Agricultural remote sensing big data: Management and applications

Yanbo Huang¹, CHEN Zhong-xin², YU Tao³, HUANG Xiang-zhi³, GU Xing-fa³

¹ Crop Production Systems Research Unit, Agricultural Research Service, United States Department of Agriculture, MS 38776, USA
² Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R.China
³ Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, P.R.China

Abstract
Big data with its vast volume and complexity is increasingly concerned, developed and used for all professions and trades. Remote sensing, as one of the sources for big data, is generating earth-observation data and analysis results daily from the platforms of satellites, manned/unmanned aircrafts, and ground-based structures. Agricultural remote sensing is one of the backbone technologies for precision agriculture, which considers within-field variability for site-specific management instead of uniform management as in traditional agriculture. The key of agricultural remote sensing is, with global positioning data and geographic information, to produce spatially-varied data for subsequent precision agricultural operations. Agricultural remote sensing data, as general remote sensing data, have all characteristics of big data. The acquisition, processing, storage, analysis and visualization of agricultural remote sensing big data are critical to the success of precision agriculture. This paper overviews available remote sensing data resources, recent development of technologies for remote sensing big data management, and remote sensing data processing and management for precision agriculture. A five-layer-fifteen-level (FLFL) satellite remote sensing data management structure is described and adapted to create a more appropriate four-layer-twelve-level (FLTL) remote sensing data management structure for management and applications of agricultural remote sensing big data for precision agriculture where the sensors are typically on high-resolution satellites, manned aircrafts, unmanned aerial vehicles and ground-based structures. The FLTL structure is the management and application framework of agricultural remote sensing big data for precision agriculture and local farm studies, which outlooks the future coordination of remote sensing big data management and applications at regional and farm scale.

Keywords: big data, remote sensing, agricultural information, precision agriculture

1. Introduction

The development of earth observation technology, especially satellite remote sensing, has made massive remotely-sensed data available for research and various applications (Liu 2015; Chi et al. 2016). There are more than a thousand active satellites orbiting earth nowadays and quite a number of the satellites are for remote sensing. Satellites are equipped with one or more sensors or instruments
depending on the purpose. Various sensors are onboard to collect various observation data from earth surface, including land, water and atmosphere. These satellites with their sensors usually acquire images of earth surface unceasingly at different spatial and temporal resolutions. In this way, huge volume of remotely sensed images are available in many countries and international agencies and the volume grows every day, every hour and even every second (Rosenqvist et al. 2003; Anonymous 2015).

Precision agriculture has revolutionized agricultural operations since the 1980s established on the basis of agricultural mechanization through the integration of global positioning system (GPS), geographic information system (GIS) and remote sensing technologies (Zhang et al. 2002). Over the past thirty years, precision agriculture has evolved from strategic monitoring using satellite imagery for regional decision making to tactical monitoring and control prescribed by the information from low-altitude remotely sensed data for field-scale site-specific treatment. Now data science and big data technology are gradually merged into precision agricultural schemes so that the data can be analyzed rapidly in time for decision making (Bendre et al. 2015; Wolfert et al. 2017) although research remains for how to manipulate big data and convert the big data to “small” data for specific issues or fields for accurate precision agricultural operation (Sabarina and Priya 2015). Agricultural remote sensing is a key technology that, with global positioning data, produces spatially-varied data and information for agricultural planning and prescription for precision agricultural operations with GIS (Yao and Huang 2013). Agricultural remote sensing data appear in different forms, and are acquired from different sensors and at different intervals and scales. Agricultural remote sensing data all have characteristics of big data. The acquisition, processing, storage, analysis and visualization of agricultural remote sensing big data are critical to the success of precision agriculture. With the most recent and coming advances of information and electronics technologies and remote sensing big data support, precision agriculture will be developed into smart, intelligent agriculture (Wolfert et al. 2017).

The objective of this paper is to overview the theory and practice of agricultural remote sensing big data management for data processing and applications. From this study, a new scheme for management and applications of agricultural remote sensing big data is formulated for precision agriculture.

2. Agricultural remote sensing big data

Remote sensing technology has been developed today for earth observation from different sensors and platforms. Sensors are mainly for imaging and non-imaging broad-band multispectral or narrow-band hyperspectral data acquisition. Platforms are space-borne for satellite-based sensors, airborne for sensors on manned and unmanned airplanes, and ground-based for field on-the-go and laboratory sensors. Objects on the earth continuously transmit, reflect and absorb electromagnetic waves. In principle, remote sensing technology differentiates the objects through determining the difference of the transmitted, reflected and absorbed electromagnetic waves. Remote sensing typically works on the bands of visible (0.4–0.7 mm), infrared (0.7–15 mm), and microwave (0.75–100 cm) in the electromagnetic spectrum. All the factors with geospatial distribution and data acquisition frequency result in remote sensing big data with huge volume and high complexity.

Remote sensing technology has been developing with new, high-performance sensors with higher spatial, spectral and temporal resolutions. Agricultural remote sensing is a highly specialized field to generate images and spectral data in huge volume and extreme complexity to drive decisions for agricultural development. In the agricultural area, remote sensing is conducted for monitoring soil properties and crop stress for decision support in fertilization, irrigation and pest management for crop production. Typical agricultural remote sensing systems include visible-NIR (near infrared) (0.4–1.5 mm) sensors for plant vegetation studies, SWIR (short wavelength infrared) (1.5–3 mm) sensors for plant moisture studies, TIR (thermal infrared) (3–15 mm) sensors for crop field surface or crop canopy temperature studies, and microwave sensors for soil moisture studies (Moran et al. 1997; Bastiaanssen et al. 2000; Pinter et al. 2003; Mulla 2013). LiDAR (Light Detection and Ranging) and SAR (Synthetic Aperture Radar) have been enabled to measure vegetation structure over agricultural lands (Zhang and Kovacs 2012; Mulla 2013). Remote sensing is the cornerstone of modern precision agriculture to realize site-specific crop field management to account for within-field variability of soil, plant stress and effect of treatments.

