Innovation on Geolocation and Pattern Recognition for Paddy Growth Stages Reporting in Indonesia

Agustan 1*, Swasetyo Yulianto 1, Lena Sumargana 1, Heri Sadmono 1, Fauziah Alhasanah 1

1Center of Technology for Regional Resources - Agency for the Assessment and Application of Technology (PTPSW - BPP Teknologi), Jakarta - Indonesia

agustan@bppt.go.id

Abstract. Information on paddy growth stages is important for rice yield prediction. By utilizing statistical approach such as Area Frame Sampling method, data on paddy growth stages in the certain area may be used to estimate its harvest potential. The method of area frame sampling is based on segment observations determined from a stratified random sampling. Normally, information of paddy growth stages is obtained from terrestrial and remote sensing method. Since the life cycle of paddy is around 120 days, temporal resolution for observation becomes the main consideration. There is a need to develop a robust location-based reporting system for paddy growth stage. This system is designed by controlling the observer to report paddy condition on a determined location, called segment; the observer must report from the center of each segment. This paper discusses the innovation on the use of mobile phone for geolocation and pattern recognition to collect paddy growth stage data. The GPS on the mobile phone is explored for geolocation whereas the camera on the mobile phone is utilized to capture the paddy images. This information is then sent to the server for automatic pattern recognition. The statistical method with pre-processing, feature extraction, classification, feature selection and learning were applied on pattern recognition. Testing on geolocation was conducted in the Java area since May 2017 and installed for 2,356 observers. It is found that 73% of observers successfully reported the paddy growth stage from the center of each segment, which is locked in 10-meter accuracy, and 27% reported not from the center of segment due to field conditions. It is also found that a combination of GPS and mobile network or assisted-GPS can speed up the positioning.

1. Introduction

Indonesia has established the Law No. 18/2012 for national food security policy, which then followed by the Government Regulation No. 17/2015. These legal bases mandate the institution responsible for providing information about food, that cover food information system, area planted, area harvested, productivity, land use, and machinery.

In Indonesia, rice is the most staple food compared to corn, sago, and cassava [1], therefore paddy becomes one of the priority crops to monitor. Paddy field in Indonesia is commonly found throughout the country with a total area around 8 million hectares. The most productive paddy field is located in Java, Sumatra and Sulawesi islands with different varieties. Most paddy varieties have 120 days life cycle and its growth stage is classified into three main stages: vegetative, reproductive and ripening [2].
Paddy growth stage is an important parameter for rice crop monitoring [3], [4], that by knowing the area and distribution of paddy growth stage, the time and quantity of harvest can be estimated.

Recently, two methods have been developed for paddy growth stage monitoring which is satellite-based remote sensing and statistic-based terrestrial observation. Satellite-based remote sensing method has the advantage of area coverage and repeats observations (temporal resolution), whereas statistic-based terrestrial observation method has the advantage of actual paddy condition. Both methods are assessed to improve paddy reporting mechanism in Indonesia that conducted on the monthly, quarterly and yearly basis.

In Indonesia, statistic-based terrestrial observation method collects two basic information of paddy crop: complete report of harvested area and amount of productivity with field survey. In data acquisition method based on statistical approach, Area Frame Sampling (AFS) was developed to improve the data quality of the harvested area and was used by Statistics Indonesia, a government agency responsible for statistical data. AFS may be defined as a listing of land areas from which the sample is selected as the basis for identifying all the statistical units to be enumerated [5]. Since AFS is linked with location-based [6] reporting mechanism, it is important to develop a system that can handle the reporting directly from the field in a simple way using a compact and mobile device. Compact and mobile device means the integration of required and functional sensors such as communication, positioning, and camera in one and handy device, which is found in mobile phones [7]. This paper describes the innovation on utilizing android based mobile phones by developing geolocation for data and image acquisition, and pattern recognition as a control system in the main server to support the implementation of AFS in paddy growth stages reporting in Indonesia.

2. Data and Method

The goal of this research is to improve the data quality in paddy growth stages reporting in Indonesia, therefore the AFS method is the foundation of this research. The ability of smartphone in data transfer through the internet connection combined with positioning and image capturing functions is assessed and developed in one integrated reporting system. Due to the compatibility and availability of smartphone, this research exploits Android-based smartphone by developing a system with Java and Eclipse programming languages [8], whereas MySQL, PHP and Python script are developed for database server application. The system has consisted of two components: operational and office. The operational component is developed for the observer and installed on their smartphone to collect and report data. The office component is developed to receive, classify and archive data then disseminate the information.

![Figure 1. Basic data observation, communication and reporting workflow](image-url)
2.1. Operational component with geolocation function

The operational system is developed to combine the geolocation or navigational function with several features including accommodating segment with “google map image” as background or base map, determining location based on GPS and assisted-GPS [9] facility, acquiring image by embedded camera, and sending data to server as reporting through data/internet connection.

![Diagram of coordinates determination based on AFS Method](image)

AFS method adopts the locational-based system, which means data reported must contain the real condition from a pre-determined location. To ensure that the observer reports from the correct position, this system enables reporting paddy growth stage only from coordinates in the center of the sub-segment as a final product from AFS. This means that the observer must find the pre-determined coordinates using the navigation function in the system, and after reaching the exact location, functions for image capturing and data reporting are then activated automatically.

