Research on Germplasm Resources Informationization Based on Mobile Information Technology

Wenhui Pu, Hongli Li*, Yun He, Qingmei Hong, Wenbin Hu

Tropical Crops Genetic Resources Institute of Chinese Academy of Tropical Agriculture Science, Dan Zhou City, Hainan Province, China, 571737

Abstract. The traditional collection of germplasm resources adopts the form of paper collection tables. It takes a lot of manpower and material resources to field operations, and then collects, records, and analyzes the collected data. This operation method is not only time-consuming and labor-intensive, but also cannot perform images. And other important multimedia is data collection. The demand for images by germplasm researchers is not limited to its visibility. They urgently need a real-time, mobile, multi-source acquisition method to achieve a higher level of integration, including basic content, geographic location, and location information. And other multimedia information and other image data information. With the further development of information technology, the Internet has begun to be applied to various industrial fields, and has played a huge role. In order to solve the above-mentioned problems, it is necessary to apply information technology to achieve rapid collection of bamboo germplasm resources. The use of mobile computing and other related information technologies to build a data collection system for bamboo germplasm resources based on mobile computing has not only improved the theoretical application level of mobile computing technology, but also promoted the information collection process of germplasm resources. Therefore, in order to achieve the information collection of germplasm resources, it is necessary to closely integrate this work with mobile computing technology.

1. Introduction

Collection is a very important part of scientific research work. Maintaining collection of plant germplasm data is the most important part of scientific research. The data of plant germplasm resources is maintained on the traditional manual collection method. It is the use of the earliest and most extensive data on the quality of resources in the traditional manual collection method. It uses the earliest and most extensive method of collecting data on quality resources. This to some extent restricts the way the line is acquired. This to some extent restricts the accuracy and completeness of the collection of plant germplasm resources. Although this collection method is far from the informatization of bamboo resource data, it is rooted in China's comprehensiveness [1-3]. Although this collection method is far from the informatization of bamboo resource data, it is rooted in China's comprehensiveness. Although this collection method is far from the informatization of bamboo resource data, but according to the actual situation in China, the main collection technology of plant germplasm resources is still maintained in the artificial situation. The plant germplasm resources data are mainly Acquisition techniques are still maintained in manual mode [5]. In figure 1, we introduce the structure of mobile computing.
However, there are many problems with manual collection. For example, manual collection has many problems such as: (1) Inaccuracy in position, inaccuracy, and inaccuracy in position and accuracy; Position and accuracy are not accurate; (2) Plotting of graphics is not allowed; Plotting is not allowed; Plotting is not allowed; Plotting is not allowed; (3) Filling in paper forms requires a large amount of work; Filling in paper forms requires a lot of work; Filling in paper forms requires a lot of work; (4) Inconvenient for checking; It is very inconvenient to consult. It is very inconvenient to consult. It is very inconvenient to consult. It is very inconvenient to consult. (5) The relationship between paper forms is erroneous; (6) The amount of indoor work data is large, easy to create artificial errors. Large amount of industry data input, easy to create human error; (7) Long collection period, slower data; Long, slower data; (8) Destructiveness of the material growth environment, etc.; (9) The acquisition records are easily lost, the data collection records are easily lost, the data are incomplete; (10) The collection standards may vary from person to person [6-10].

2. Real-time Location Image Acquisition Methods

2.1. Research on mobile positioning method

The location service based on the mobile terminal brings great convenience to people's daily and field collection operations, which also enables the rapid development of mobile location technology. According to different positioning methods and positioning strategies, mobile positioning can be classified into two types based on terminals and networks. The following describes the techniques commonly used in these two positioning methods according to the different positioning methods.

Firstly, let’s talk about GPS positioning technology. GPS is the global positioning system developed by the United States in the second half of the last century. It uses the instantaneous position of the satellite as known data to determine the position of the point to be measured. The system consists of 24 satellites that are evenly distributed over six orbits and exhibit an inclination of 55 degrees, with each orbital plane separated by 60 degrees and about 20,000 kilometers from the ground.

| Table1 positioning technology |
|--------------------------------|
| Technical solutions | Advantage | Disadvantage |
|---------------------|-----------|--------------|
| Terminal-based mobile positioning technology |
| GPS                 | High precision and wide range | Easily affected |
| A-GPS               | Short positioning time      | Expensive      |
| BDS                 | Open to the outside world  | Too expensive  |
Network-based mobile positioning technology

|            | Easy operating and inexpensive | Low precision |
|------------|-------------------------------|--------------|
| COO        |                               |              |
| Wi-Fi      | Short responding time         | weak signal  |

The second method is A-GPS positioning technology. That is auxiliary positioning technology. That is, the assisted GPS positioning technology is to overcome the problem that the first positioning time of the GPS is longer than the first positioning time and develops the technology. A reference network of a GPS reference network is established by using a communication network of a mobile terminal such as a mobile phone, and the GPS position information is used. Etc. passed to the reference network for calculation by its server.