With the rapid development of remote sensing technology, especially the use of new sensors with higher resolutions, the volume of remote sensing data will dramatically increase with a much higher complexity. Now, a major concern is determining how to effectively extract useful information from such big data for users to enhance analysis, answer questions, and solve problems. Remote sensing data are a form of big data (Ma et al. 2015). Storage, rapid processing, information extraction, information fusion, and applications of massive remote sensing data have become research hotspots at present (Rathore et al. 2015; Jagannathan 2016).

Agricultural remote sensing big data have the same features as all remote sensing big data. The specialty of
agricultural remote sensing is that it not only helps manage planning and management of agricultural production strategically at regional, national and even global scales, but also provides control information tactically for precision agricultural operations at the scale of farm fields. Therefore, agricultural remote sensing may produce higher spatial and temporal resolution data. In recent years, unmanned aerial vehicle (UAV) has become a unique platform for agricultural remote sensing to provide coverage of crop fields with multiple images from very low altitude (Huang et al. 2013b). The images can, in turn, be converted into not only a two dimensional visualization of the field but also a three-dimensional surface representation of the field (Huang and Reddy 2015). So, UAV-based agricultural remote sensors are contributing significantly to agricultural remote sensing big data. UAV-based remote sensing is a special kind of airborne remote sensing with possible monitoring of crop fields at ultra-low altitude. How to rapidly and effectively process and apply the data acquired from UAV agricultural remote sensing platforms is being studied widely at present (Huang et al. 2013b; Suomalainen et al. 2014; Cardiga et al. 2015).

3. Management of agricultural remote sensing big data

3.1. Remote sensing data processing and products

Remote sensing data have to be processed before it can be used. In general, it is not suggested to use raw images directly acquired from remote sensors on satellites and aircraft because the data have to be corrected due to deformations from interactions between sensors, atmospheric conditions and terrain profiles. The corrections typically include radiometric and geometric corrections. A complete radiometric correction is related to the sensitivity of the remote sensor, topography and sun angle, and atmospheric scattering and absorption. The atmospheric correction is difficult for agronomists and agro-technicians, in general, because it requires the data and information of atmospheric conditions during image acquisition. However, the data and information typically vary with time and location. The geometric correction is aimed at correcting squeezing, twisting, stretching and shifting of remotely sensed image pixels relative to the actual position on the ground, which are caused by remote sensing platform’s angle, altitude and speed, sensitivity of the remote sensor, and earth surface topography and sun angle, atmospheric scattering and absorption, and rotation of the earth. The raw and corrected remote sensing images can be summarized into data products at different levels as explained in general in Table 1 (Di and Kobler 2000; Piwowar 2001). The programs of MODIS (moderate resolution imaging spectroradiometer) (National Aeronautics and Space Administration (NASA), Washington, D.C.), Landsat (NASA and United States Geological Survey (USGS), Reston, VA), European satellites such as SPOT (SPOT Image, Toulouse, France), and Chinese satellites such as Ziyuan (China Centre for Resources Satellite Data and Application, Beijing, China) all provide products at different levels more or less depending on different applications. Table 2 shows the remote sensing data characteristics of the main medium- and high-resolution land satellites of different countries in the world. Besides there are ocean observation satellite, such as OrbView of United States, and meteorological satellites, such as AVHRR (Advanced Very High Resolution Radiometer), NOAA (National Oceanic and Atmospheric Administration, Washington, D.C.) of United States and FY-3A/B of China. Thus, as illustrated, one remote sensing image can have many products at different levels and the same image product could be resampled up or down scale to meet the practical requirements and transformed for different applications so that the volume and complexity of remote sensing data are rapidly increased as big data.

