Chapter 8
Agricultural Information Processing Technology

Hui Fang, Bingquan Chu, Yong He, and Xiya Zhang

Abstract Agriculture, as it grows increasingly information-based, will enter the development stage of smart agriculture. At the same time, research on agriculture-related information also demonstrates that it is the future trend for agriculture to become intelligent, precise, standardized, and digitized. This chapter offers a detailed elucidation on the concept of agricultural information technology, key technologies in agricultural information processing technology, and multi-source agricultural information fusion processing technology.

Keywords Agricultural Information · Information processing technologies · Multi-source · Information fusion technologies

8.1 Introduction

Agricultural information processing technology, used to obtain information that relate to agricultural activities through various channels, relies on information processing technology to sort, analyze, process, and mine information, thus facilitating decision-making and providing the theoretical basis for intelligent agricultural control. Agricultural information processing technology is the intersection of multiple disciplines. It shares the complexity of agricultural production and relies on the basic theories and technologies of information processing.

According to the level of intelligence, agricultural information processing is divided into two categories: basic agricultural information processing technology and intelligent agricultural information processing technology. This section, focused on these two categories, illustrates the basic concepts, key technologies, and development trend of agricultural information processing.

H. Fang · B. Chu · Y. He (✉) · X. Zhang
College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, Zhejiang, China
e-mail: hfang@zju.edu.cn; bqchu@zust.edu.cn; yhe@zju.edu.cn; 21913015@zju.edu.cn

Y. He et al. (eds.), *Agricultural Internet of Things*, Agriculture Automation and Control, https://doi.org/10.1007/978-3-030-65702-4_8

8.2 Basic Concepts of Agricultural Information Processing Technology

8.2.1 Characteristics and Types of Agricultural Information

Agricultural information features both generality and particularity:

- Region-specific. Agricultural operations, susceptible to the external environment, are region-specific. On the one hand, the differences in terms of regional and natural conditions determine the scope and extent of agriculture; and on the other hand, the differences in socioeconomic conditions determine the distribution, structure, management style, and production level of agricultural resources. Therefore, agricultural information is subject to a large number of regional constraints.
- Timeliness. In the process of production, the timeliness of information is critical. For example, with relation to the prevention and control of diseases and pests, though prevention and control measures that are adopted too early may be effective, those that are adopted too late means unnecessary prevention and control which often causes crop losses. Therefore, the timeliness of agricultural information is particularly important.
- Periodic. Most agricultural information regarding the life process of crops are presented in the form of different growth periods, and the requirements for the same element in different growth periods differ significantly. For instance, the general absorption patterns of nitrogen, phosphorus, potassium, and other elements in different growth stages of corn can be summed up in the following way. At the seedling stage, growth is minimal, and the amount of absorption is also small; with the increase of the growth at the early stage, absorption increases and accelerates, reaching the peak at flowering stage; and at last, the organic nutrients are concentrated and transported to the grains, from the flowering to the filling stage.
- Comprehensiveness. Agricultural activities are complex. An agricultural phenomenon is the synthesis of multiple factors. For example, the growth of crops is a comprehensive reflection of multiple factors, such as soil, fertilizer, moisture, and temperature, and the estimated price trend of agricultural products is the result obtained via comprehensive analysis of massive data in multiple markets in a certain period.
- Hysteresis. As organic and living bodies, crops are resistant to environmental changes, manifested as lags of information. For example, it takes a period of time for fertilizers to be effective, and it also takes time for deficiencies of a given nutrient to manifest itself. Hence, the processing of related information must take timeliness into consideration.

8.2.2 *Agricultural Information in IoT*

8.2.2.1 Composition of Agricultural Information

Agricultural IoT mainly includes three levels: the perception layer, the transmission layer, and the system application layer. The first layer is the perception layer. Sensor nodes including RFID barcodes, sensors, and other devices are used to achieve real-time and dynamic sensing, rapid identification, and information collection. More specifically, the primary information collected by the perception layer include information that relate to farmland environment, soil, and nutrients and physiology of crops. The second layer is the transmission layer, which realizes long-distance wireless transmission of information collected. In agricultural IoT, this layer is mainly reflected as the collection and transmission of large-scale farmland information. The third layer is an application system that provides intelligent management through data processing. When it is adopted in combination with agricultural automation equipment, the system will be able to save resources, protect the environment, and improve product quality and yield. The three levels of agricultural IoT allow it to fully sense information, reliably transmit data, effectively optimize systems, and intelligently process information (Nie 2012).

Based on the analysis of the major areas for the application of agricultural information technology and the main sources that generate big data, it can be concluded that agricultural IoT information mainly involves data in the following areas:

Big data on production process management, including facility planting, facility breeding (livestock and aquaculture, etc.), precision agriculture, etc. It is an urgent task for information-based agriculture to improve precise monitoring, intelligent decision-making, science-based management, and regulation of the entire production process.

Big data on agricultural resource management, covering land, water, biological resources, production materials, etc. Agricultural resources in China are scarce, and the environment and biodiversity are degraded; hence, we must further optimize the allocation and use of resources to achieve sustainable agricultural development with high yield, high quality, and energy efficiency.

Big data on agricultural environment management, covering soil, atmosphere, water quality, meteorology, pollution, disasters, etc. Comprehensive monitoring and precise management are needed in this regard.

Big data on agricultural and food safety management, including production environment, industrial chain management, pre-production, post-production, storage and processing, market circulation, logistics, supply chain, and traceability systems.

Big data on agricultural equipment and facility monitoring, covering the monitoring of equipment, remote diagnosis, service scheduling, etc. In the above applications, the environment and resources, production process, product safety, market conditions, and monitoring and prediction of consumption are the key.

Big data generated by research activities, including remote sensing data on space and ground, and data generated by biological experiments, such as gene maps, large-scale sequencing and genomes, macromolecules, drug design, etc.

8.2.2.2 Basic Processing Methods of Agricultural Information

The basic processing method of agricultural IoT information is the concrete embodiment of agricultural information technology in agricultural applications. In view of the characteristics of the agricultural information, the basic methods of agricultural IoT information processing include data storage, format conversion, data query, retrieval, deeper data analysis, and mining. According to the degree of intelligence of agricultural information processing technology, it can be divided into basic agricultural information technology and intelligent agricultural information technology.

Basic agricultural information technology refers to the establishment of various types of agricultural databases and the combination with computer networks, mobile Internet, "3S" technology, and other technologies. Its main function is to provide dynamic information. Intelligent agricultural information technology is an important means of combining modern information technology and agricultural science and technology on the basis of basic agricultural information technology to realize the effective transmission, rational analysis, and intelligent application of agricultural information. The specific content includes agricultural expert system and agricultural intelligent decision support system.

This chapter mainly deals with basic agricultural information technologies.

8.3 Key Agricultural Information Processing Technologies

Key technologies mainly involve agricultural data storage technology, data search technology, cloud computing technology, geographic information system, image processing technology, and standardization of agricultural IoT data.

8.3.1 Data Storage Technology

In agriculture-related production and management, a large amount of data will be generated. These data, while guiding current production, come with reference values for production in later stages. The application of IoT has brought agriculture into the era of big data. If information is not properly stored, it will lead to loss of data or ineffectiveness thereof. In data processing, the basic links are the collection, storage, classification, retrieval, and transmission of data, which are also called data management. Data storage technology is a specialized technology for managing complex data that come in lager numbers.

8.3.1.1 Concept of Database

A database is a warehouse that organizes, stores, and manages data according to the data structure. Data in the database is organized, described, and stored according to a certain data model that features less redundancy, higher data independence, and easy scalability, and such data can be shared among users. The following constitutes a brief introduction of five types of databases, namely, document database, time series database, spatial database, relational database, and graphic database.

Document Database

As a type of database system that stores and manages a large number of structured documents, document database not only provides functions for document expression, organization, storage, and access but also offers deep processing functions for documents, such as text mining and automatic abstracts.

The major difference between document database and traditional database systems is that the operations in the traditional database are deterministic data queries, while those in the document database are based on semantically relevant queries in addition to the exact matching of strings. Furthermore, in traditional databases, information is divided into discrete data segments, while in document databases, documents are themselves the basic unit for processing information.

Time Series Database

Time series database is composed of time series, which is an ordered set of sequence values or events that change over time.