The third method is COO positioning technology. COO (Cell of Origin) originates from cells. The COO positioning technology is located through the CELL-ID where the mobile terminal is located. As long as the angle, time and time difference of the signal to the CELL-ID are measured, the approximate location can be calculated.

The last method is Wi-Fi positioning technology. Based on Wi-Fi (Wireless Fidelity) high-fidelity wireless LAN positioning technology is also based on WLAN positioning technology, which is currently using RSSI technology. The positioning principle is that the RSSI information can be obtained only by the mobile phone acquiring the signals sent by each AP (Access Point), and then combined with each AP's address to perform calculation and positioning through different algorithms. See Table 1.

2.2. Establishment of Real-time Location Image Model
Image is the carrier used to describe things in people's social life production. Most of them are the carriers used to describe things in people's social life production. Most of them are obtained by sight, so they rely on vision. To obtain, therefore, images have become the main source of information for people, and irreplaceable possession in life products has become a major source of information for people, and they occupy an irreplaceable position in life products. Digital image is also called digital image. Its expression and tradition, its expression and traditional image are different, and the use of two-dimensional numbers is different. It uses two-dimensional numbers or arrays and matrix to express the content of the image. EXIF (Exchangeable Image File) is a format of digital image files, "exchangeable image files." Its standard was developed by Japan's Electronic Industry Development Association, which has a relatively developed image technology.

2.3. Real-time positioning image acquisition
Image location has always been a hot topic in the field collecting industry. With the deepening of research and the maturation of technology, it is not difficult to solve the positioning problem. The difficulty lies in how to increase the positioning speed to achieve real-time positioning. This section is for real-time positioning. The positioning algorithm is studied. When positioning, there are different divisions according to different classification criteria. For example, GPS positioning can be divided into pseudo-range positioning and carrier-phase positioning according to observations; according to positioning patterns, it can be divided into single-point positioning and differential positioning; according to time, it can be divided into Real-time positioning and non-real-time positioning; according to the motion status can be divided into dynamic positioning and static positioning.

(1) The least square method. One of the most commonly used localization algorithms is the Least Square (LS) method, which is also called the least square method. It is itself an optimization algorithm. The reason why the algorithm is simple and the amount of calculation is small is the reason why it uses a wide range. The formula is as follows:

\[ \sum_{j=1}^{n} X_j \beta_j = y_i, (i = 1, 2, 3, \cdots, m) \] (1)

In which,
It is clear that the above procedure is unsolvable, so we also need to introduce the sum of squared residuals.

\[ S(\beta) = \|X\beta - y\|^2 \]  

(5)

When \(\hat{\beta} = \beta\), \(S(\beta)\) has the minimum value:

\[ \hat{\beta} = \text{arg min} \left( S(\beta) \right) \]  

(6)

And then take derivative for \(S(\beta)\), we have:

\[ X^T X \hat{\beta} = X^T y \]  

(7)

If \(X^T X\) is nonsingular, then:

\[ \hat{\beta} = (X^T X)^{-1} X^T y \]  

(8)

Although the least square method is widely used, it also has its own drawbacks, such as the fact that the algorithm cannot remove noise well, so that the accuracy is not high, which is relatively unfavorable for achieving fast, real-time positioning.

(2) Unscented Kalman Filter method. The Unscented Kalman Filter (referred to as Unscented Kalman Filter) is also referred to as UKF, also known as de-fragrant Kalman filtering or no-track Kalman filtering. The principle is as follows:

\[
\begin{align*}
X_{0,k-1} &= \hat{X}_{k-1} \\
X_{i,k-1} &= \hat{X}_{k-1} + \left( \sqrt{(N+k)P_{i,k-1}} \right)_i \\
X_{i+N,k-1} &= \hat{X}_{k-1} - \left( \sqrt{(N+k)P_{i,k-1}} \right)_i
\end{align*}
\]  

(9)

First-order weight coefficients are:

\[ W_i = \begin{cases} 
\lambda / (N + \lambda), & i = 0 \\
1/2 (N + \lambda), & i \neq 0
\end{cases} \]  

(10)

Second-order weight coefficients are:
The non-destructive Kalman filter method and the extended Kalman filter method all belong to the Kalman filter algorithm derivation algorithm, which is an improvement on the Kalman filter method, carried out on the basis of the UT (Unscented Transform) transform, using the sampling approach to the nonlinear close. The non-destructive Kalman filtering algorithm is a much optimized algorithm. Compared to other improved Kalman filtering methods, the non-destructive Kalman filtering method does not require the calculation of the sub-comparison matrix and the Hessian matrix, and the calculation amount is less, and because it is in the UT transformation. On top of this makes the accuracy higher and even up to the third-order effect. However, the non-destructive Kalman filtering algorithm also has its disadvantages. For example, when positioning, there are always some values that differ greatly from other values. Obviously, they are not reliable data. These values are extremely extreme and are called abnormal observation values. Therefore, eliminating the non-destructive Kalman filtering algorithm's outliers and improving the algorithm's “anti-wildness” ability have become an urgent problem to be improved.