**Table 1** Product levels of satellite remote sensing data

<table>
<thead>
<tr>
<th>Level</th>
<th>Product description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Reconstructed, unprocessed instrument and payload data at full resolution, with any and all communication artifacts (e.g., synchronization frames, communications headers, duplicate data) removed</td>
</tr>
<tr>
<td>1A</td>
<td>Reconstructed, unprocessed instrument data at full resolution, time-referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters (e.g., platform ephemeris) computed and appended but not applied to the Level 0 data (or if applied, in a manner that Level 0 is fully recoverable from Level 1A data)</td>
</tr>
<tr>
<td>1B</td>
<td>Level 1A data that have been processed to sensor units (e.g., radar backscatter cross section, brightness temperature, optical, etc.); not all instruments have Level 1B data; Level 0 data are not recoverable from Level 1B data</td>
</tr>
<tr>
<td>2</td>
<td>Derived agro-geophysical variables (e.g., ocean wave height, ice concentration, soil moisture/temperature, canopy temperature, etc.) at the same resolution and location as Level 1A source data</td>
</tr>
<tr>
<td>3</td>
<td>Variables mapped on uniform spatial grid scales, usually with some completeness and consistency (e.g., missing points interpolated, complete regions mosaicked together from multiple orbits, etc.)</td>
</tr>
<tr>
<td>4</td>
<td>Model output or results from analyses of lower level data (i.e., variables that were not measured by the instruments but instead are derived from these measurements)</td>
</tr>
<tr>
<td>Satellite</td>
<td>Payload name</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>EOS-TERRA/</td>
<td>MODIS</td>
</tr>
<tr>
<td>AQUA</td>
<td>Enhanced Thematic Mapper Plus (ETM+)</td>
</tr>
<tr>
<td>Landsat 8</td>
<td>OLI (15–30 m)</td>
</tr>
<tr>
<td>Quickbird</td>
<td>Panchromatic 0.65 m Multispectral 2.62 m</td>
</tr>
<tr>
<td>GeoEye-1</td>
<td>Panchromatic 0.46 m Multispectral 1.84 m</td>
</tr>
<tr>
<td>IKONOS</td>
<td>Panchromatic 1 m Multispectral 4 m</td>
</tr>
<tr>
<td>WorldView-1</td>
<td>Panchromatic 0.5 m Multispectral 1.8 m</td>
</tr>
<tr>
<td>WorldView-2</td>
<td>Panchromatic 0.5 m Multispectral 1.8 m</td>
</tr>
<tr>
<td>WorldView-3</td>
<td>Panchromatic 0.31 m Multispectral 1.24 m</td>
</tr>
<tr>
<td>WorldView-4</td>
<td>Panchromatic 0.31 m Multispectral 1.24 m</td>
</tr>
<tr>
<td>Pleiades-1/2</td>
<td>Panchromatic 0.5 m Multispectral 2 m</td>
</tr>
<tr>
<td>Spot-5</td>
<td>HRS 2.5–20 m</td>
</tr>
<tr>
<td>Spot-6</td>
<td>Panchromatic 1.5 m Multispectral 6 m</td>
</tr>
<tr>
<td>RapidEye</td>
<td>Multispectral 5 m</td>
</tr>
<tr>
<td>CBERS-02B</td>
<td>Multispectral 19.5 m WFI Multispectral 258 m</td>
</tr>
<tr>
<td>ZY-1 02C</td>
<td>Panchromatic 5 m Multispectral 10 m</td>
</tr>
<tr>
<td>ZY-3</td>
<td>Panchromatic 2.1 m Multispectral 5.8 m</td>
</tr>
<tr>
<td>HJ-1A</td>
<td>Multispectral 30 m Hyperspectral 100 m</td>
</tr>
<tr>
<td>HJ-1B</td>
<td>Multispectral 30 m</td>
</tr>
<tr>
<td>JL-1</td>
<td>Panchromatic 0.72 m Multispectral 2.88 m</td>
</tr>
<tr>
<td>GF1</td>
<td>Panchromatic 1 m Multispectral 4 m and 16 m</td>
</tr>
<tr>
<td>GF2</td>
<td>Panchromatic 0.8 m Multispectral 3.2 m</td>
</tr>
<tr>
<td>GF3</td>
<td>SAR-C 1-500 m</td>
</tr>
<tr>
<td>Sentinel-1</td>
<td>SAR-C4-80 m</td>
</tr>
<tr>
<td>Sentinel-2</td>
<td>Multispectral 10 m/20 m/60 m</td>
</tr>
<tr>
<td>Sentinel-3</td>
<td>OLCI 300 m</td>
</tr>
<tr>
<td></td>
<td>SLSTR 500 m/1 km</td>
</tr>
<tr>
<td></td>
<td>SYN 300 m/1 km</td>
</tr>
<tr>
<td></td>
<td>SRAL</td>
</tr>
<tr>
<td>RadarSAT-2</td>
<td>SAR-C 1–100 m</td>
</tr>
</tbody>
</table>
Remote sensing began through aerial photography in nineteenth century. Satellite remote sensing has been developed and dominating since the 1960s. In the last decade, airborne remote sensing, especially UAV-based remote sensing, has been significantly developed and applied for monitoring natural resources and managing agricultural lands. With such development of remote sensing systems and methods, the new remote sensing data product system can be formulated in more detail to be referred for airborne remote sensing data processing flow.

Image fusion is often needed as well to fuse the images in multiple sources, multiple scales and multiple phases but in the same coverage to enhance the later image analysis. For example, fusion of a high-resolution panchromatic image with a low-resolution multispectral image can produce a high-resolution multispectral image. Data fusion is at data level, feature level and decision level (Hall and Linas 2009). Remote sensing image data fusion is mostly at pixel level. Feature level and decision level fusions can be used in image classification and model-based parameter inverse. In addition, image fusion can be used in the process of image mosaic to remove exposure difference from image pieces and mosaic artifacts.

UAV remote sensing systems are often operated at very low altitude, especially in precision agriculture (Huang et al. 2013b), for very high-resolution images (1 mm–5 cm/pixel). The images from UAV systems, in one hand, could be less dependent on weather condition, which could simplify or even omit atmospheric correction. On the other hand, each image would cover much less area compared to the images from satellites and aircrafts. To cover a certain area, mosaicking of multiple UAV remote sensing images is required in the processing of the images. The coverage with multiple images, in turn, provides an opportunity to build three dimensional models of the ground surface with structure from motion (SfM) point clouds through stereo vision (Rosniall and Honkavaara 2012; Turner et al. 2012; Mathews and Jensen 2013) for extraction of surface features such as plant height and biomass volume (Bendig et al. 2014; Huang et al. 2017).