Time series database is defined relative to the static database. There are two differences between them. First, the static database includes a series of records, and the order of the records is arbitrary. However, in a time series database, records cannot be arranged arbitrarily, and certain attributes in the records are time-stamped; thus the arrangement is inherently related. Second, in a static database, attributes are independent of each other, while in a time-series database, attributes are a function of time and are related to each other.

Spatial Database

Spatial database is the sum of application-related geospatial data stored on computers via a geographic information system. Spatial database is generally organized in the form of a series of files with a specific structure. Traditional databases are not suitable for the representation and storage of spatial data, and cannot support complex objects (such as graphics and images). Compared with traditional database systems, spatial databases suffer from the disadvantages of massive data sets, high accessibility, and complex spatial data models.

Spatial data are express in two ways: vector and raster. The raster method divides the two-dimensional space of the geospatial entity into regular subspaces. Therefore, raster data at different levels share a common basic element (subspace) division. Its advantages are simple structure and easy operation, and its disadvantages include large data storage, low accuracy, difficulties in establishing the topological relationship between features, and complicated operation of a single target. The vector method describes the spatial elements of geospatial entities with points, lines, and surfaces and establishes clear spatial relationship between geospatial entities. Its advantages are high accuracy, small data storage, easy expression of the topological relationship between entities, and easy manipulation of a single target. The disadvantage is that the data structure is complex, and it is generally time-consuming to perform a large number of calculations during spatial analysis and overlay operations (Liu 2012).

Relational Database

The relational model of relational database is "one-to-one, one-to-many, many-to-many." The model refers to the two-dimensional table model; hence a relational database is a data organization composed of two-dimensional tables and their connections. As they are based on the Structured Query Language (SQL) used, these database systems are widely known as SQL databases. The current mainstream relational databases include Microsoft SQL Server,[1] Microsoft Access, Oracle, My SQL, DB2, etc.

Data in a relational model is usually represented by a database table (schema), and objects of the same type (i.e., the same number of attributes with the same type and format) are combined in a table for data structuring. Relational databases store data in rows, and each row has the same number and type of data columns. In addition, data in a relational database is typically normalized, resulting in the creation of multiple tables. Moreover, querying data distributed across multiple tables requires reading and integrating information from one or more different tables. The process of integrating information is based on the matching values of the primary and foreign keys of multiple tables in a relational database, a process called joining tables.

In addition, a major property of relational databases is the ACID principle, which defines a set of rules to ensure data integrity. ACID is an abbreviation of the four basic elements for the correct execution of database operations, that is, atomicity, consistency, isolation, and durability (Wu 2015).

8.3.1.2 Open-Source Database and Its Agricultural Application

The open-source database features low costs, strong performance, open-source code, easy usage, and multi-platform support. Furthermore, this type of database ensures high reliability, high scalability, and multi-language support; thus it is widely used in small and medium-sized enterprises. The scales of agricultural information are large, and the sources are restricted by regions. Generally speaking, the distribution of agricultural information is relatively scattered, and distributed organization and management have become a critical approach. At present, most agricultural resource management platforms still depend on traditional distributed

[1] *Disclaimer*: Commercial products are referred to solely for the purpose of clarification and should not be construed as being endorsed by the authors or the institution with which the authors are affiliated.

database to manage data. Its efficiency is low, the storage capacity of the system is limited by the capabilities of the database management system it depends on, and the support for the management and release for data resources is weak (Yang et al. 2011). Because of the fact that distributed storage and management of various data on a distributed open-source database are scalable and highly fault-tolerant, this approach has become the trend of agricultural databases in the future.

8.3.1.3 Mainstream Database Products and Their Agricultural Applications

The market currently abounds with database products, and the choice of appropriate databases for IoT products is something that ought to be taken into consideration.

In the twenty-first century, agricultural information technology plays an increasingly important role in agricultural development. The application of agricultural technology, especially IoT technology, will accelerate agricultural development, allowing it to enter the age of information. Information in agriculture are relatively large, wide, and scattered. Modern information technology composed of computer technology and communication and network technology has fundamentally transformed the production, transmission, and acquisition of information, and networks have become the largest distribution center for information. The prerequisite of building an information network is to establish a layer of database management software, located between users and databases, that provides a way for users or applications to access databases for functions that include the establishment, query, and update of data. Databases play an important role in strengthening the foundation of agriculture, and the construction of database is one of the major building blocks for information systems in agriculture. Such databases achieve computer-enabled management of agricultural resources, thereby improving the speed of querying, processing, and sharing of information and providing services for production management, policy formulation, and promotion of related research in agriculture.

8.3.2 Data Search Technology

With the development of Internet technology, agricultural information are becoming ever more abundant in both number and type. For example, agricultural information in China is distributed and stored in the form of databases, research websites, and government web pages. Although great progress has been made with regard to information-based agriculture, there are still many problems, including insufficient utilization of data, data stored in the physical state that lie idle, etc. The rational and effective use of these information resources would provide considerable technical support for agricultural development. Data search technology is a technology that adopts search engines to automatically collect information from the Internet and organize and process the original document for users to query. It is noteworthy that what lie at the core of data search technology are search engines.

8.3.2.1 Search Engine

Search engine refers to a system that collects information from the Internet using specific computer programs in accordance with certain strategies. After organizing and processing the information, search engines provide users with retrieval services and display the retrieved information. Search engine, as a retrieval technology that works on the Internet, aims to improve the speed at which people obtain information and to provide people with a better network environment.

Finding information in such a vast ocean of information on the Internet is like "seeking a needle in a haystack," and search engine is the technology that emerged to solve this problem. Search engines use certain strategies to collect and discover information on the Internet; to understand, extract, organize, and process information; and to provide users with retrieval services for navigation. Search engines work in the following steps.

Collecting Information
The collection of information is basically automatic for search engines, as they use automated search robot programs called spiders to connect to hyperlinks on every web page. The robot program is based on the hyperlinks of web pages that lead to other web pages. These robots start from a few web pages and become connected to all other web pages. Theoretically, a robot can traverse most web pages if there are proper hyperlinks on the web pages.

Organizing Information
The process by which search engines organize information is called "indexing." Search engines save the information they gather and arrange them according to certain rules. By doing so, search engines are able to identify the required information without having to reexamine all the information in a timely fashion.

Receive Inquiry
Users issue a query to the search engine, which then accepts the query and returns information to users. Search engines receive queries from many users at all times, almost simultaneously. They check their index according to the requirements of each user, find the information the user needs within short periods of time, and return it to the user. Currently, the results returned by search engines are mainly in the form of web links. Through these links, users can reach the web page where the information they need is located.

Search engines use certain strategies to collect and discover information on the Internet; to understand, extract, organize, and process information; and to provide users with retrieval services for navigation. Depending on different methods of collecting information and providing services, search engines are divided into directory search engines, robot search engines, and metasearch engines.

Directory search engines collect and classify Internet information according to certain standards, compile them into corresponding directories, and manage directories in a hierarchical and successively itemized manner. Searching information can be entered layer by layer, and the desired information can be identified at last. Yahoo is a typical categorized directory search engine. Such engines are characterized by the addition of artificial intelligence, which leads to higher levels of information accuracy, but the range of such search engines is small and low. Such engines are extremely convenient when retrieving information with obvious professional characteristics. However, the quality of information classification and the user's understanding of information categories will directly affect the query results.

Robot search engines use their internal search robot spider programs to automatically search the content of large and small websites on the Internet. According to the principle of webpage relevance, a corresponding relationship is established between each keyword and all related web pages and is stored in their web servers' database. As long as the user enters a keyword, he can find all indexed web pages that match the characteristics of the keyword, and a brief introduction to the search result is provided in the form of a hyperlink. Users may click on the link to enter the corresponding web resource website and acquire the required information. The search results are usually millions, but the more relevant the information, the higher its position in the search result list. The typical keyword full-text search engine is Google.

Metasearch engine do not have its own index data; instead, it submits user query requests to multiple search engines at the same time; processes the returned results after repeated elimination, reordering, and so on; and returns them to the user as its own results. This type of search engines includes worldwide search, Search, Dogpile, etc. The advantage of metasearch engines is that they can search multiple search engines at the same time, which improves the scope of the query to a certain extent. The disadvantage is that sometimes it is not possible to check all information, and critical information can be omitted.

