is uncertain, the data from the single sensor node is unreliable, it is difficult to guarantee the accuracy of the information. In this way, the data can be aggregated in a region which can ensure the accuracy and precision of the data. The last one is that it can improve the efficiency of data collection. In general, there are thousands of nodes arranged in the sensor network. If a single node directly communicate with aggregate nodes in the large-scale network, it will result in transmission congestion and conflict collision. We can effectively alleviate the phenomenon through the data aggregation.

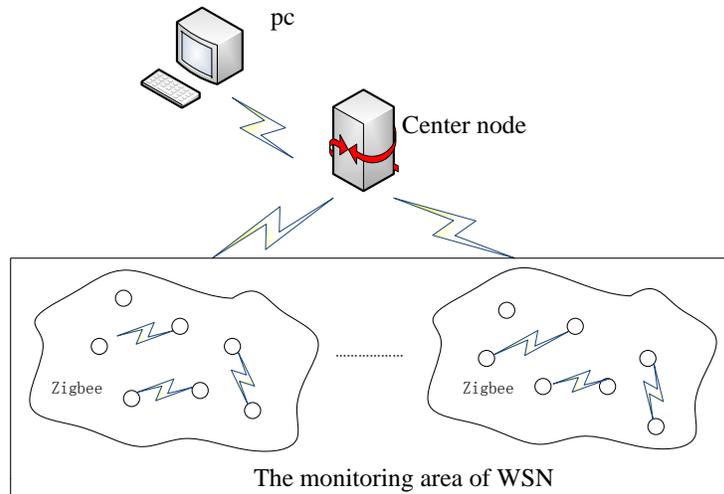


Figure 1. The Structure of WSN

3. The Information Aggregating Algorithm based on Matching Pursuit

The sparse decomposition theory of the Internet of Things signals is essentially the same as the image sparse decomposition theory. Moreover, as it contains the network topology information, the measuring data in Internet of things are more complicated. [10] We can directly adopt the matching pursuit algorithm to sparser the network data.

Matching pursuit algorithm is greedy algorithm of repeat iterative approximation. [11] This algorithm can overcome the limitations of Window Fourier transform and wavelet transform. The Window Fourier transform can not express the non-stationary signals. [12] Because all of the waveforms have the same fixed scale parameter that is in direct proportion to the window size. Wavelet transformation decomposed signals in variable metric time-frequency atoms but the wavelet transform class is designed to that the frequency parameter is defined as the inverse relationship to the scale parameter. So the frequency of waveform cannot be accurately expressed in the expansion coefficients of wavelet frame. The Fourier transformation can realize this function. In the matching pursuit algorithm, signals structure is represented by the waveforms which can match its time-frequency characteristics.

The topologies of the Internet of things can be seen as image signals, where the sensor nodes are equal to the vertices in the graph and the communication links between the nodes correspond to the edges in the graph. Take this fact into consideration that only there are communications links can the data to be exchanged and transmitted in the Internet of things.

Then only these measurement values of sensor nodes which are in the Internet of things are relevant.

We suppose that $F = (Z, L)$ is a connection diagram corresponds to a topology of the Internet of Things. $Z = [z_1, z_2, \dots, z_N]$ is the fixed points collection of this diagram and the sensor nodes collection. L means the edges collection, namely the collection of communication links. Given an over-complete dictionary used in signal sparse decomposition named $D = \{g_\gamma\}$, g_γ is defined as atoms by the parameter set. $r = (s, u, v, w)$ is the time-frequency parameter. s is scale factor, u is translation factor, v is frequency of atom and w is phase of the atom. The dilation and the translation may be represented with the distance between 2 points in the same space with Z . Because the atom dictionary with high redundancy is needed to contains a wide time-frequency property, we directly give the definition to(1), Then the time-frequency clusters shows.

$$g_\gamma = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i\zeta t} \quad (1)$$

Then $g(t)$ is normalized as $\|g_\gamma\| = 1$. If $g(t)$ is an even function, the latitude will be close to the horizontal axis u . Energy is concentrated in the vicinity of u , the size is proportional to s . As for the connection diagram F , first we choose a best matching atom g_{γ_0} from the over-completed dictionary D . and it has to meet the equation below:

$$\left| \langle f, g_{\gamma_0} \rangle \right| = \sup \left| \langle f, g_\gamma \rangle \right| \quad (2)$$