(3) An improved algorithm LSUTF. The least square method is used to eliminate the outliers in the non-destructive Kalman filter, resulting in an improved algorithm LSUTF. According to the principle of non-destructive Kalman filtering method, a nonlinear to linear close-up transformation is performed as soon as an algorithm calculation is performed, and an observation and update correction is performed. However, if a problem occurs in the sensor, the observation data is used to update and correct. It will only make the data error bigger and more anomalous.

\[
W_i^c = \begin{cases} 
\lambda / (N + \lambda) + 1 + \beta - \alpha^2, & i = 0 \\
1/2(N + \lambda), & i \neq 0 
\end{cases}
\]

\[
\lambda = \alpha^2(k + N) - N
\]

The non-destructive Kalman filter method and the extended Kalman filter method all belong to the Kalman filter algorithm derivation algorithm, which is an improvement on the Kalman filter method, carried out on the basis of the UT (Unscented Transform) transform, using the sampling approach to the nonlinear close. The non-destructive Kalman filtering algorithm is a much optimized algorithm. Compared to other improved Kalman filtering methods, the non-destructive Kalman filtering method does not require the calculation of the sub-comparison matrix and the Hessian matrix, and the calculation amount is less, and because it is in the UT transformation. On top of this makes the accuracy higher and even up to the third-order effect. However, the non-destructive Kalman filtering algorithm also has its disadvantages. For example, when positioning, there are always some values that differ greatly from other values. Obviously, they are not reliable data. These values are extremely extreme and are called abnormal observation values. Therefore, eliminating the non-destructive Kalman filtering algorithm’s outliers and improving the algorithm’s “anti-wildness” ability have become an urgent problem to be improved.

(3) An improved algorithm LSUTF. The least square method is used to eliminate the outliers in the non-destructive Kalman filter, resulting in an improved algorithm LSUTF. According to the principle of non-destructive Kalman filtering method, a nonlinear to linear close-up transformation is performed as soon as an algorithm calculation is performed, and an observation and update correction is performed. However, if a problem occurs in the sensor, the observation data is used to update and correct. It will only make the data error bigger and more anomalous.

\[
x_k^\wedge = \begin{cases} 
x_{k|k-1}^\wedge + L_k(z_k - z_{k|k-1}), & \text{if no outliers} \\
p x_{k|k-1}^\wedge + q, & \text{if has outliers}
\end{cases}
\]

\[
x_m = \frac{1}{n} \sum_{i=1}^{n} x_{(k-i)|(k-i-m)}
\]

\[
y_m = \frac{1}{n} \sum_{i=1}^{n} x_{k-i}
\]

3. Cache Data Synchronization Method Based on Mobile Environment

3.1. Cache in a mobile environment

From life to production, to scientific research, with the development of network technology and communication technology, people do not want to be fixed only at a certain place or time to access data information, but they are eager to operate at anytime and anywhere. The data transmission of the germ-palms resource mobile acquisition system is based on such a mobile computing environment. The distributed computing environment corresponds to the mobile computing environment. The so-called distributed computing environment refers to the host, network, and terminal that have continuous connectivity. It is a fixed computing environment. Relative to the distributed computing environment, the mobile computing environment is more flexible and mobile, which facilitates the access and transmission of information. The mobile computing environment is obviously unable to adapt to distribute computing. It is imperative to optimize and improve the traditional distributed computing systems and technologies to make them suitable for mobile computing environments.
3.2. Data Synchronization Method Optimization Based on a Cache Replacement Algorithm

In order to verify the feasibility of the replacement algorithm, the performance was analyzed and an analog simulation experiment was designed here. The experimental system consists of a cloud server, a client, and a network link. The server is used to maintain the database. The client is used to place the semantic cache and all queries are issued by the client. If a request is sent, it will be processed locally first. If the client cache can be satisfied, it will respond directly from the cache; if it cannot respond directly from the cache, it will be delivered to the server for processing. At this time, the return value will be received through the wireless network. Store this value in the local cache.