8.3.2.2 Data Search and Agricultural Application

More than 30,000 agricultural-related websites exist in China, all of which provide rich information on production, markets, regulations, policies, and technologies. The above three types of search engines could provide a handy tool for collecting and using such information, which could bring new vitality to the development of agriculture. Because users of such information are farmers, and considering their educational level and computer operation skills and the complexity of relevant information, the above three agricultural search engines have the following two shortcomings in practical application. First, users are unable to use simple keywords to find the required information. Second, search engines only return thousands of links of web pages, and there is still a long way to go before users can get the information they need. Therefore, specialized and intelligent engines suitable for agricultural information search will be the mainstream of future agricultural information search engines.

8.3.3 Cloud Computing Technology

Cloud computing technology, originated from search engines, is a computing technology developed by Internet companies in the pursuit of low cost and high efficiency. It has now become an important platform for Internet services.

Through cloud computing technology, network providers can form and process tens of millions of data in a matter of seconds, giving rise to network servers with the same powerful functions as supercomputers. On the one hand, cloud computing has changed the traditional resource allocation and distribution pattern and transformed the mode of providing information services on the other. After years of development, cloud computing has gradually substantialized (Zheng 2015).

8.3.3.1 Principles of Cloud Computing

The basic principle of cloud computing is to distribute computing to a large number of distributed computers, rather than local computers or remote servers, and the functioning of corporate data centers is similar to the Internet, allowing companies to allocate resources to corresponding applications and obtain access to computers and storage systems in a demand-specific manner.

Figure 8.1 illustrates a typical cloud computing platform. Users can select service items in the service catalog from the interactive interface services provided by the cloud user terminal, schedule the corresponding resources through the request management system, and distribute requests through configuration tools and configure web applications.

- The service directory is the directory listing that users can access. Users can also modify their own inventory directories.
- The management system and configuration tools are mainly responsible for user login, authentication and authorization, management of available computers and resources, procedures for receiving user requests and forwarding them to the corresponding applications, and dynamic deployment, configuration, and reclaiming of resources.
- The monitoring and statistics module is responsible for monitoring the usage of cloud system resources, so that node synchronization configuration, load balancing configuration, and resource monitoring can be completed in a reasonable and timely manner to ensure the rational allocation of resources.
- The computing/storage resources are virtual or physical servers that are used to respond to user needs.

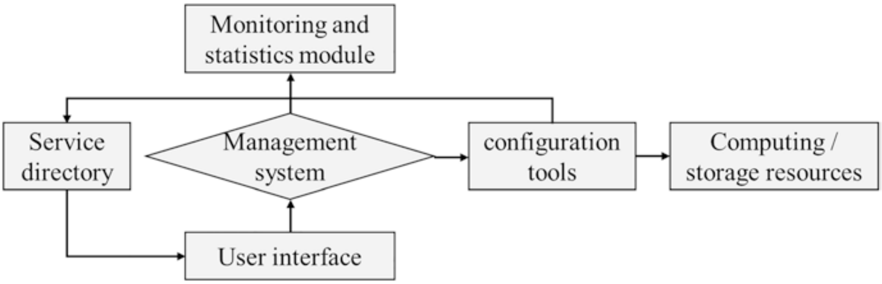

Fig. 8.1 Typical cloud computing platform

8.3.3.2 Type of Cloud Computing

Cloud computing includes two components, namely, cloud platform and cloud service. Cloud platform refers to hardware-based services that provide computing, network, and storage capabilities, whereas cloud services refer to services that are based on abstract infrastructure and can be flexibly extended. It is noteworthy that cloud services are not necessarily based on cloud platforms.

Depending on whether the service is publicly released, cloud computing can be divided into public cloud, hybrid cloud, and private cloud. There is no essential difference between the three in terms of technology, and they only differ in the operation and scope of application. Public cloud refers to companies using cloud platform services operated by other companies or organizations, private cloud refers to companies operating and using cloud platform services, and hybrid cloud shares some of the features of both (Zheng 2012).

Depending on the type of service, cloud computing can be divided into infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS). As shown in Fig. 8.2, each layer in the computer network implements a certain function, and there is a certain relationship between layers. The layers in the cloud computing system can be divided, and each layer can complete a user's request independently without the support and services of the other layers.

IaaS

Infrastructure as a service (IaaS) provides virtual hardware resources in the form of services, such as virtual hosting, storage, networking, and database management. Users do not need to purchase servers, network equipment, and storage equipment; instead, their own application systems can be built simply by renting the abovementioned equipment on the Internet. Typical IaaS includes Amazon's elastic cloud (Amazon, EC2) (Zheng 2012). The IaaS layer mainly provides low-cost and high-performance data centers and reliable infrastructure services, including IBM's Wuxi Cloud Computing Center, Century Interconnection's CloudEx cloud host, etc.

PaaS

Platform as a service (PaaS) provides application service engines, including Internet application programming interfaces and operating platforms. Based on application

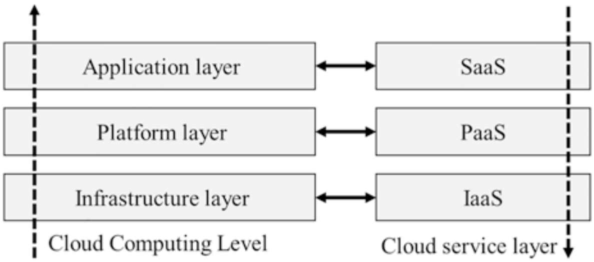

Fig. 8.2 Cloud computing service types

service engines, users can build this type of applications, such as Google App Engine and Microsoft Azure. At the PaaS level, service providers provide encapsulated IT capabilities, such as databases, file systems, and application operating environments, which are usually billed based on user logins.

SaaS

The software as a service (SaaS) layer, focused on the end users of cloud computing, provides Internet-based software application services. SaaS is the most common form of cloud computing at present. In SaaS, service providers are responsible for maintaining and managing software and hardware facilities, and users are able to get access to software through the Internet. Furthermore, users do not have to make any purchase; rather, they only need to rent the software they need. Typical services of this type include Google's online office software and 800 apps (Zheng 2012).

8.3.3.3 Cloud Computing and IoT

A large number of data is produced in agricultural production, covering seed selection, cultivation, harvesting, classification, and processing. However, the cost of implementing complex algorithms on hardware or software is relatively high, and because farmers have lower incomes, they cannot afford the costly modern equipment. The introduction of cloud computing into agriculture, in combination with IoT technology, builds an agricultural data cloud that reduces costs, improves efficiency, customizes as needed, saves resources, and promotes the modernization of agriculture.

Advantages of Reliability

As the amount of data generated in agricultural operations continues to ramp up, the number of servers in IoT has also increased, and as the number of servers continues to increase, the probability of errors in server nodes rises. Via cloud computing, redundant backup technology can be used to repair error server information, which significantly improves the reliability of agricultural operations.

Advantages of Cost

The server's hardware capacity is limited, and it crashes when the trafficking exceeds the limit. Additionally, access to data in IoT is uncertain and dynamic. Because decision-making in production requires comprehensive analysis of data from previous years, as time goes by, the huge amount of information accumulated can be very demanding on server hardware. The use of cloud computing technology can dynamically increase the number of servers in the cloud, which not only reduces production costs but also meets the need of IoT access at any time.

Advantages of Computing Power and Storage Power

Cloud computing uses parallel technology, distributed computing technology, and grid computing technology to integrate multiple computer entities into a powerful computer system through the network. It is capable of providing powerful computing functions just like supercomputers, and cloud storage facilitates the storage of massive agricultural information.

Advantages of Data Mining

The data generated by IoT (including sensor data, RFID data, two-dimensional code, video, pictures, etc.) is fairly complicated. In the process of production, IoT data are real-time and uninterrupted. As time goes by, the amount of data continues to increase, and there might not be any upper limit to it. Cloud computing, the brain of IoT mining, guarantees distributed parallel data mining and efficient real-time mining. As a common mode of data mining, cloud services ensure the efficiency of data mining, realize the sharing of data mining, and lower the technical threshold of data mining. Cloud computing, the core of IoT, achieves the real-time and dynamic monitoring and IoT-enabled management, making intelligent analysis possible.