In this way, the signal is decomposed into two components: the component in the best atom g_{γ_0} and the residual component. As the following formula:

$$f = \langle f, g_{\gamma_0} \rangle g_{\gamma_0} + R^1 f \quad (3)$$

R^1 is the residual component after the matching of best atom. If we continue to decompose the residual component according to the equation (3), after N times decomposition, the signal be decomposed as:

$$R^N f = \langle R^N f, g_{\gamma_N} \rangle g_{\gamma_N} + R^{N+1} f \quad (4)$$

in which the g_{γ_N} meet the follow:

$$\left| \langle R^N f, g_{\gamma_N} \rangle \right| = \sup \left| \langle R^N f, g_\gamma \rangle \right| \quad (5)$$

It can be known from the equation (3) and (4) that after M times decomposition, the signal becomes:

$$f = \sum_{N=0}^{m-1} \langle R^N f, g_{\gamma_N} \rangle g_{\gamma_N} + R^m f \quad (6)$$

In which the $R^m f$ is the residual component after M times decomposition, it also named as approximation errors. Since the atom that is best matching the residual signal is chosen in every step of decomposition. Every residual component will decay rapidly as the

decomposition times increases. It was proved in some documents that as the N increases, the norm $\|R^N\|$ will fall to 0. The signal is decomposed as:

$$f = \sum_{N=0}^{\infty} \langle R^N f, g_{\gamma_N} \rangle g_{\gamma_N} \quad (7)$$

It can be known from the equation 7 that the signal low-energy component is set to zero by matching pursuit algorithm and then sparser transformation signals are obtained. The data after sparser are sent to the central processing nodes of the Internet of Things. The central processing nodes aggregate and send the data.

4. The Simulation of Information Aggregating Algorithm based on Matching Pursuit Algorithm

The parameters captured by data acquisition system of WSN are shown in Table 1. Taking humidity signal for an example, we validate the characteristics of data aggregation transmission in the WSN based on matching pursuit algorithm.

Table 1. The Parameters Captured by Data Acquisition System of WSN

measurement Items	measurement range	measurement precision
Temperature/°C	-40~120	± 0.4
Humidity RH/%	0~100	± 3
Light intensity /lk	0~200 000	± 20
Wind velocity/m/s	0~60	± 0.6
The concentration of CO ₂ /mg/L	0~2000	± 50
pH	0~14	0.1

We choose a greenhouse sized $300 \times 300 \text{ m}^2$ as the monitoring area. 512 wireless nodes are installed in the WSN to collect data. The sensor nodes are used to periodically measure the soil humidity. Transmit the measured data of soil humidity from all sensor nodes to the processing nodes in WSN center, and then we can get the overall information about the soil humidity in the monitoring area. We adopted the data aggregating algorithm based on matching pursuit algorithm to achieve the data aggregating of all sensor nodes in this experiment.

According to the topology connections characteristics of the Internet of Things, we use the Gabor over-complete dictionary in the process of sparse decomposition to the collected data as the Gabor atoms can represent the singularity signal correctly. $D = \{g_{\gamma}\}$ consists of the Gabor atoms and each atom consists of a Gaussian window function. The function is shown below.

$$g_{\gamma}(x, y) = \frac{1}{2\pi\alpha\beta} e^{-\pi \left[\frac{(x-x_0)^2}{\alpha^2} + \frac{(y-y_0)^2}{\beta^2} \right]} e^{i(ux+vy)} \quad (8)$$

You can have the access to the atoms which combines Gabor over-complete dictionary through time-frequency parameters discretization.

Figure 2 is the information about the collected soil moisture data. The main program of the matching pursuit algorithm is programmed in Matlab. The number of the iterations is 80, as the reconstruction signal basically approximates the original signal after 80 times of iterations. Figure 3 shows the residual signal after 80 times of iteration. Figure 4 is the main signal, also named the reconstruction signal.