3.2. Remote sensing data analysis

Remote sensing data can be analyzed qualitatively and quantitatively. Qualitative remote sensing data analysis is classification-based. Remote sensing classification is critical for remote sensing data being converted to useful information for practical applications. Conventional remote sensing image classification is based on image pixel classification in unsupervised and supervised modes. The commonly used methods includes ISODATA self-organizing, maximum likelihood, and machine learning algorithms such as artificial neural networks and support vector machine (Huang 2009; Huang et al. 2010a). With the development of remote sensing technology for high-resolution data, the conventional pixel-based classification methods cannot meet practical requirements because the high-resolution images with more details may be classified into some unknown “blank” spots, which may have negative impact on later analysis. Object-based remote sensing image classification (Walter 2004; Blaschke 2010) provides an innovative idea to perform image segmentation to merge the neighboring pixels with similar spectral signatures into objects as “pixels” to classify. Quantitative remote sensing data analysis is model-based. Features, such as vegetation indices, extracted from remote sensing data, have been modeled empirically with biophysical and biochemical measurements, such as plant height, shoot dry weight and chlorophyll content, and with the calibrated models the biophysical and biochemical parameters can be predicted for estimation of biomass amount and crop yield. Radiative transfer is the fundamental theory for development of remote sensing data analysis (Gong 2009). Corresponding physically-based model simulation and parameter inverse have been the research focus to understand the mechanism of interaction between remote sensing and ground surface features. In the last few decades, the PROSPECT leaf optical properties model (Jacquemoud and Baret 1990) and the SAIL canopy bidirectional reflectance model (Verhoef 1984) have been representative in radiative transfer studies of remote sensing plant characterization. The two models have been evolved, expanded and even integrated (Jacquemoud et al. 2005, 2009) to lay the foundation leading to more advanced studies in this aspect (Zhao et al. 2010, 2014). Furthermore, remote sensing data assimilation with process-based models such as crop growth models, soil water models, is an emerging technology to estimate agricultural parameters which are very difficult to inverse only from a model or remote sensing data (Huang et al. 2015a, b, c).

Recent years deep learning has been developed from machine learning for remote sensing image classification (Mohanty et al. 2016; Sladojevic et al. 2016). It was strongly believed that deep learning techniques are crucial and important in remote sensing data analysis, particularly for the age of remote sensing big data (Zhang et al. 2016). With the low-level spectral and textural features in the bottom level, the deep feature representation in the top level of the deep artificial neural network can be directly fed into a subsequent classifier for pixel-based classification. This hierarchy handles deep feature extraction with remote sensing big data, which can be used all parts in remote sensing data processing and analysis.
3.3. Remote sensing data visualization

Visualization of remote sensing data and products are critical for users to interpret and analyze. GIS as a platform of remote sensing data visualization has been developing in the last decade in four aspects (Song 2008):

- Modularization
  Modular GIS is organized of components with certain standards and protocols.
- Web enabling
  Web GIS (Fu and Sun 2010) has been developed to publish geospatial data for users to view, query and analyze through Internet.
- Miniaturization and mobility
  Although desktop GIS applications still dominate, mobile GIS clients have been adopted with personal digital assistant (PDA), tablets and smart phones.
- Data-based
  GIS spatial data management has been developed from flat file management, file/database management, to spatial database management. Spatial data management provides the capabilities of massive data management, multi-user co-current operation, data visit permission management, and co-current visit and systematic applicability of database clusters.

The integration of remote sensing data with GIS has been developed in the past two decades. Techniques such as machine learning and deep learning offer great potential for better extraction of geographical information from remote sensing data and images. However, issues remain as data organization, algorithm construction and error and uncertainty handling. With the increased volume and complexity of remote sensing data acquired from multiple sensors using multispectral and hyperspectral devices with multi-angle views with the time, new development is needed for visualization tools with spatial, spectral and temporal analysis (Wilkinson 1996; Chen and Zhang 2014).

3.4. Remote sensing data management

The generalization, standardization and serialization of remote sensing data and loads are the inevitable trend in future remote sensing technological development. It is the basis for solving the problem of inconsistent remote sensing data. It is also the prerequisite for promoting the application of remote sensing big data. At present, remote sensing satellites are developed and operated by independent institutions or commercial companies; therefore, basically all the satellites have developed their own product system standards, but the lack of a set of a unified product standard system has resulted in misperception of data and hindered development of remote sensing data applications.

The technical committee of Geographic Information/Geomatics of International Organization for Standardization (ISO/TC 211), Defense Geospatial Information Working Group (DG/IC), American National Standards Institute (ANSI), Federal Geographic Data Committee (FGDC), and German Institute for Standardization (DIN) all have established and published standards related to remote sensing data. Examples are <<ISO/TS 19101-2 Geographic information — Reference model — Part 2: Imagery>>, <<ISO/TS 19131 Geographic information — Data product specifications>>, <<ISO/DIS 19144-1 Geographic information — Classification systems — Part 1: Classification system structure>>, <<ISO/RS 19124 Geographic information — Imagery and gridded data components>> and <<ISO 19115 Geographic information — Metadata>>. However, more standards are needed for consistent applications of remote sensing data from multiple sources.

For geometric information retrieval, MODIS creates a MOD 03 file of geolocation data to store latitude/longitude information corresponding to each pixel in addition to the MOD 02 file of calibrated geolocated radiance. The image of each MODIS scene covers fairly large area. If a specific geographic region is the focus, the data of the whole image to cover a large more area has to be loaded and mapped to the surface of the 3D sphere of the earth. This leads to unnecessary large-volume data handling, difficult data visualization at different levels and scales, and ineffective data transmission. Furthermore, mapping of MODIS images to the sphere in general only uses image four corners as reference points so that mapped images often have relatively large geometric deformation.