Integration of Cloud Computing and IoT

IoT is the result of deepening developments of information, and cloud computing is bound to be created as IT continues to mature. Cloud computing provides various services in a virtualized way, and even IoT itself exists in a cloud-based manner. In a way, IoT needs the help of cloud computing to solve certain problems, and IoT itself is a form of cloud computing application. The integration of cloud computing and IoT in the future will certainly provide strong support for the transformation of agricultural production.

8.3.4 Artificial Neural Network Technology

Artificial neural network (ANN) is a complex network with many simple processing units (i.e., neurons) interconnected with learning, memory, and induction functions. It is a mathematical simulation of the nervous system of the human brain that is relied on to achieve learning. Moreover, ANN imitates the way in which signal processing is done in human brains.

Artificial neural network is a highly nonlinear, mega-scale, continuous, and dynamic system. In addition to the common characteristics of nonlinear dynamic systems, other major characteristics of ANN include continuous nonlinear dynamics, holistic concept of networks, large-scale parallel distributed processing, high robustness, and ability of learning and associations.

8.3.4.1 Principle of Neural Network Technology

Artificial neural network (ANN), a mathematical model that simulates biological neural networks for information processing, consists of simple and adaptive information processing units connected through a topological structure in a massively parallel manner. Depending on the differences in network topology, neuron transfer functions, learning algorithms, and system characteristics, ANN models come in many forms. Among them, back propagation (BP) neural network is currently one of the most widely used artificial neural networks. Its structure is divided into input layer, hidden layer, and output layer. The core idea is to adopt error correction learning algorithms (also known as delta learning rules) for the adjustment of network parameters, thus continuously improving the accuracy of the network's response to input patterns.

8.3.4.2 Mathematical Model of Neuron

The structure and model of artificial neurons are shown in Fig. 8.3. Regarding the *i*th neuron in the figure, when the sum of time is not considered, the change of the membrane potential u_i is equal to the linear combination of the input signal I and the weight$W$:

$$u_i = \sum_{k=1}^{n} W_k I_k \tag{8.1}$$

According to the change of membrane potential u_i and the activity x_i at this time, the activity $x_i(t)$ of the *i*th neuron at time t can be determined:

$$Q_i(t) = f_i(x_i) \tag{8.2}$$

$f_i(x_i)$ is an arbitrary function containing the threshold value, also known as the output function, and is determined according to different models. $Q_i(t)$ is the output of the *i*th neuron.

The mathematical theory of neural networks is essentially a nonlinear one, and BP neural network can be regarded as a nonlinear mapping between input and output sets. To realize such a nonlinear mapping relationship, it is not necessary to study the entire internal structure of the system, and one only needs to simulate the internal structure of the system by examining a limited number of samples.

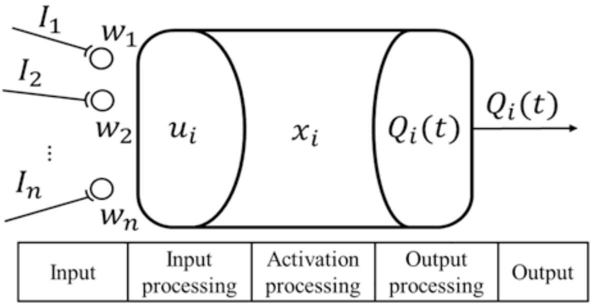

Fig. 8.3 Artificial neuron model

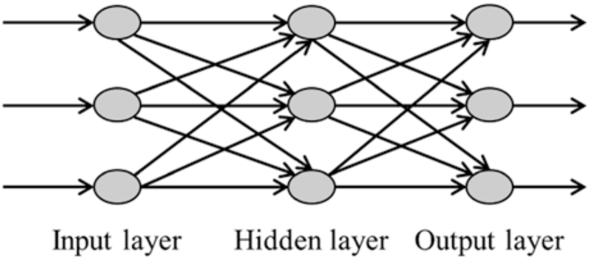

Fig. 8.4 BP network structure diagram

As shown in Fig. 8.4, the standard BP model consists of three neuron layers: the input layer, the hidden layer, and the output layer. BP network, deemed as a nonlinear mapping between input and output sets, uses offline multilevel error correction gradient descent to learn offline. Error back propagation learning only implements gradient descent on the surface of the cost function. Due to the existence of nonlinear hidden units in BP networks, multiple minima exist in the cost function. Therefore, gradient descent cannot guarantee the holistic minimum. It can be proved that in the case where the hidden layer nodes can be freely set as required, a three-layer BP network can be used to approximate any continuous function with arbitrary precision. In addition, there are many classic CNN models, such as LeNet5, AlexNet, VGG, GoogLeNet, and ResNets.

8.3.4.3 Application of Artificial Neural Network in Agriculture

A neural network model is able to learn from past events, and a superior network model is capable of making predictions based on experience. In order to improve the performance of agricultural remote sensing image retrieval, Ye et al. (2019) proposed a remote sensing image retrieval method using convolutional neural network classification capabilities. In Ye's research, the fine-tuned convolutional neural network model is used to extract the retrieval features of the query image and the weight of each category. Then, the category accuracy and the initial ranking results of the retrieved image determined by the CNN model are used to calculate the category precision. The test results demonstrated that the average accuracy of the retrieval method in the PatternNet data set reached 97.56%.

Neural networks can perform classification processing. In agricultural engineering, many classification problems exist, including the hierarchical classification of agricultural products. Given that the classification index involves volume, size, shape, weight, color, biological characteristics, and so on, many products could only be classified and graded manually in the past. The computer vision system is used to input the product appearance information into the computer. Images are preprocessed according to the classification characteristics of the image and are then forwarded to the neural network for learning. The learning internal network system, used for classification, yields satisfactory results.

Neural networks can be applied to process control. Though it is difficult to describe the production process of many processing industries with accurate mathematical models, neural networks can learn from the data accumulated in the production process, allowing neural networks to fully reflect the impact of process control variables on the production process and predict product quality. Therefore, neural networks can be used to control the production process in real time, thus optimizing the production process. Wang (2016) extracted the acoustic signal characteristics related to egg cracks and used artificial neural networks to construct a discriminant model to verify the discriminant effect and designed an FPGA + DSP-based online detection system for egg cracks. The artificial neural network is trained with all 26 features and filtered features, respectively, and the results are compared. The results demonstrate that this method significantly reduces the amount of calculation while ensuring the accuracy of detection, which bears practical significance.

8.3.4.4 Artificial Neural Network Promotes Smart Agriculture

With the development of the mobile Internet, IoT, and cloud computing, we are bound to be surrounded by a vast ocean of data. In addition, the level of intelligence of today's artificial neural network technology is getting closer to human intelligence; as a result, the overall structure of modern societies is becoming more and more "smart." In the future, human beings will have a more comprehensive ability to perceive the world. But because of the huge amount of information that is generated every moment, ordinary people and traditional technologies fail to cope with such data and are increasingly enslaved by it. What this means is that further innovations are required in the way information is communicated between humans and nature. Relying on the new generation of information technology, information can be more "smartly" mined and used to provide more "smart" decision-makings and responses. The emergence of "smart agriculture," achieving efficient, high-quality, energy-efficient, environmentally friendly, and sustainable development in agricultural production, will inevitably become the future trend in agriculture.

"Smart agriculture" is a huge system or a top-level virtual concept. The specific implementation of smart agriculture is composed of subsystems in various fields, covering intelligent facility agriculture, intelligent agricultural conditions, intelligent crop protection, intelligent irrigation, intelligent market management, and other subsystems. As smart agriculture deepens, advanced technologies such as IoT, cloud computing, big data, and semantic networks will be increasingly applied. The key to smart agriculture is to fully integrate knowledge bases and model bases in various fields; to adopt inference, analysis, and other mechanisms to make predictions; and to provide intelligent control and decision-making management. At present, smart agriculture is far from real intelligence, but as the Chinese saying goes, the journey of thousands of miles begins with the first step. For example, in facility agriculture, through the automatic network monitoring system, multiple environmental parameters in the greenhouse are collected in real time, including the temperature and humidity of air, soil temperature and humidity, CO2, light, dew temperature, and other environmental parameters. Furthermore, the specific equipment (such as those used for irrigation, shading, ventilation, heating, and cooling) is automatically turned on or off. Additionally, according to user needs, automatic monitoring for the comprehensive information on facility agriculture is offered, and a scientific basis for automatic control and intelligent management is also provided. By monitoring information on crops and their growth environment, various disaster information, and so on, the appropriate time for irrigation and fertilization can be determined, early warning can be issued, and effective disaster prevention and mitigation measures can be adopted. Although these are only the preliminary applications of smart agriculture, they do contribute to the transformation of production methods.

