RICE FIELD MONITORING USING INTRINSIC IMAGES DECOMPOSED FROM FIELD SERVER IMAGERY

Panu SRESTASATHIERN, Narut SOONTRANON and Preesan RAKWATIN
Geo-informatics and Space Technology Development Agency (GISTDA)
120 The Government Complex, Ratthaprasasanabhakti Building
6th and 7th Floor, Chaeng Wattana Road, Lak Si, Bangkok, Thailand 10210
E-mail: panu@gistda.or.th; narut@gistda.or.th; preesan@gistda.or.th

KEY WORDS: Field server, Rice monitoring, Intrinsic image, Vegetation index

ABSTRACT: Field server is the equipment used for monitoring the rice field using RGB images acquired daily e.g., at 10 a.m. or 2p.m. The obtained RGB images are used to compute the vegetation indices which are used to analyze and monitor the rice growing stage and health. However, the reflectance information on each RGB channel is contaminated by illumination changing which can affect the computed vegetation indices such as the fluctuation of their time series. In this paper, I proposed to use the intrinsic images of the scene to compute the vegetation indices. The term intrinsic images was proposed by Barrow and Tenenbaum (Barrow and Tenenbaum (1978)) as the midlevel description of a scene. Precisely, an image of the scene is defined as the product of an illumination image and a reflectance image. Assuming that the reflectance of the scene is constant and illumination changes, the decomposition of the illumination and reflectance images uses the image sequences of rice paddy field recorded at different times. The decomposed reflectance image is then independent from the illumination change and used for computing the vegetation index. To test the performance of the proposed method, the images used was from the rice fields located in Suphanburi provinces, Thailand. The time series of the vegetation index computed from the decomposed intrinsic image was much smoother than that of fluctuating vegetation index computed from original RGB images. The smoother time series of vegetation index can improve the interpretation of rice growing state.

1. Introduction

The goal of remote sensing is to analyze or measure physical quantity of an object or phenomenon without physical contact. Especially, the most commonly used information in the remote sensing applications is the spectral information which can be corrected by receiving the reflected or direct radiation from objects. The radiometric information corrected by the sensor is distorted by the sensor characteristic and atmosphere (Schowengerdt (2006)). Therefore, the radiometric correction consists of the sensor calibration and atmospheric correction. The sensor calibration deal with finding the relation between the digital number stored at each image pixel with the recorded signal. Particularly, the gain and offset of the sensor are estimated in sensor calibration in order to convert digital number at the pixel (x,y) i.e., $DN = I(x, y)$ to the radiance value:

$$R_s = g \cdot DN + offset = gI(x, y) + offset. \tag{1}$$

The atmospheric correction is the process to transfer the at-sensor radiance to radiance at the earth's surface:

$$R_s = \tau R + R_p, \tag{2}$$

where τ is the view path transmittance and R_p the path radiance. The path radiance is the from the atmospheric scattering.

Field server is an equipment used for monitoring the rice's growing state. The primary data is the RGB image which is used for computing the vegetation index. The rice growing state can then be monitored by using the vegetation index's time series. Similar to the vegetation index's time series obtained from satellite imagery, the time series are fluctuated according to the radiometric distortion. The source of distortion is especially from the illumination change. In Figure 1, the effect of illumination change is illustrated. Therefore, in order to obtain more accurate, vegetation index's time series, the effect from illumination change must be removed similar to the case of satellite imagery. That is, we seek to find the radiance at earth surface L in (2). In this work, we do not deal with the sensor calibration in order to simplify the problem. However, the radiometric correction can be performed before the installation of the field server. The recovery of reflectance at the ground surface is based on the concept of intrinsic image decomposition. The decomposition technique only uses the image sequence with natural image statistics. The advantage of this technique is that no physical parameter e.g., viewing angle is required. The application of the intrinsic image for rice field monitoring will be demonstrated as the reduction of fluctuation in the vegetation index's time series which will help in the interpretation of the rice growing state.

The rest of this paper is organized as follow. The concept of intrinsic image and the decomposition technique used in this work is reviewed in Section 2. The use of intrinsic image in the monitoring of the rice growing state will

(a) (b)

Figure 1: Effect of different illuminations on the rice field's image. The image in Figure 1b was acquired one day after.

be introduced in Section 3. In Section 4, the results from using intrinsic image for rice growing state monitoring will be illustrated. The conclusion of this work will be discussed in Section 5.