In order to solve the problem of image mapping on the 3D sphere, the earth surface can be divided into blocks. In each block the pyramid of images is created and the scene can be visualized with image blocks at different resolutions with the altitude of viewing point. Widely used World Wind (NASA), GoogleEarth (Google Inc., Mountain View, CA) and BingMaps (Microsoft) visualize geospatial data through such a method. For example, World Wind, which is more specialized with remote sensing data, expands the 3D sphere into a 2D flat map through Plate Carrée projection (Snyder 1993). Then, the flat map is cut into blocks globally. Map cutting occurs through the division of the map evenly at different levels. The first level has the cutting interval of 36°. The second level has the cutting interval of half of the first level, i.e., 18°. And so on, each follow-up level has the interval of half of the previous level. In this way with the global projection of 3D sphere to 2D flat map, the longitude is as the horizontal axis with the range of -180° to +180° and the latitude is as the vertical axis with the range of
3.5. Five-layer-fifteen-level remote sensing data management

Scientists in the Institute of Remote Sensing and Digital Earth (former Institute of Remote Sensing Applications), Chinese Academy of Science (CAS) developed an innovative five-layer-fifteen-level (FLFL) remote sensing data management structure (Wang et al. 2012; Yu et al. 2012; Gu et al. 2013). The FLFL structure is used to block the sphere surface of the earth with each block filled with a 1000×1000 image. Each of the five layers has three levels with a size proportion of 2:5:1. The size difference between layers is 10. Therefore, the sizes of the three blocks in the first layer are 50°×50°, 25°×25° and 10°×10° in sequence. The block sizes in the second layer are 5°×5°, 2.5°×2.5° and 1°×1°. And so on for the layers of 3, 4 and 5. This data management structure is well fitted for remote sensing data organization with the grid of latitude and longitude to match up with the commonly used map scales. Table 3 illustrates the scheme for remote sensing data blocking with the FLFL structure.

Fig. 1 shows the image blocks of a MODIS image covering the area of longitude in [270°, 320°] and latitude in [130°, 160°] with the FLFL scheme (the origins of the latitude and longitude were redefined to avoid negative values for the blocking algorithm). With 1 km image resolution, the corresponding block size is 10° and the

![Fig. 1 MODIS image FLFL (five-layer-fifteen-level) blocking and naming convention.](image-url)
image contains fifteen $10^4 \times 10^4$ blocks at the third level of the first layer (Table 3). This figure also shows the naming convention of the FLFL scheme. The standardized image naming can help speed up image retrieval. In general, the FLFL structure can be used for different kinds of remote sensing images. The structure has been used for managing MODIS, Landsat and Gaofen (China National Space Administration) data. Through FLFL, data management, massive remote sensing data are stored in the distributed mode. Moreover, the data path can be directly determined through the method of data retrieval through direct access to data address. In this way, any new-generated massive remote sensing data can be blocked and sliced into image tiles on the surface of the earth sphere and the data of the image tiles can be stored and queried quickly. In such a way, the FLFL structure can provide powerful support for rapid storage, production, retrieval and services of remote sensing big data to the public. In 2014, a similar blocking/tile remote sensing data analysis and access structure called Australian Geoscience Data Cube has been established by Geoscience Australia (GA), CSIRO (Commonwealth Scientific and Industrial Research Organization) and the NCI (National Computational Infrastructure) to analyze and publish Landsat data GA archived covering the Australian continent (http://www.datacube.org.au). The Data Cube has made more than three decades of satellite imagery spanning Australia’s total land area at a resolution of 25 square meters become available, and provides over 240,000 images showing how Australia’s vegetation, land use, water movements and urban expansion have changed over the past 30 years (Lewis et al. 2016; Mueller et al. 2016).

Open Geospatial Consortium (OGC) (Wayland, MA, USA) defined a Discrete Global Grid Systems (DGGS) as "A form of Earth reference that, unlike its established counterpart the coordinate reference system that represents the Earth as a continual lattice of points, represents the Earth with a tessellation of nested cells." The DGGS is “A solution can only be achieved through the conversion of traditional data archives into standardized data architectures that support parallel processing in distributed and/or high performance computing environments. A common framework is required that will link very large multi-resolution and multi-domain datasets together and to enable the next generation of analytic processes to be applied. A solution must be capable of handling multiple data streams rather than being explicitly linked to a sensor or data type. Success has been achieved using a framework called a discrete global grid system (DGGS)." (http://www.opengeospatial.org/projects/groups/dggsdgwg). Therefore, the GRID Cube based on the FLFL scheme agrees with the definition and purpose description of the OGC-DGGS. With the FLFL scheme, a new FLFL GRID Cube was created to become a general DGGS Software entity with unique interfaces, standardized integration of multiple-source, heterogeneous and massive spatial data, distributed storage, parallel computing, integrated data mining, diverse applications and high-efficiency network services while OGC only provides the definition and description of DGGS and OGC-DGGS Core Standard for defining the components of DGGS data models, methods of frame operation and interface parameters in general.

The uniqueness of the FLFL structure is in that the resolution structure of the blocking tiles is 500 m/250 m/100 m /50 m/25 m/10 m/5 m/2.5 m/1 m and this structure can be expanded up and down indefinitely. The layer/level are divided for one layer with three levels with the ratio of 5:2.5:1 and the ratio of layers is 10:1. This diving ratios can match up commonly used map scales like 1:1 000 000, 1:500 000, 1:250 000, 1:100 000, 1:15 000, 1:25 000, 1:10 000, etc. Compared to the traditional ratio of $m^2$, such as 2$^2$/2$^2$/2$^2$, 2$^5$/2$^5$/2$^5$, 3$^2$/3$^2$/3$^2$, 2$^3$/2$^3$/2$^2$ and 5$^2$/5$^2$/5$^2$, 10$^2$/10$^2$/5$^2$ of FLFL offers the uniformity of layer/level mesh size, better fit to application scales and reduced redundancies. The decimal system is the most popular numeric system people use everyday. The ratio of 10 between layers in FLFL is convenient for being memorized and conversed.

3.6. Remote sensing data management for precision agriculture

The FLFL remote sensing data management structure has been developed for satellite imagery at different resolutions. Precision agriculture mostly deals with tactical variable-rate operations prescribed by the site-specific data and information extracted from remote sensing data in the scale of farm fields. In this way, the low-resolution (such as MODIS) and medium-resolution (such as Landsat OLI) data cannot play a role directly in precision agriculture.