![Figure 2: The corresponding time for query set](image)

According to the main parameters in the simulation system, the client's resources include the client processor and the disk. The experiment mainly tests the average corresponding time and network overhead of the query. The specific results are shown in Figure 2 and Figure 3.

From the above results, it can be seen that both the average response time and the network overhead, the CRF-FAR algorithm has better performance than the FAR.

![Figure 3: The overhead for query set](image)

In the mobile network environment, under the condition of poor network connection, the cache technology can effectively improve the system performance. A good cache replacement strategy can improve the cache hit rate. However, simply caching is not enough. When the network is disconnected, if the user is allowed to operate on the client cache, it will cause differences between the cached data and the server database data. In order to change this difference, two If the data is consistent, it is necessary to synchronize the data when the network connection is normal to regain consistency.

4. Conclusion

In this paper, we focus on the cache settings in the mobile environment based on research needs, discusses the differences between the traditional cache replacement strategy and the semantic cache replacement strategy, and proposes a new CRF-FAR cache replacement for the lack of semantic cache
replacement strategy. Based on the strategy, a cache data synchronization method model is given. Finally, the problem of synchronization of cached data in the case of poor network stability in the mobile environment during the information collection process of bamboo germplasm resources was solved. However, the basic data acquisition module functionality can be further enhanced. The current basic data fields of the system are limited and are limited to some common fields. If you want to collect richer content, you will have some embarrassment. Therefore, the next phase can significantly increase the number of acquired fields, and it can also increase the data dictionary's functionality.

References
[1] Wei, Dong, Xiao-jiang CHEN, and Ding-yi FANG. "Research on the software development based on SOA [J]." Microelectronics & Computer 6 (2005): 020.
[2] Xu, Jing, and Wei XU. "Summarization of message-oriented middleware [J]." Computer engineering 16 (2005): 029.
[3] Youbing, Zhang Rongbiao Gu Guodong Feng, and Lian Chengfei. "Realization of Communication in Wireless Monitoring System in Greenhouse Based on IEEE802. 15. 4 [J]." Transactions of the Chinese Society for Agricultural Machinery 8 (2008): 029.
[4] Shan-wen, L. I. U. "Establishment of a Web-based crop germplasm resource management system [J]." Fujian Journal of Agricultural Sciences 4 (2009): 019.
[5] Jinxin, BIAN Fuling1 WANG. "Engineering and Technical System of Digital City [J]." Editorial Board of Geomatics and Information Science of Wuhan University 12 (2004): 002.
[6] Han, Yan-yun, et al. "A crop trait information acquisition system with multitag-based identification technologies for breeding precision management." Computers and Electronics in Agriculture 135 (2017): 71-80.
[7] Zheng, Xiaofei, et al. "Study on image retrieval based on image texture and color statistical projection." Neurocomputing 215 (2016): 217-224.
[8] Zhen, Yang, et al. "Research and design on conveying system for tractive orchard picking platform." Journal of Chinese Agricultural Mechanization 7 (2017): 005.
[9] Li, XingWu, et al. "Application research of GM two-dimensional bar code in germplasm database information management." Journal of Southern Agriculture 45,11 (2014): 1927-1931.
[10] Ye, Sijing, et al. "Development of a highly flexible mobile GIS-based system for collecting arable land quality data." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7.11 (2014): 4432-4441.
[11] WANG, Hongyun, et al. "Design and implementation of tobacco leaf production management system in dali using google maps." Chinese Tobacco Science 1 (2012): 018.
[12] Chuyou, L. I. N., Z. H. A. N. G. Xiaoji, and X. I. Yanyong. "Application of SMS in Agricultural Information Service [J]." Agriculture Network Information 6 (2011): 042.
[13] Song, Qijian, et al. "Fingerprinting soybean germplasm and its utility in genomic research." G3: Genes, Genomes, Genetics 5.10 (2015): 1999-2006.
[14] Sheng, WenTao, et al. "Research progress on utilization of wild rice germplasm in rice high-yield breeding." Journal of Southern Agriculture 48.2 (2017): 222-230.
[15] Casler, Michael D., Kenneth P. Vogel, and Melanie Harrison. "Switchgrass germplasm resources." Crop science 55.6 (2015): 2463-2478.
[16] Horler, R. S. P., et al. "SeedStor: a germplasm information management system and public database." Plant and Cell Physiology 59.1 (2017): e5-e5.
[17] Gao, Yuan, et al. "Genetic diversity of Malus cultivars and wild relatives in the Chinese National Repository of Apple Germplasm Resources." Tree genetics & genomes 11.5 (2015): 106.
[18] Davies, L. R., and C. J. Allender. "Who is sowing our seeds? A systematic review of the use of plant genetic resources in research." Genetic Resources and Crop Evolution 64.8 (2017): 1999-2008.