2. Intrinsic images

The term "intrinsic image" was first introduced by Barrow and Tenenbaum (Barrow and Tenenbaum (1978)). The concept of the intrinsic image is to decompose image into reflectance and illumination. That is, the image is represented as the product between reflectance and illumination:

$$I(x,y) = L(x,y)R(x,y), \tag{3}$$

where $I(x,y)$ is the input image, $R(x,y)$ the reflectance image and $L(x,y)$ the illumination image. The goal of intrinsic image decomposition is to extract the reflectance image $R(x,y)$ from the input image $I(x,y)$. The decomposition of (3) is ill-posed problem on its own because the number of unknowns is two times the number of equation. The trivial solution for decomposing the above equation is to set $L(x,y) = 1$. However, this solution is not applicable because the illumination effect is not removed from the image.

There have been efforts on decomposing the image into illumination and reflectance by exploiting the structure in the image to reduce the number of unknowns. In (Shen et al. (2008)), Shen et al. took the texture information into account in order to give constraints on reflectance among pixels. That is, texture information was used to regularize the decomposition problem. Daniilidis et al. (Jiang et al. (2010)) proposed a new feature called local luminance amplitude which was integrate with hue and texture to extract reflectance images. In (Rother et al. (2011)), Gehler et al. used probabilistic approach to decompose the reflectance image i.e., conditional random field with the assumption that the reflectance value for each pixel is drawn from some mixing components. In (Laffont et al. (2012)), Laffont et al. used images taken from multiple view to reconstruct 3D points and normals which were used to derive relationships between reflectance values at different locations.

In this work, the approach proposed by Weiss (Weiss (2001)) is adopted because only image sequence of a static scene is required and the computation is not expensive. It made an assumption that the illumination variations are sparse and can be approximated by Laplacian distribution. Moreover, the concept of this approach relies on the temporal structure such that:

$$R(x,y,t) = R(x,y). \tag{4}$$

That is, this approach assume that the camera is stationary and the reflectance of the scene does not change. However, the problem is still ill-posed because the number of unknowns is reduced from $2T$ to $T+1$, where T is the number of equation. To obtain the solution, the problem is slightly changed to extracting the stationary background from the image sequence.

By assuming that the camera is stationary, in general situation, an image of static background can be decomposed as:

$$i(x,y,t) = \text{static background} + \text{dynamic foreground}. \tag{5}$$

In our case, the static background is referred to as the reflectance image and the dynamic foreground the illumination image. A popular approach for extracting the background is to use the (pixel-wise) median of the image sequence:

$$\text{static background} = \underset{t}{median}\; i(x,y,t). \tag{6}$$

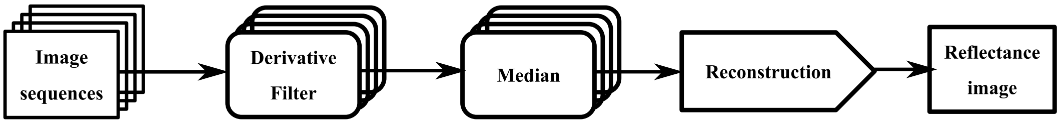


Figure 2: The flow of the intrinsic image decomposition.

The concept of using median to extract the reflectance image can be applied by using the logarithm of (3) :

$$\log(I(x,y,t)) = \log(L(x,y,t)R(x,y,t)) \tag{7}$$
$$i(x,y,t) = l(x,y,t) + r(x,y) \quad (\because R(x,y,t) = R(x,y)) \tag{8}$$

Another assumption is that the gradient of the illumination image are Laplacian distribution in time and space. Let $\{f_n\}$ be a set of N filter. In this case, $\{f_n\}$ is the set of derivative filter i.e., the derivative in the vertical and horizontal direction. The filter outputs are denoted by:

$$o_n(x,y,t) = f_n * i(x,y,t) \tag{9}$$
$$= f_n * (l(x,y,t) + r(x,y)) \tag{10}$$
$$= l_n + r_n \tag{11}$$

The maximum likelihood estimation of the reflectance image r_n can be obtained by using the median overtime:

$$\widehat{r}_n(x,y) = \underset{t}{median}\ o_n(x,y,t) \tag{12}$$

Once the filtered reflectance image $\widehat{r}_n$ is computed. The reflectance image $\widehat{r}$ can be obtained:

$$\widehat{r} = g * \left(\sum_n f_n^r * \widehat{r}_n \right), \tag{13}$$

where f_n^r is the reverse filter of f_n i.e., $f_n(x,y) = f_n^r(-x,-y)$. The function g can be computed by solving the equation:

$$g * \left(\sum_n f_n^r * f_n \right) = \delta \tag{14}$$

3. Rice field monitoring using intrinsic images

Assuming that the scene of rice paddy field is statics filed, the image sequences of the scene is used to extract the intrinsic image following the technique presented in Section 2. As it is already mentioned, the sensor calibration is not concerned in this work. This mean that the digital number DN is used as the radiance at the sensor i.e., $R_s = DN = I(x,y)$, see (1). Since the field server is installed at the earth surface, the path radiance does not have any effect. In other word, the radiance at the sensor is then:

$$I(x,y) = \tau R. \tag{15}$$

The above radiometric distortion is analogous with the model used for intrinsic image decomposition (3). Namely, the intrinsic image can be used as the estimation of the reflectance at the ground surface i.e., rice field. Given the image sequence from the field server $\{I_i\}$ where $j = 1, \ldots, N$, the reflectance on the j^{th} image I_j can be decomposed following (12):

$$\widehat{r}_n(x,y) = \underset{j}{median}\ f_n * \log(I_j(x,y)), \tag{16}$$

where $i = i - w, \ldots, i + w$ and w is the parameter of the moving window with size $2w+1$. The reflectance can then be recovered using (13). Once the intrinsic image is obtained, it is then used to compute the vegetation index.

In literature, there exists various vegetation indices. Some of them that can be computed using only RGB image are presented in this Section. In (Guijarro et al. (2011)), Guijarro et al. utilized various indices in the segmentation of agricultural images. Such indices include Excess Green (ExG), Excess Red (ExR), Excess Blue (ExB) and Excess Green minus excess Red (ExGR). Given an image in the RGB color space, the computation of those indices starts from the normalization of spectral red (R), green (G) and blue(B) components at a certain pixel:

$$r = \frac{R}{R+G+B}\ ,\ g = \frac{G}{R+G+B}\ ,\ b = \frac{B}{R+G+B} \tag{17}$$

As a consequence, the normalized spectral r, g and b are in the range $[0, 1]$. The indices are then computed as following:

$$
\begin{aligned}
ExG &= 2g - r - b && (18)\\
ExR &= 1.4r - g && (19)\\
ExGR &= ExG - ExR. && (20)
\end{aligned}
$$

In (Woebbecke et al. (1993)), Woebbecke et al. used green and red bands to compute Normalized Difference Index (NDI) for distinguishing plants from the soil:

$$NDI = \frac{g - r}{g + r}. \quad (21)$$

The aforementioned indices are very useful for segmenting vegetation on an image when only red, green and blue spectral bands are available.

4. Experimental results

To test the performance of the proposed method, the images used were from two rice fields located in Suphanburi and Roi-Et provinces, Thailand. The image sequence from Suphanburi province consists of 81 days i.e., one image per day; while that of Roi-Et province 105 images. The images were taken at 10 a.m. In Figure 3, the histograms of horizontal derivative filter outputs using the image shown in Figure 1a is illustrated. It can be observed that the filter response is sparse and its shape is similar to the Laplacian distribution. That is, the characteristics of the rice field image satisfy the assumption of the algorithm used for decomposing the reflectance from the image.

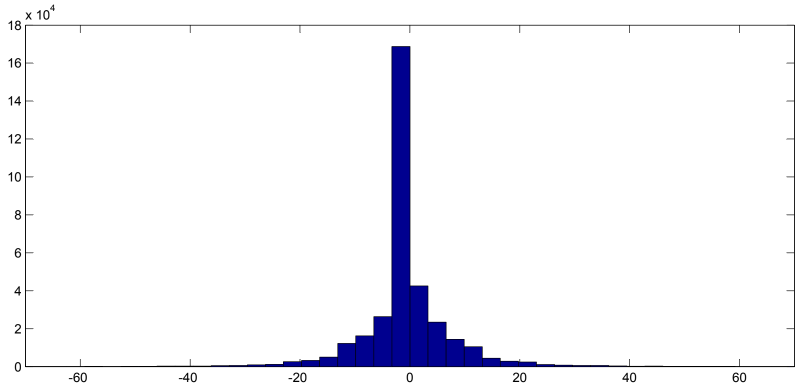


Figure 3: The histogram of horizontal derivative output using image shown in Figure 1a.

In this experiment, the size of moving window is 9 $(w = 4)$. For example, the images from 16^{th} to 24^{th} days were used to extracted the reflectance image for the image taken in 20^{th} day. The example of decomposed reflectance image is illustrated in Figure 4. It can be observed that obtain reflectance is blur especially on the rice field compared to the original one. The reason is that the rice field is not static because of wind. The image in Figure 1a looks brighter than the image in Figure 1b because the first one has less cloud. Namely, the rice field in the first image was more illuminated and looked yellowish. As a result, the vegetation indices computed from the first image (Figure 1a is less than from its true value. Note that the cloud is removed from the reflectance images. This means that the intrinsic image decomposition can remove changes that are not purely due to lighting.

The reflectance image of each image in the sequence was extracted and used in the computation of vegetation indices. The time series of vegetation indices from field server installed in Suphanburi and Roi-Et provinces are illustrated in Figures 5 and 6, respectively. In those Figure, the ExG, ExGR and NDI are plotted. The time series computed using image sequence without any pre-processing are fluctuated because of light changing. It can be observed that the time series computed from decomposed reflectance images are much smoother as promise.