Agricultural remote sensing is conducted on airplanes at different altitudes for different resolutions. Low altitude remote sensing (LARS) is very effective for precision agriculture with manned airplanes at 300–1 000 m and UAVs at 10–300 m. Another platform of agricultural remote sensing is ground on-the-go with sensors mounted on tractors or other movable structures for proximal sensing over crop fields. Therefore, high-resolution satellite remote sensing (such as Worldview and Quickbird), airborne remote sensing and ground-based remote sensing are integral to agricultural remote sensing. Specific to precision agriculture LARS and ground on-the-go remote sensing are major data sources for prescription of variable-rate applications of seeds, fertilizers, pesticides and water.

Based on the characteristics of remote sensing data for precision agriculture, a four-layer-twelve-level (FLTL) remote sensing data management structure can be built by
expanding down the FLFL remote sensing data management structure. The FLTL structure is used to block the images on the sphere surface of the earth and the images are from high-resolution satellite sensors, high-resolution airborne sensors and ultra-high-resolution UAV-based sensors with the pixel resolution from 5.57 m to 1 mm. Any images in coarser or finer resolutions can be further handled with expansion of the high-resolution and low-resolution ends of the FLTL structure. Table 4 illustrates the scheme for remote sensing data blocking with the FLTL structure. From the table, it can be seen that at the end of the highest resolution the pixel resolution is 1 mm, with which a 1:1.5 image map can be made. Now this can be created with a small UAV flying super low at like 1–3 m with a high-resolution portable camera such as GoPro Hero4 action camera (GoPro, San Mateo, CA, USA).

Fig. 2 shows the FLTL image blocks of a CIR (Color-Infrared) mosaic image of a number of images taken by a MS 4100 3-CCD (Charge-Coupled Device) multispectral camera (Optech, Inc., West Henrietta, NY, USA). The images were taken at an altitude over about 300 m to cover part of the mechanization research farm of USDA Agricultural Research Service (ARS), Crop Production Systems Research Unit in Stoneville, Mississippi. The area

<table>
<thead>
<tr>
<th>Layer</th>
<th>Level</th>
<th>Block size (°)</th>
<th>Sphere dimension (km)</th>
<th>Pixel size (m)</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>50</td>
<td>5566.00</td>
<td>5566.00</td>
<td>1:50,000,000</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>25</td>
<td>2783.00</td>
<td>2783.00</td>
<td>1:25,000,000</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>10</td>
<td>1113.20</td>
<td>1113.20</td>
<td>1:10,000,000</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>5</td>
<td>556.60</td>
<td>556.60</td>
<td>1:500,000</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>2.5</td>
<td>278.30</td>
<td>278.30</td>
<td>1:250,000</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1</td>
<td>111.32</td>
<td>111.32</td>
<td>1:100,000</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>0.5</td>
<td>55.66</td>
<td>55.66</td>
<td>1:50,000</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.25</td>
<td>27.83</td>
<td>27.83</td>
<td>1:250,000</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.1</td>
<td>11.13</td>
<td>11.13</td>
<td>1:100,000</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>0.05</td>
<td>5.57</td>
<td>5.57</td>
<td>1:50,000</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>0.025</td>
<td>2.78</td>
<td>2.78</td>
<td>1:250,000</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.01</td>
<td>1.11</td>
<td>1.11</td>
<td>1:100,000</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>0.005</td>
<td>0.56</td>
<td>0.56</td>
<td>1:50,000</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>0.0025</td>
<td>0.28</td>
<td>0.28</td>
<td>1:250,000</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>0.001</td>
<td>0.11</td>
<td>0.11</td>
<td>1:1000</td>
</tr>
</tbody>
</table>

![Fig. 2 MS 4100 CIR image FLTL (four-layer-twelve-level) blocking.](image-url)
ranges at longitude in (−90.872°, −90.867°) and latitude in (33.4425°, 33.450°). With the FLTL scheme, the image is divided into six 0.0025°×0.0025° blocks at the fifth level in the second layer (Table 4). After resampling, the entire image has a resolution of 0.28 m/pixel. Such processed and organized images can serve for decision support for local precision agriculture and at the same time the images can be shared and referred in the global perspective.

The complexity and frequency of remote sensing monitoring for precision agriculture, especially LARS monitoring, have produced massive volume of data to process and analyze. This is the new horizon of big data from agricultural remote sensing. In the past, the management of the remote sensing data for precision agriculture was organized in file-based systems and separated from processing and analysis tools. With the accumulation of the data from all the dimensions in time (yearly, quarterly, monthly, daily, hourly and even minutely), spatial location and spectral range, the management of the remote sensing data for precision agriculture requires that the data are organized, processed and analyzed in a unified framework so that the stored, processed and analyzed data and products can be shared locally, nationally and even globally. To meet the requirements, the data can be fed into the FLTL framework and streamline into the flow of data processing, analysis and management as shown in Fig. 3. In the flow image coverage clustering is the key for best use of the images to reduce redundancy and get rid of the coverage of non-agricultural areas.

4. Applications of agricultural remote sensing big data

4.1. China agricultural remote sensing monitoring service

China Agriculture Remote Sensing Monitoring System (CHARMS) was originally developed by the Remote Sensing Application Center in the Ministry of Agriculture (MOA) of China (Chen et al. 2011). The system has been operational since 1998. It monitors crop acreage change, yield, production, growth, drought and other agro-information for 7 major crops (wheat, corn, rice, soybean, cotton, canola and sugarcane) in China. The system provides the monitoring information to MOA and related agriculture management sectors according to MOA’s Agriculture Information Dissemination Calendar with more than 100 reports per year.