![Diagram of Coordinates Determination based on AFS Method](image)
2.2. Office component with pattern recognition function

The office system applied to pattern recognition to distinguish paddy growth stage that works in the database server. Various paddy images are taken from the field by utilizing image capture feature that is built into the system. This feature is enhanced with image compression capability hence the image size is standardized into approximately 500 kilobytes.

![Figure 4](image)

**Figure 4.** Captured images from field and paddy phenological stage (modified from IRRI-Rice Knowledge Bank)

![Figure 5](image)

**Figure 5.** Strategy for Image Processing
Paddy images received and stored in the server are classified into phenological stage (Figure 4). A digital number of color is extracted based on Hue-Saturation-Value (HSV) color space and grouped by K-means clustering. Then, these data are discriminated based on decision trees algorithm [10]. The first step is classified as brown and not brown. Image with 90% brown, 5% green, 5% others are classified as land preparation; image with 60% brown, 20% dark blue or black, 10% green, 10% others are classified into early vegetative (V1); image with 20% brown, 60% green, 10% yellow, 10% others is classified into late vegetative (V2); image with 40% green, 40% yellow, 10% brown, 10% others is classified into generative; whereas image with 75% yellow, 15% green, 10% others are classified into ripening or harvest.

3. Result and Discussion
Testing was conducted and installed for 2,356 observers, and it is found that 73% of observers successfully reported the paddy growth stage from the centre of segment, which is locked in 10 meter accuracy, whereas the other 27% reported not from the centre of segment due to field conditions. It is also found that a combination of GPS and mobile network or assisted-GPS can speed up the positioning.

For the office component, 200 images were captured and stored in the server, which was processed with the developed algorithm. The output of this program is a text file that contains location and file name input, segment code, observation value, result value, similarity matching, ranks of the dominant color, rank of pixel number. Since the classification is related to physical features of paddy growth stage, therefore the output is presented without land preparation class. Figure 6 illustrates the process of image segmentation based on HSV value to extract the true and dominant color for each image.
Figure 6 illustrates the color spectrum distribution based on pattern recognition developed using Python script. The two parallel red lines in Figure 7 represent the threshold of lower and upper limit of color digital number to be classified into phenological stages. By implementing the rules into the data, the accuracy of this system is around 60%. It is found that for vegetative phase the images can be segmented easily, however for generative and harvest phases, the images are more difficult to classify.

4. Conclusions
In urban areas or areas with well-developed communication infrastructure, assisted-GPS mode reaches the segment center faster. However, in areas with no communication signals, assisted-GPS mode displays unsettled coordinates, hence it often causes confusion. Smartphone-based positioning is useful for ensuring the quality of data reported for AFS method. The automatic image classification based on HSV image segmentation still needs improvement since the accuracy is only 60%. The similarity of the prime color of paddy growth stage and non-standardized camera used may be the cause these difficulties.

References
[1] Hartini T N S, Padmawati R S, Lindholm L, Surjono A and Winkvist A 2005 The importance of eating rice. Changing food habits among pregnant Indonesian women during the economic crisis, Soc. Sci. Med., 61 (2005), pp. 199–210.
[2] Yoshida S 1981 Fundamental of Rice Crop Science, The International Rice Institute, Los Banos, Laguna, Philippines, 269 p.
[3] Shao Y, Xiangtao F, Liua H, Xiaojn J, Rossc S, Brisco B, Brownic R and Staples G 2001 Rice monitoring and production estimation using multitemporal RADARSAT, Remote Sensing of Environment, 76, pp. 310-325.
[4] Tan C P, Koay J Y, Ewe H T, Chuah H T and Bahari S 2006 Applications of remote sensing in the monitoring of rice crops, IEM Journal, Vol. 67, No. 4, pp.1-11.
[5] Gallego F J 2015 Area Sampling Frames for Agricultural and Environmental Statistics, EUR 27595 EN, doi:10.2788/88253
[6] Lin C Y and Hung M T 2014 A location-based personal task reminder for mobile users, Personal Ubiquitous Computing, vol. 18, pp. 303-314.
[7] Casey E and Turnbull B 2011 Digital evidence on mobile devices. In Eoghan Casey, Digital Evidence and Computer Crime. Third Edition. Forensic Science, Computers, and the Internet, Academic Pres.
[8] Dexter M 2007 Eclipse and Java for Total Beginners Companion Tutorial, Licensed under the Educational Community License version 1.0, 45 p.
[9] LaMance J, DeSalas J and Järvinen J 2002 Assisted GPS: A Low-Infrastructure Approach. GPS World. March 1, 2002.
[10] Jiang L, Wang W, Yang X, Xie N and Cheng Y 2011 Classification Methods of Remote Sensing Image Based on Decision Tree Technologies. In: Li D., Liu Y., Chen Y. (eds) Computer and
Computing Technologies in Agriculture IV. CCTA 2010. *IFIP Advances in Information and Communication Technology*, vol 344. Springer, Berlin, Heidelberg

**Acknowledgments**

Ministry of Research, Technology and Higher Education – Directorate General of Research and Development for INSINAS-IRPI 2017 funding scheme; Statistics Indonesia for supports, observers and facilities