As a primary industry in China, agriculture is facing severe challenges, such as increasing demand for agricultural products, scarce resources, frequent disasters caused by climate change, fragile ecological security, and continued decline in biodiversity. In light of all this, we should consolidate the information-based agriculture with the agricultural IoT and cloud computing technology as its core, thus improving the services of information-based agriculture supported by big data, making new progresses in smart agriculture, and delivering leap-forward development.

8.3.5 Geographic Information System

8.3.5.1 Overview of GIS

Geographic information system (GIS), a specific spatial information system, is a technical system that collects, stores, manages, calculates, analyzes, displays, and describes relevant geographical distribution data in the space of all or part of the Earth's surface (including the atmosphere). Supported by computer hardware and software systems, GIS technology is used in multiple industries to establish spatial databases and decision support systems of various scales, providing users with multiple forms of spatial query, spatial analysis, and auxiliary planning, and decision-making functions. Moreover, GIS performs comprehensive processing, integrated management, and dynamic access to spatio-temporal data from multiple sources. It serves as the fundamental platform for newly integrated systems and provides geological knowledge for intelligent data collection. At the same time, the application of geographic information systems has also transformed the way in which geographic information is released and exchanged, providing a new way to understand geographic information. In this manner, new geographic information can be analyzed and generated more effectively, and geographic information systems can evolve into a discipline that specifically deals with spatial data.

8.3.5.2 GIS Classification

Thematic geographic information system is a geographic information system designed for specific services. This type of systems, such as crop yield information system, comes with limited goals and special features.

Regional geographic information system, a geographic information system that is based on geographical areas, conducts comprehensive research and provides integrated information services. Such systems, like the National Information System of Canada, can be divided by administrative area; they can also be divided by natural area, such as the Yellow River Basin Information System.

The system of general software tools is a set of software packages with the basic functions of geographic information system, including graphic image digitization, storage management, query inspection, analysis calculation, and multiple outputs. The system is mainly used as support software for geographic information system to build thematic or regional geographic information systems.

8.3.5.3 GIS Composition

A functioning GIS should support the collection, management, processing, analysis, modeling, and display of spatial data. GIS is mainly composed of five parts: system hardware, system software, spatial data, personnel, and application model.

- System hardware includes computer host, data input device, data storage device, data output device, and data communication transmission device.
- System software refers to various programs necessary for the operation of a geographic information system, usually including computer system software, geographic information system software, and application analysis programs.
- Spatial data is a term that refers to the object and content within the scope of the geographic information system.
- Personnel refers to developers. Skilled operators are able to overcome the deficiencies of GIS software functions, whereas those who are not fail to do the

same. The superiority of software cannot make up for the damages that ignorant GIS operators may bring.

- Application model refers to the expression of the rules summarized on the basis of extensive research on the specific objects and processes in the professional field.

The GIS software system consists of six modules, namely, data input module, graphics and attribute editing module, data storage and management module, data analysis and processing module, data output and presentation module, and user interface module.

- Data input module is used to transfer raw data (data from multiple sources and multiple forms) to the system and convert these data from the external format to an internal format that would be convenient for system processing. It mainly consists of vector tracking digitization by hand tracking digitization, raster scanning digitization by scanning digitization, and keyboard input.
- After data input is completed, the graphics and attribute editing module is used to layer, classify, and encode data. It is also used for edition, transformation of coordinates, trimming, edge joining, attribute input and attribute connection, etc.
- Data storage and management module mainly organizes and manages the location, spatial topological relationship, and attribute data of geographic elements (usually points, lines, and surfaces). The specific functions of the module include seamless layer connection, partitioning, spatial indexing, and data access mechanism.
- Data analysis and processing module mainly performs spatial analysis operations and preprocessing operations on data. Including query and extraction operations, raster overlay analysis and vector overlay analysis, spatial clustering, path analysis, resource allocation address matching, and other basic spatial operation functions.
- Data output and presentation module is primarily responsible for the thematic mapping output of data, statistical report output, visual representation of scheme selection, etc.
- User interface module mainly provides user interface, program interface, data interface, and development environment.

8.3.5.4 GIS Application in Agriculture

Based on the characteristics and powerful functions of geographic entities of information, GIS is extensively adopted in the management of resources and environment, planning and zoning, disaster monitoring, and environmental protection for agriculture.

Management of Agricultural Resources and Environment
Management of agricultural resources and environment brings together all kinds of resources and uses GIS statistics and coverage analysis functions to provide resource

statistics and the reproduction of original data under diverse conditions according to boundary and attribute conditions. Such management surveys agricultural resources and environmental data through GIS and adopts GIS statistical analysis function to provide the basis for decision-making in resource and environmental management.

Agricultural Planning and Zoning

To achieve agricultural planning, one needs to deal with many problems of different natures and characteristics and conduct comprehensive analysis of multiple factors. The GIS database merges these data into a unified system for professional application models to achieve multi-factor and multi-objective planning, including the suitability evaluation of different types of land and the optimal allocation of resources.

Monitoring of Agricultural Disasters

Comprehensive evaluation of major meteorological disasters, using remote sensing, GIS, and computer technologies, is able to calculate the approximate disaster area based on GIS spatial information and to estimate the potential economic losses. Via the analysis of historical environmental conditions and disasters, one is able to predict the basic pattern of disaster occurrence, spatio-temporal distribution, probability distribution, and hazard levels. Additionally, comprehensive evaluation simulations are made, and the specifics of disasters are predicted, thus facilitating decision-making for prevention and mitigation.

Agricultural Environmental Protection

GIS is used to analyze and process the information obtained by remote sensing. Relying on GIS, the relevant problems are identified in time, and early warning is given accordingly. Moreover, an environmental spatial database is established. The database is then analyzed, and a thematic map of an indicator is drawn, which can express the changes in the environment. The GIS model is used to establish an environmental model for simulation of the dynamic changes of the agricultural environment. The model also provides services for decision-making and management. Any substance in the material world is firmly branded in time and space, and IoT cannot be separated from GIS's location services. In terms of processing spatial information and attributes of matter, GIS is more advantageous. As a spatial geographic information system, GIS is a computer system that can visually support decision-making. It will become the fundamental support system for building IoT and will play an increasingly important role in information society of the future.

8.3.5.5 Development of GIS in Agriculture

Geographic information system is a technology in research on geography that has rapidly matured since the 1960s. It is a new marginal discipline that integrates computer science, geography, mapping and remote sensing, space science, information science, and management science. Furthermore, GIS is also a spatial information system that collects, stores, manages, analyzes, and describes data related to space and geographic distribution based on a geospatial database with the support of computer hardware and software systems. It provides services for geographic research and decision-making. The objects processed and managed by GIS are data on geospatial entities and their relationships, covering spatial positioning data, graphics, remote sensing images, attributes, etc. Such data are used to analyze and process various phenomena and processes distributed in a certain geographical area and to solve the problems of complex planning, decision-making, and management.

From a technical point of view, GIS is currently in a critical period of development. New concepts and new products of GIS are emerging endlessly, and the trend of componentization and network of GIS are becoming increasingly apparent. Component GIS (abbreviated as ComGIS) is based on a standard component platform. In ComGIS, each component can be freely and flexibly reorganized. Additionally, it also has a visual interface and a convenient standard interface. The component platform mainly includes Microsoft's COM (component object model) and DCOM (distributed component object model) and OMG's common object request broker architecture, and Microsoft's COM/DCOM dominates the market status. Based on COM/DCOM, Microsoft introduced ActiveX technology. ActiveX controls are the most widely used standard components in today's visual programming, and most of the new generation of component GIS are ActiveX controls or their predecessors – OLE controls. It can be concluded that component GIS represents the current trend of GIS development.

As the relevant theory improves, GIS is becoming more integrated and intelligent, making it an important tool for decision support. By combining GIS technology with GPS technology, RS technology, expert systems technology, and Internet technology, the functions of the GIS system can be extended, and various geospatial data, attribute data, images, and files on the web can be browsed and obtained from any node on the Internet across regions. Geospatial analysis is then conducted, so that the design of information systems is transformed from the traditional model of self-sufficient data to an open model of online data resources and that timely and reliable information decision support is made available.