The CHARMS system is a comprehensive operational crop monitoring system in the Remote Sensing Application Center in MOA of China (Fig. 4). The system consists of a database system and six modules for crop acreage change monitoring, crop yield estimation, crop growth monitoring, soil moisture monitoring, disaster monitoring and information service, respectively. Now, the 7 crops as mentioned above are monitored. More crops are being added into this system progressively.

The monitoring intervals of the system are every 10 days for crop growth and soil moisture monitoring, every 30 days for crop yield estimation, 20–30 days before harvest for sowing area and yield prediction, every 30 days for grass growth monitoring, and once a year for aquacultural area estimation.

A space and ground integrated network system for agricultural data acquisition is operated to coordinate multiple sources of remotely sensed crop parameters from satellites and ground-based systems with WSNs (wireless sensor networks). UAVs have been utilized to capture agricultural data in very low altitudes to complement to the ground-based systems.

With the use of the CHARMS system, a series of crop remote sensing monitoring analysis have been accomplished. Remote sensing data analysis was conducted for the survey of planting areas of rice, wheat, and corn in China using high-resolution satellite imagery such as RapidEye imagery.
Fig. 3 Remote sensing image processing, analysis and management flow for supporting precision agriculture. RGB, red, green and blue; CIR, color infrared; VI, vegetation index.

With medium-resolution SPOT image analysis, crop (rice, soybean and corn) area change was determined (Chen et al. 2011). Wheat, corn, soybean and rice growth dynamics were monitored through analysis of high temporal resolution data by the integration of ground observation, agronomic models and low-resolution MODIS remote sensing imagery. Rice growth was monitored through analysis of MODIS imagery (Huang et al. 2012). Soil moisture of farm land in China was monitored through analysis of MODIS imagery (Chen et al. 2011). Crop yields were estimated through integration of remote sensing analysis with crop growth models, agricultural meteorological models and yield trend models (Ren et al. 2008, 2011). Remote sensing data analysis of MODIS imagery was conducted for drought, flood, snow and wild fire monitoring and loss assessment. Remote sensing data analysis was conducted for pest management in crop production in farm land of north and northeast of China. Remote sensing-based dynamical...
monitoring of grass growth and productivity was conducted in China (Xu et al. 2008, 2013). Besides, global agricultural remote sensing monitoring was conducted through GEO (Group on Earth Observations) and APEC (Asia-Pacific Economic Cooperation) networks to cover rice, corn and wheat in USA, Canada, Australia, Philippine, Thailand, and Vietnam (Chen et al. 2011; Ren et al. 2015).

4.2. Agricultural remote sensing systems operated for precision agriculture in Mississippi Delta

Global and national agricultural remote sensing monitoring is a coordination of satellite remote sensing and ground-based remote sensing. Remote sensing for regional and local farm monitoring and control is a coordination of airborne remote sensing and ground-on-the-go proximal remote sensing. For precision agriculture, LARS plays a critical role in providing prescription data for controlling variable-rate operations. UAVs provide a unique platform for remote sensing at very low altitudes with very high resolution images over crop fields to be supplemented to airborne remote sensing on manned aircraft and ground-based proximal remote sensing.

The Mississippi Delta is the section in the northwest of the state of Mississippi in the United States. The section lies between the Mississippi River and the Yazoo River. Agriculture is the backbone of the Mississippi Delta’s economy. This area has many advantages for massive commercial crop production with plain topography, extensive surface and ground water resources and nutrient-rich soils. Major crops produced in the Mississippi Delta are cotton, soybean, corn and rice. Major agricultural companies such as Monsanto Company (St. Louis, MO, USA), Syngenta AG (Basel, Switzerland) and Dow AgroSciences LLC (Indianapolis, IN, USA) have local office for research and farm consultation or collaborated with local companies to farming enhancement for sustainable development of agriculture in this area. In the middle of 1960s, the USDA ARS established the Jamie Whitten Delta States Research Center at Stoneville, Mississippi. The center consists of seven research units, with scientists conducting basic and applied research in the areas of biological control, crop entomology, crop genetics, cotton ginning, pest biological control, crop genomics and bioinformatics, and crop production systems to aim at agricultural problems of the Mid South area of the United States centered in the Mississippi Delta. The Mississippi Delta is well-suited for mechanized agriculture for large-scale crop production in typical of large flood plains with the area ranging from nearly flat to undulating, gentle slopes (Snipes et al. 2005). As a major research task, scientists in the Crop Production Systems Research Unit have been conducting research on developing techniques of precision agriculture on the basis of mechanized agriculture. The research has focused on two research farms located in the area of Stoneville, Mississippi (A: 33.441803°, -90.886169° and B: 33.446753°, -90.872211°, respectively) with the areas of 65 and 49 ha, respectively (Fig. 5) for engineering development and technical evaluation of aerial application technology, aerial remote sensing systems building and methods development for crop growth monitoring and stress detection, and system evaluation for aerial variable rate application with prescription from remote sensing monitoring. The valuable results and information from the researchs in the farms have been extended to support Mississippi Delta agricultural development. Since 2008 three terabyte data from multispectral, hyperspectral and thermal imaging sensors have been accumulated with an average increase of 30 gigabytes per year for various studies of crop stress from herbicide damage (Huang et al. 2010b; Huang Y et al. 2015), weed herbicide resistance (Reddy et al. 2014), water deficiency (Thomson et al. 2012) and nutrient deficiency (Huang et al. 2013a). In image processing and analysis, various methods and algorithms have been developed and applied and with the increase of the data volume and complexity the challenges at data storage, computation and system input and output (I/O) in data management and application will become serious issues. The adoption of the FLTL structure for such data management would help distributed data storage, computational decomposition for parallel processing, and relief of limited system I/O capability, which would result in effective applications driven by processing and analysis of
the data with such management. Fig. 6 shows the original mosaicked RGB image and blocked and resampled images of the two farms of USDA ARS Crop Production Systems Research Unit with the FLTL scheme in the blocks of 0.0005°×0.0005° at the fourth level in the second layer. The raw RGB images were acquired using a 10MP GoPro HERO3+ camera (GoPro Inc., San Mateo, California). The GoPro camera was used on UAVs for low-altitude small field imaging. In order to cover larger areas at the scale of research farms the camera was mounted and operated on an Air Tractor 402B airplane (Air Tractor Inc., Olney, TX, USA) with a 2.97 mm f/4.0 non-distortion lens at the altitude about 900 m for a spatial resolution at about 55 cm/pixel. Multiple images were acquired to be mosaicked to cover the farm areas and the resulting images were georectified and resampled to fit into the FLTL structure.