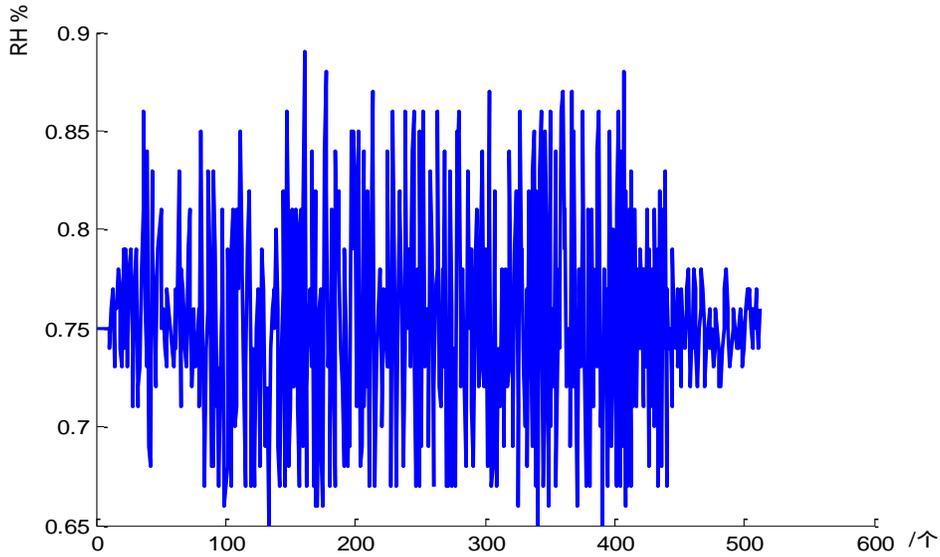


Figure 2. Original Signal

Figure 3 shows the residual signal after 80 times of iteration.

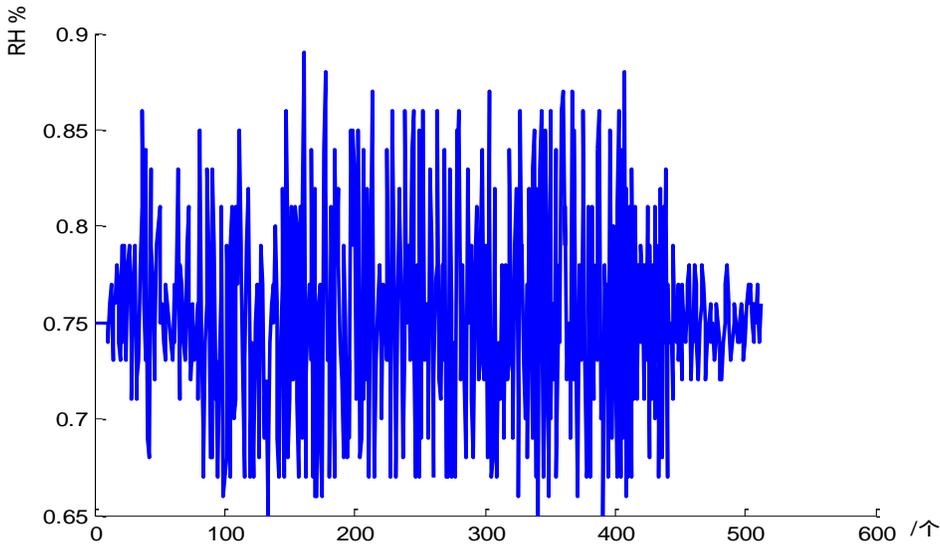


Figure 3. Residual Signal

Figure 4 is the main signal, also named the reconstruction signal.

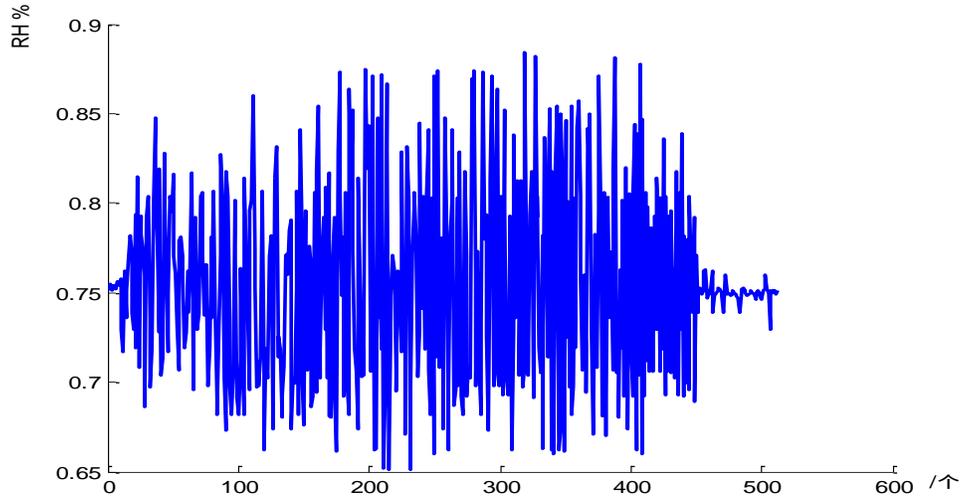


Figure 4. Reconstruction Signal

As it can be seen from Figure 2 and Figure 4 that the reconstruction signal after iteration and the original signal is pretty much the same. 512 sensor nodes were numbered to verify the accuracy of the reconstruction signal. We did 100 times periodic measures using 5 feature nodes. The average reconstruction error of sensor network data was calculated according to the following equations:

$$\sqrt{\sum_{i=1}^{100} \delta_i^2 / 100} \quad (9)$$

$$\delta_i = \chi - \varphi \quad (10)$$

In this equation, χ stands for the sensor nodes after reconstruction and φ is the measured sensor nodes. The result is shown in Table 2.