8.3.6 Image Processing Technology

Image processing is a method for removing noise. Moreover, it is also dedicated for enhancing, restoring, segmenting, and extracting features from an image using a computer. The adoption of image processing technology in agriculture is fairly recent. Computer multimedia technology, as it matures in recent years, has been extensively adopted in crop growth monitoring; diagnosis of diseases, pests, and weeds; automatic harvest of crops; seed quality detection; identification of deficiencies in crops; and classification of agricultural products. Common computer image processing methods include image enhancement, morphological processing, and image segmentation. The following constitutes a brief introduction of the related image processing technologies.

8.3.6.1 Image Enhancement

Image Histogram Processing

Histograms are the basis of many spatial domain processing technologies. Histogram operations can be used for effective image enhancement. In addition to providing useful image statistics, the information inherent to histograms is also useful in image compression and segmentation. Additionally, histograms are easily calculated in software and are also suitable for commercial hardware devices.

The detection of white foreign fibers is a difficult problem in cotton online detection. Wang et al. (2012) improved the two-dimensional Otsu algorithm by analyzing the gray histograms of white heterosexual fibers and lint cotton. When calculating the probability sum of the target and the background, the probability sum of the sub-diagonal regions of the two-dimensional gray histogram was considered, which reduced the value range of the two-dimensional Otsu algorithm threshold pair. Experiments demonstrated that compared with the one-dimensional Otsu algorithm and the fast-two-dimensional Otsu algorithm, the accuracy and real-time performance of the improved two-dimensional Otsu algorithm are markedly superior, and the algorithm has been successfully applied to actual production. Combining multiple support vector machine classifiers, Zhang et al. (2016) proposed an intelligent rapeseed deficiency analysis and diagnosis method based on non-uniform histogram of HSV color space. In Zhang's research, first, the active contour model was used to segment the rape leaf region, then the HSV color histogram features of the segmented rape leaf region were extracted, and the method of non-uniform quantization was adopted to characterize the color difference of different deficiency rapeseed leaf images. Finally, one-to-many scheme was used to train multiple support vector machine classifiers that are used for the classification and recognition of different lack of rapeseed leaf images. The results demonstrated that this method can identify the types of deficiency in common rapeseed in a more accurate manner, and the overall recognition rate of the five types of deficiency was 93%.

Image Filtering

Image enhancement methods fall into two broad categories: spatial domain methods and frequency domain methods. The former is based on the direct processing of pixels in an image, whereas the latter is based on the Fourier transform of the modified image. Neighborhood processing is adopted to control the image values of the neighborhood and the corresponding sub-images with the same dimensions. These sub-images can be called filters, masks, kernels, templates, or windows. The mechanism of spatial filtering is to move the mask point by point in the image to be processed. At each point, the filter's response is calculated through a predefined relationship. With regard to linear spatial filtering, the response is determined by the product of the filter coefficient and the corresponding pixel value of the area swept by the filter mask.

The output (response) of the smoothed linear spatial filter is a simple average of the pixels contained in the filter mask. The smoothing filter replaces the value of each pixel of the image with the average gray value of the pixels in the neighborhood determined by the filter mask. This process reduces the “sharp” change in image gray. A common smoothing application is noise reduction; however, because the edge of the image is affected by sharp changes in terms of gray level, there is a negative effect of edge blurring in median filtering process, whose major function is to remove the irrelevant details from the image.

A statistical filter is a nonlinear spatial filter whose response is based on the ordering of pixels in the image area surrounded by the image filter, and the value determined by the statistical ordering results replaces the value of the central pixel. A common median filter is counted on to replace the value of the pixel with the median gray level of the pixel. Such filters feature excellent de-noising ability for certain types of random noise and are very effective for processing impulse noise (pepper salt noise).

In order to reduce the noise of the images about crop diseases and pests, Li et al. (2019) proposed an improved adaptive Gaussian filtering algorithm. This method is used to determine the Gaussian standard deviation by calculating the ratio of the center point neighborhood variance in the pixel matrix area of the image to the two-dimensional Gaussian filter function. Additionally, it is also used to dynamically generate a Gaussian convolution kernel and perform noise reduction and smoothing on the lesion image. The improved filtering algorithm provides an optimized means of pre-processing for the image of leaf diseases and pests, thereby improving the accuracy of diagnosis. Yu et al. (2019) proposed a new filtering method to solve the problem that the measured SAR image is widely submerged by noise and that the edge tends to blur when the traditional filtering method is used. In this new method, on the multi-scale wavelet components of the image, the components after de-noising are obtained by setting different thresholds in different coefficients and different directions based on Bayesian theory and the wavelets corresponding to the image edges and other structures extracted based on multi-scale edge detection. Eventually, these reconstructions are fused to filter noise.

8.3.6.2 Image Morphology Processing

Image morphology processing methods mainly include expansion, erosion, open operation, and closed operation. Dilation or erosion is used to convolve the image with the kernel. Dilation is an operation that seeks a local offset, and corrosion, as opposed to swelling, is an operation that seeks to localize. Open operation is performed to etch and then expand. Generally speaking, open operation will smooth the outline of the object and eliminate small burrs, but the local shape of the object remains unchanged. Reversely, closed operation is to first expand and then erode. Similarly, the closed operation will also smooth the contours and eliminate edge burrs, but in contrast to open operation, it usually eliminates isolated spots and fills in cracks and holes in the contour lines.

In order to solve the problem of inaccurate fruit recognition during the operation of tomato picking robots, Sun et al. (2019) proposed a target extraction algorithm based on a combination of geometric morphology and iterative random circles. This algorithm can effectively segment and identify the fruits that are stuck in the image. The accuracy rate of fruit recognition using this algorithm reached 85.1%. To some extent, this algorithm solves the problem of fruit segmentation when multiple fruits are stuck or blocked in a complex environment. Wang et al. (2019) used the method of morphological erosion to optimize the segmentation results when implementing the segmentation of corn stubble rows and removed false segmentation caused by threshold errors, such as burrs, isolated points, and certain inter-row noise. The average error rate of Wang's experiment relative to the target area was 24.68%, which is far lower than the iterative method (90.67%) and the OTSU method (86.42%).

8.3.6.3 Image Segmentation

RGB Segmentation Model

Due to the inherent characteristics of the human eye, the colors we see are various combinations of the so-called primary colors: red (R), green (G), and blue (B). In the RGB color model, the image represented is composed of three image components, and each component is its primary color image.

In agricultural production, image segmentation technology is used to separate soil and vegetation, so that weeds and crops can be further differentiated. In general, for crops whose main pigment is chlorophyll (such as corn, soybeans, wheat, etc.), the green component G of the crop area is much greater than the red component R and the blue component B. Therefore, a 2G-R-B (ExG) filtering algorithm that emphasizes the green component and suppresses the remaining two components can be used to grayscale the color image (García-Santillán et al. 2018; Vidović et al. 2016; Arroyo et al. 2016). The existing navigation line extraction algorithm of the automatic navigation system based on machine vision is easily interfered by the external environment, and the processing speed is slow. To solve these problems, Zhai et al. (2016) proposed a multi-crop line detection algorithm based on binocular vision. Relying on this algorithm, color images of different crops were obtained, and the 2G-R-B algorithm was used to grayscale the color pictures for the acquisition of pictures with clear contrast between the crop rows and the soil background.

Zhang et al. (2014) proposed an online recognition method for apple defects and pedicel calyx based on brightness correction and AdaBoost. During Zhang's research, focused on Fuji Apple, first the RGB image and NIR image of apple were collected online, and the NIR image was segmented to obtain the binary mask of apples. Second, the brightness correction algorithm was used to perform brightness correction on the R component image, and the corrected image was segmented to obtain the defect candidate area (stalks, flowers, defects, etc.). Then using each candidate area as a mask, the information of the 7 pixels inside it were randomly extracted to represent the characteristics of the candidate area, and the seven groups of features were sent to the AdaBoost classifier for classification and voting. The category of the candidate area was determined using the final voting results. In order to study the distribution of cantaloupe surface texture features, Ma et al. (2014) collected a variety of cantaloupe sample images. Ma performed algebraic operations on the R, G, and B components of RGB color images, converted them into grayscale images, and then carried out background segmentation. The image was then decomposed using double tree complex wavelet transform (DT-CWT), high-frequency sub-images were obtained, and neighborhood operations were carried out. Furthermore, using iteration method, the optimal threshold was selected to complete the texture extraction. Finally, the gray difference statistical method and texture spectrum analysis method were used to describe and analyze the cantaloupe texture features, and a classification model based on support vector machine (SVM) is established. The research demonstrated that using the combination of DT-CWT and neighborhood operation, more continuous and complete cantaloupe texture images could be available. Additionally, it is noteworthy that the texture feature values of the four kinds of cantaloupe differ significantly, and the accuracy rate of classification using texture feature values reached 89.3%.