5. Comments and outlook

Agricultural remote sensing big data will be developed and used for the studies at the global, regional and field scales. Agricultural studies face the challenge of uncertainties from the variations of weather conditions and management strategies. Remote sensing big data are the valuable resource for precision agriculture to potentially make robust distributions of agricultural variables, such as yield and other biotic and abiotic indicators of crops, to tackle the uncertainties from experiments and analysis from different sites and farms. The global and regional trends identified from the big data are definitely important for the studies of global and regional agriculture, but they are definitely not capable of addressing the issues of individual farms. In this way, local remote sensing data management is as important as large-scale remote sensing data management. Large-scale big data could tell the general trends while the local data provides specific features of the farm and fields with the weather information. For site-specific recommendations, large-scale data and local data have been coupled together to balance between big data and local conditions (Rubin 2016). The established FLTL structure provides the management and application framework for large-scale remote sensing big data while the proposed FLTL structure in this paper is the management and application framework of agricultural remote sensing big data for precision agriculture and local farm studies.

Science and technology of remote sensing are advancing with the advancement of information and communication technology. Remote sensing data industry has been boomed up and the supply chain of remote sensing data from raw data to products is going to establish with the explosion of data along the chain. Working with big data for extraction of useful information to support decision making is one of the competitive advantages for organizations today. Enterprises are balancing the analytical power to formulate the strategies in every aspects in the operations to reduce business risk (Biswa and Sen 2016). The developing market of remote sensing data requires the industry to define and establish the supply chain management for remote sensing big data. For this purpose, an agricultural remote
sensing big data architecture for remote sensing data supply chain management will be built with the state of the art technology of data management, analytics and visualization. The sharing, security, and privacy requirements of the remote sensing big data supply chain system will be defined and developed accordingly.

The temporal dimension of remote sensing big data generates two insight information, historical trend and current status. The historical trends can be derived with the archived data while the current status has to be determined with real-time data acquisition, processing and analysis. Remote sensing for earth observation generates massive volume of data everyday. For the information that has a potential significance, the data have to be collected and aggregated timely and effectively. Therefore, in today’s era, there is a great deal added to real-time remote sensing big data than it was before (Rathore et al. 2015). Rathore et al. (2015) proposed a real-time big data analytical architecture for remote sensing satellite application. Accordingly, it would be expected that a real-time big data analytical architecture for precision agriculture will appear soon.

As evolved from artificial neural networks, deep-learning (DL) algorithms have been widely studied and used for machine learning in recently. DL learns and identifies the representative features through a hierarchical structure with the data. Now, DL is being used for remote sensing data analysis from image preprocessing, classification, target recognition, to the advanced semantic feature extraction and image scene understanding (Zhang et al. 2016). Chen et al. (2014) conducted DL-based classification of NASA Airborne Visible/Infrared Imaging Spectrometer hyperspectral data with the hybrid of principle component analysis, DL stacked autoencoders, and logistic regression. Basu et al. (2015) proposed a classification framework that extracts features from input images from the National Agricultural Imagery Program dataset in the United States, normalizes the extracted features and feeds the normalized features into a Deep Belief Network for classification. Artificial neural networks have been developed and applied for processing and classification of agricultural remote sensing data (Huang 2009, 2010a). With the development of artificial neural networks in deep learning, agricultural remote sensing will share the results of the studies of deep learning in remote sensing data processing and analysis and develop unique research and development for precision agriculture.

Overall, agricultural remote sensing has a number of requirements from big data technology for further development:

- Capability of global remote sensing acquisition and service.
- Rapid location and retrieve of remote sensing data for specific application.
- Data processing capability at the scales of global, national, regional and farm.
- Standardized agricultural remote sensing data interactive operation and automated retrieve.
- Tools of agricultural remote sensing information extraction.
- Visual representation of agricultural remote sensing information.

To meet the requirements, the following works have to be accomplished in the next few years:

- Standardize agricultural remote sensing data acquisition and organizing.
- Agricultural information infrastructure building, especially high-speed network environment and high-performance group computing environment.
- Agricultural information service systems building with better data analysis capability with improved, faster and complete mining of agricultural remote sensing big data in deeper and broader horizons.
- Building of agricultural remote sensing automated processing models and agricultural process simulation models to improve the quality and efficiency of agricultural spatial analysis.
- Building of computationally intensive agricultural data platform in highly distributed network environments for coordination of agricultural information services to solve large-scale technical problems.

Acknowledgements

This research was financially supported by the funding appropriated from USDA-ARS National Program 305 Crop Production and the 948 Program of Ministry of Agriculture of China (2016-X38). Assistance provided by Mr. Ryan H. Poe of USDA ARS was greatly appreciated for acquiring and processing images of USDA ARS.

References


Jacquemoud S, Baret F. 1990. PROSPECT: A model of leaf


Thomson S J, Ouellet-Plamondon C M, DeFauw S L, Huang


Section editor ZHANG Wei-li
Managing editor SUN Lu-juan