Table 2. The Error Between the Measurement Data and the Reconstruction Data

feature nodes	100	200	300	400	500
average error	0.047	0.068	0.056	0.087	0.137

As it was known from the above table that the matching pursuit algorithm can realize the data aggregation transmission and greatly reduces the data traffic. In addition, the low average error of the reconstruction data can meet the demand of agriculture monitor data.

5. Conclusion

This paper primarily focuses on the precision agriculture. To meet the actual demand of precision agriculture, the environmental monitoring system based on wireless sensor networks is designed. And according to the characteristics and function of data aggregation in wireless sensor networks, matching pursuit algorithm is proposed. The research on matching pursuit algorithm proves that this algorithm can achieve the data rapid aggregation transmission. The application of the algorithm can not only reduce the amount of network

data transmission and network energy consumption but also prolong the network life. The testing and simulation shows that the after the data aggregation using matching pursuit algorithm, the data reconstruction accuracy is high and the validity of data transmission is ensured. It can be applied into large-scale data collection in farmland environment.

Environmental monitoring system of precision agriculture based on wireless sensor networks is still in the prototype stage, so it does not be used widely yet. As the network scale expands, it will followed by many problems like the response time of the system, data transmission delays and other issues. Further discussion and researches are needed.

Acknowledgment

Heilongjiang Province College Education Engineering Projects. Grant No: JG2012010282 and JG2013010302.

References

- [1] K. R. Varshney, M. Ceti, J. W. Fisher and Alanswillsky, "Sparse Esentation in Structured Dictionaries with Application to Synthetic Aperture Radar", IEEE Transactions on Signal Processing, vol. 8, (2008).
- [2] S. Sulaiman, A. Manut and A. R. Nur Firdaus, "Design,fabrication and testing of fringing electric field soil moisture sensor for wireless precision agriculture applications", International Conference on Information and Multimedia Technology, (2009).
- [3] A. Jayaashree, G. S. Biradar and V. D. Mytri, "Review of Multipath Routing Protocols in Wireless Multimedia Sensor Network", Internarional Journal of Scientific&Engineering Research", vol. 3, (2012).
- [4] Y. Li, X. Xu, B. Bai and Y. Zhang, "An Improved Combined Muti-scale Method for Image De-noising", Electronic Technology, vol. 4, no. 17, (2008).
- [5] J. Emmanuel, M. Candes and B. Wakin, "An Introduction to Compressive Sampling", Signal Processing Magazine, IEEE, (2008).
- [6] Y. H. Wang, "Seismic time frequency spectral decomposition by Matching Pursuit Geophysics", vol. 1, no. 72, (2007).
- [7] D. Needel and R. Vershynin, "Uniform Uncertainty Principle and Signal Recovery via Regularized Orthogonal Matching Pursuit", Foundations of Computational Mathematics, vol. 3, (2009).
- [8] R. Baraniuk, V. Cevher, M. Duarte and C. Hegde, "Model-based Compressive Sensing", IEEE Transactions on Information Theory, (2010).
- [9] E. J. Cand'es, Y. C. Eldar and D. Needell, "Compressed Sensing with Coherent and Redundant Dictionaries", Applied and Computational Harmonic Analysis, (2011).
- [10] M. A. Davenport, P. T. Boufounos and M .B. Wakin, "Signal processing with Compressive Measurements", IEEE Journal of Selected Topics in Signal Processing, (2010).
- [11] W. Dai and O. Milenkovic, "Subspace Pursuit for Compressive Sensing Signal Reconstruction", International Symposium on Turbo Codes and Related Topics, vol. 5, (2008).
- [12] T. D. Thong, L. Gan, N. Nguyen and T. D. Tran, "Sparsity adaptive matching pursuit algorithm for practical compressed sensing", Asilomar Conference on Signals, Systems, and Computers, (2008).