HSI Segmentation Model

Hue is a property that describes solid colors. Saturation gives a measure of the degree to which pure color is diluted by white light; while intensity, the most useful descriptor for monochrome images, can be measured and easily interpreted. The HSI (hue, saturation, intensity) color model conforms to the way people describe and interpret colors. This model can remove the influence of intensity components from the color information (hue and saturation) carried in a color image, making it more suitable for gray processing technology. HSI is an ideal tool for developing image processing methods based on color description.

The complex and changeable color of grape stalks and irregular contours make it difficult for the picking robot to accurately locate the picking points. To solve this problem, Luo et al. (2015) proposed a new method for picking point positioning based on improved clustering image segmentation and the minimum distance constraint of point and line.

Firstly, by analyzing the color space of grape images, the HSI color space component H that can best highlight summer black grapes is extracted, and an improved optimization method with artificial bee colony and fuzzy clustering was used to segment grape fruit images. Morphological de-noising processing is carried out on the segmented image, the largest connected area was extracted, and the centroid, contour extreme points, and circumscribed rectangles of the area were calculated. Then, the region of interest of the picking point was determined based on the centroid coordinates and information on the edge of grape cluster. Furthermore, the cumulative probability Hough straight line detection was performed in the area, and the distances between all the detected straight lines and the centroid were determined. Finally, the straight line with the smallest point-line distance was selected as the line where the picking point was located, and the coordinates of the midpoint of the line segment were used as the picking point.

Texture Segmentation Model

A large number of detailed information is contained in texture, and such information can be used to significantly improve image segmentation. In plateau pika target tracking, the target and background color are similar in natural habitats. To solve this problem, Chen et al. (2015) proposed a plateau pika target tracking method based on local texture difference operator. A new visual descriptor, named the local texture difference operator (LTDC), was constructed to reflect the subtle differences between the target and the background. The method demonstrated its strong capacity to distinguish target from background. Additionally, this descriptor can accurately locate the plateau pika target in scenes with similar target and background colors.

Forest vegetation is an important target in remote sensing image segmentation, and effectively determining the texture scale of forest vegetation is an essential aspect of texture segmentation. Liu and Yang (2015) proposed a method for describing the texture characteristics of forest vegetation in remote sensing images using the blue noise theory. This was a new method for vegetation texture characterization and texture scale calculation. In their research, the correspondence between the scale and the texture of the vegetation was studied. For the selected detection area, the blue noise characteristics were iteratively searched. The iterative process included reducing the size of the area by geometric transformation, using the fast Fourier transform to obtain the spectral response of the area, and extracting the blue noise features from the spectral response. For areas with blue noise characteristics, the gray distribution of the forest vegetation texture was calculated, and the size of the texture was calculated according to the current area size. The experiments they conducted demonstrated that the measurement results of the scale and gray distribution of the texture units of forest vegetation are accurate; thus a reliable basis was provided for further texture segmentation. Aiming at the current problem of multi-feature utilization in remote sensing image segmentation, Wu et al. (2013) proposed a segmentation method that makes comprehensive use of spectrum, texture, and shape information. The method, based on the initial segmentation, counts the spectrum and LBP texture features of the region; then calculates the heterogeneity between adjacent regions based on the spectrum, texture, and shape features; and uses this as a basis to construct a region adjacency map. Based on the adjacency graph, a stepwise iterative optimization algorithm is adopted to perform regional merge, thus acquiring the final segmentation result. It was demonstrated by the segmentation experiments of QuickBird and SAR images that the algorithm can make full use of the spectrum, texture, and shape information of the ground features in the image and that the effect and efficiency of segmentation are satisfactory.

8.3.7 Standardization of Agricultural Data

8.3.7.1 Overview of Agricultural Data Standardization

Information-based development is driving the transformation of agriculture from traditional to modern. Agricultural economy depends not only on traditional agricultural resources but also on the extent to which modern technology is used and to which information is obtained and used. Countries around the world are building various agriculture-related databases, data platforms, and agriculture-related websites based on relevant information to provide timely and comprehensive agricultural information services to producers and researchers. Agricultural information has developed rapidly. The huge amount of data has brought rich and comprehensive information to agricultural activities. At the same time, a number of problems have emerged, including disorder, distorted content, and reduced social benefits. As Nasbitt puts it: “Loss of control and unorganized information no longer constitutes a resource in the information society. Instead, it will become the enemy of information workers” (Zhang et al. 2011). How to reasonably organize, share, and use agricultural information resources has become a major problem. Technical solutions such as automatic crawling and intelligent search are not the best ways to use agricultural information, and standardization of agricultural information is the key to this problem.

8.3.7.2 Definition of Agricultural Data Standardization

Agricultural information is a wide-ranging subject, including multiple fields and levels of production, as well as agricultural economy, markets, policies, and regulations. The standardization of agricultural information refers to the standardized management of all links of agricultural activities and the connection of the different stages of the acquisition, transmission, storage, analysis, and utilization of information. Via such standardization, agricultural information can be reasonably used, and the scope of information sharing can be expanded.

It is generally believed that agricultural information standardization refers to the most common, most regular, and most repetitive things and concepts in the field of agricultural information technology. By formulating, publishing, and implementing standardization, agricultural information technology can achieve a certain unity or consistency within a given range. This unification or consistency is the prerequisite for the implementation and sharing of information resources, which facilitate the promotion and popularization of information technology. Furthermore, it will also be conducive to the development and utilization of agricultural resources and to the formation of information-based industries in agriculture.

8.3.7.3 Contents of Standardized Processing of Agricultural Data

Standardization will make agricultural information simple, practical, open, and easy to maintain, and the information standards will be made science-based and forward-looking. The standardization of agricultural information is manifested in improving the integrity, reliability, timeliness, and applicability of information acquisition and in achieving the accuracy, reliability, versatility, and effectiveness of information processing and utilization. Based on the above principles and objectives of standardization, agricultural information standards should include the following aspects.

Agricultural Data Terminology Standards
Through extensive inquiries of and reference to agricultural standards, the basic concepts of research on and management of agriculture are collected and sorted out, and the words with high frequencies of agricultural use are standardized and unified to ensure consistency of the terms. By revising and formulating new terms, the accuracy of information is guaranteed. Based on the principles of science-based measures, easy usage, and easy maintenance, an agricultural terminology database is established to provide an informative terminology set for producers, managers, and research and technical personnel.

Classification and Coding of Agricultural Information
According to the science-based, forward-looking, open, and easy-to-maintain principles of agricultural information standardization, agricultural information is rationally classified, and the attributes are standardized.

Agricultural Data Technology Standardization
China's agricultural departments at all levels have now developed many sets of expert systems, real-time control systems, geographic information systems, and model simulation systems. These systems are related to each other and are the sources of data for each other. But without a unified technical standard, exchanges between the systems are not possible, and resource sharing cannot be achieved. Thus, standardization is the foundation for the sharing and deep processing of agricultural information.

Agricultural Data Management Standardization
In addition to agricultural information, data management in agriculture must also be standardized, which is one of the prerequisites for ensuring the reliability of agricultural information. Standardization of data management includes standardization of the acquisition, processing, exchange, and release of data.

8.4 Multi-Source Agricultural Information Fusion Technology

Multi-source information fusion is an information processing method for systems using multiple types of information sources (or sensors). The information from multiple sources or multiple sensors is comprehensively processed to obtain more accurate and reliable conclusions. The technology is also known as multi-source correlation, multi-source synthesis, sensor integration, multi-source information fusion, multi-sensor information fusion, etc. This section introduces the concept, basic principles, and applications of multi-sensor data fusion in agriculture.

8.4.1 Concepts and Principles of Multi-Source Information Fusion Technology

Information fusion, a new multidisciplinary discipline, originated from the US Department of Defense's Sonar Signal Processing System Project in 1973. In the 1980s, multi-sensor data fusion technology was developed to meet military needs. In recent years, as agricultural IoT technology matures, multi-sensors have also begun to be used in agriculture to obtain multi-point information on agricultural products before, during, and after production, hence laying the foundation for multi-source agricultural information fusion.

Information fusion technology has grown rapidly, and a complete definition is not yet available. Information fusion technology refers to a process of information processing that uses computer technology to automatically analyze, optimize, and synthesize the information of several sensors obtained in time series under certain criteria, thus completing the required decision-making and estimation tasks.

In comparison with the comprehensive analysis in which the human brain rely on to obtain vision, hearing, taste, smell, and touch using the eyes, ears, mouth, nose, skin, and other organs, the basic principle of multi-source information fusion can be concluded as making full use of the advantages of common or joint processing of multiple sensors. Redundant information is eliminated according to certain criteria, and complementary information is synthesized to generate new effective information, thereby improving the effectiveness of the sensor system. This process is not limited to simple input and output; it is also the result of the interaction among various types of information. Sensors are the hardware foundation of multi-source information fusion, while coordination optimization and comprehensive processing are the core of multi-source information fusion.

Multi-source information fusion mainly consists of multi-sensor signal detection, data signal pre-processing, feature extraction, fusion calculation, and target recognition results. Among them, feature extraction and fusion calculation are the key links. More specifically, feature extraction means to extract target information that are useful for classification and recognition from the original data, whereas fusion calculation means to properly process the extracted target information for the completion of correlation, estimation, and recognition between the feature signal and the target parameters.

8.4.2 Algorithm and Model Fusion Technology

In practical applications, the state of an object is affected by multiple factors. The object is subject to its own properties and is also disturbed by a variety of external factors. Therefore, in actual applications, the state of the object is more complicated in space and time. Most algorithms and models are based on univariate time series, that is, only univariate time series information of a certain type of information source can be used, and existing multi-source information cannot be fully utilized. In order to solve the abovementioned problems, interactive multi-model, multi-algorithm technology, and multi-sensor information fusion technology could be combined to process multi-source information and extract comprehensive information that can best represent the state of the object.

In terms of algorithm fusion, such as multi-spectral and hyper-spectral image fusion algorithms based on three-dimensional distorted discrete wavelet transform, the same size is obtained by resampling, and three-dimensional distorted discrete wavelet transform is performed. Feature selection, a reasonable and effective fusion criterion, is adopted, the source image wavelet coefficients are combined, and finally the 3D inverse discrete wavelet transform is performed on the fused transform domain data to obtain a fused image.

8.4.3 Model Transfer Technology

Model transfer refers to mathematical methods applied to new sample states that empower models established under specific environmental or equipment conditions. Model transfer usually occurs between different instruments, different accessories of the same instrument, or different measurement environments.

With regard to objects with strong uncertainties, the following two main solutions have been proposed in terms of model transfer. The first is to establish parametric modeling of complex systems, and the second is to adopt multiple models to approximate the dynamic performance of the original system relying on the multi-model fusion. The former is primarily achieved through pre-processing, such as variable selection, differentiation, wavelet transform, multivariate scattering correction (MSC), standard orthogonal variable transform (SNV), Fourier transform (FT), and orthogonal signal correction (OSC). Moreover, extended correction models, robust regression methods, RPLS, INV-PLS, and other methods are tapped into for enhancing the selection capacity of the model built by the source machine, so that it can be used for the signal measured by the target machine. Yet it takes time to calculate the dynamic parameters of the model and to test the accuracy and adaptability of the model representation. Thus, there are only a limited number of such applications. In contrast, multi-model fusion, an effective method for solving complex modeling problems, is characterized by the linearized multi-model representation of a nonlinear system, and the uncertain parameter space (or model space) is mapped into a model set. The estimators work in parallel, and the results of each estimator are optimally fused to acquire the final estimate.

8.5 Summary

This chapter is mainly an introduction of information processing technology in agriculture. Starting from its basic concepts, characteristics, and types, the chapter outlines the key technologies in agricultural information processing technology, including data storage technology, data search technology, cloud computing technology, artificial neural network technology, geographic information system, image processing technology, and standardization of IoT and agricultural data. Based on this, the concepts and principles of multi-source agricultural information fusion and processing technology are laid out.

References

Arroyo J, Guijarro M, Pajares G (2016) An instance-based learning approach for thresholding in crop images under different outdoor conditions. Comput Electron Agric 127:669–679

Chen HY, Zhang AH, Hu SY (2015) Object tracking of *Ochotona curzoniae* based on local texture difference operator. Trans CSAM 46(5):20–25

García-Santillán I, Guerrero JM, Montalvo M, Pajares G (2018) Curved and straight crop row detection by accumulation of green pixels from images in maize fields. Precis Agric 19(1):18–41

Li J, Ding XQ, Chen G, Sun Y, Jiang N (2019) Blade image denoising method based on improved Gauss filtering algorithm. J Southern Agric 50(6):1385–1391

Liu LT (2012) Design and implementation of a distributed spatial database engine. Dissertation, Jilin University

Liu XD, Yang S (2015) Forest vegetation texture measurement of remote sensing images based on the blue noise theory. Remote Sensing Land Resour 27(2):63–68

Luo LF, Zou XJ, Xiong JT, Zhang Y, Peng HX, Lin GC (2015) Automatic positioning for picking point of grape picking robot in natural environment. Trans CSAE 31(2):14–21

Ma BX, Gao GG, Wang B, Lü C, Zhang W, Zhu RG (2014) Texture extraction of Hami melon based on dual-tree complex wavelet transform and neighborhood operation. Trans CSAM 45(12):316–322

Nie PC (2012) Research on plant information perception and self-organized agricultural IoT system. Dissertation, Zhejiang University

Sun JT, Sun YF, Zhao R, Ji YH, Zhang M, Li H (2019) Tomato recognition method based on iterative random circle and geometric morphology. Trans CSAM B07:22–26

Vidović I, Cupec R, Željko H (2016) Crop row detection by global energy minimization. Pattern Recogn 55:68–86

Wang HJ (2016) Research on key technologies of eggshell crack detection and on-line detection based on acoustic method. Dissertation, Zhejiang University

Wang HP, Feng XY, Li L (2012) Detection algorithm of white foreign fibers based on improved two-dimensional maximum between-class variance method. Trans CSAE 28(8):214–219

Wang CL, Lu CY, Chen WZ, Li HW, He J, Wang QJ (2019) Image segmentation of maize stubble row based on genetic algorithm and threshold filtering noise. Trans CSAE 35(16):198–205

Wu DB (2015) Contrast research and application of SQL and NoSQL. Dissertation, East China University of Technology

Wu ZC, Hu ZW, Zhang Q, Cui WH (2013) On combining spectral, textural and shape features for remote sensing image segmentation. Acta Geodaetica et Cartographica Sinica 42(1):44–50

Yang F, Wu HR, Zhu HJ, Zhang HH, Sun X (2011) Massive agricultural data resource management platform based on Hadoop. Comput Eng 37(12):242–244

Ye FM, Dong M, Luo W, Xiao H, Zhao XQ, Min WD (2019) Agricultural remote sensing image retrieval based on convolutional neural network and reranking. Trans CSAE 35(15):138–145

Yu XW, Xue DJ, Chen FJ (2019) A wavelet Bayesian SAR image filtering with multiscale edge detection. Remote Sensing Inform 34(5):120–125

Zhai ZQ, Zhu ZX, Du YF, Song ZH, Mao ER (2016) Multi-crop-row detection algorithm based on binocular vision. Biosyst Eng 150:89–103

Zhang JF, Lou CS, Sun SF, Wei QF (2011) Thinking on the standardization of network agricultural information. Chinese Agric Sci Bull 27(1):461–465

Zhang BH, Huang WQ, Li JB, Zhao CJ, Liu CL, Huang DF (2014) On-line identification of defect on apples using lightness correction and AdaBoost methods. Trans CSAM 45(6):221–226

Zhang KB, Zhang AQ, Li CS (2016) Nutrient deficiency diagnosis method for rape leaves using color histogram on HSV space. Trans CSAE 32(19):179–187

Zheng B (2012) Agricultural image processing system design and algorithm research based on cloud computing. Dissertation, Jiangsu University