5. Discussion and conclusion

In this paper, we demonstrate the use of intrinsic image decomposition for rice growing state monitoring. The intrinsic image is used as the estimation of reflectance at the earth surface which is the rice field in our case. As a result, the vegetation indices can be obtained from the intrinsic image and the fluctuation in the vegetation indices time series is reduced as promised. In the future work, the sensor calibration will be taken into account in order to obtain the real reflectance from the rice field.

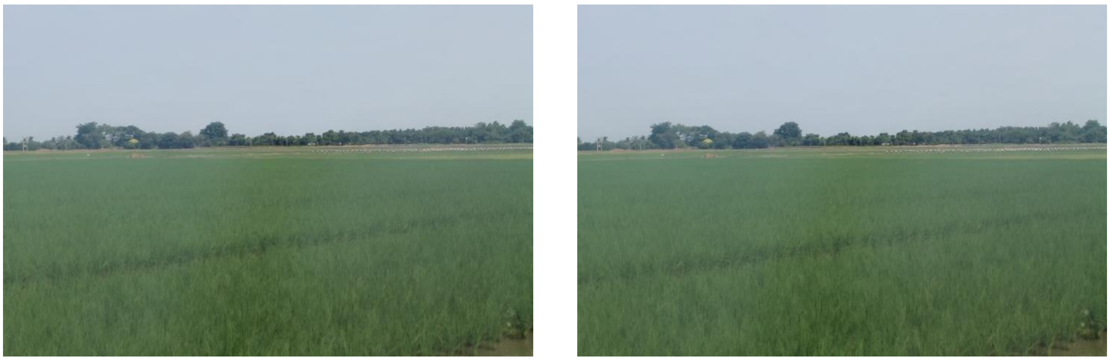
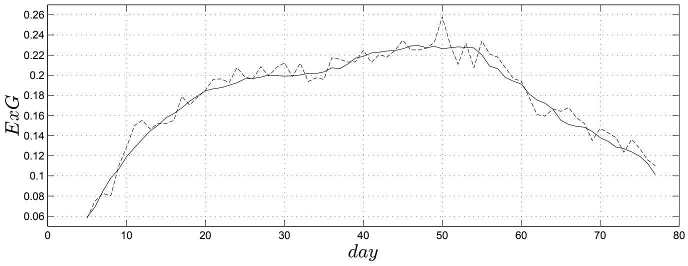
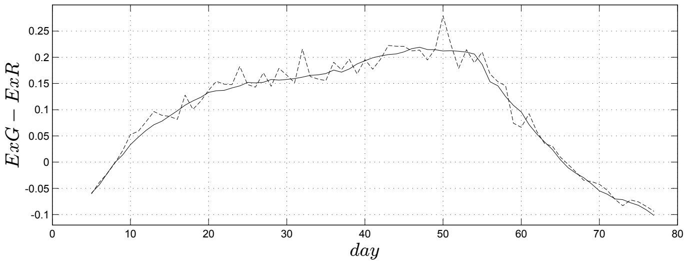


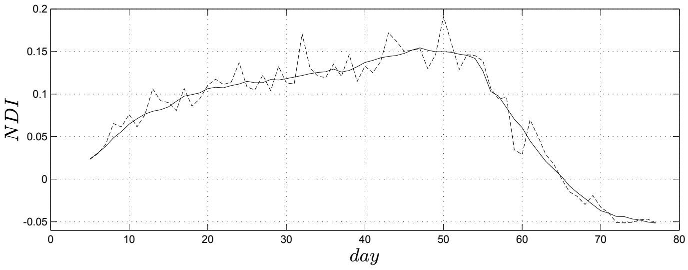
Figure 4: The reflectance image decomposed from the image shown in Figure 1a



(a) Excess Green (ExG)



(b) Excess Green minus Excess Red (ExG - ExR)



(c) Normalized Difference Index (NDI))

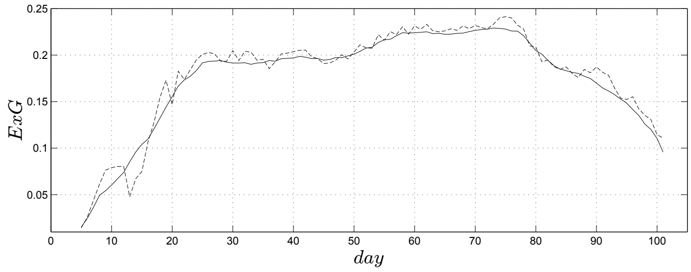
Figure 5: Vegetation index time series computed from images acquired by field server installed in Suphanburi province. The time series obtained from image sequence without any pre-processing is plotted with dash line. The time series computed from decomposed reflectance is plotted with solid line.

References

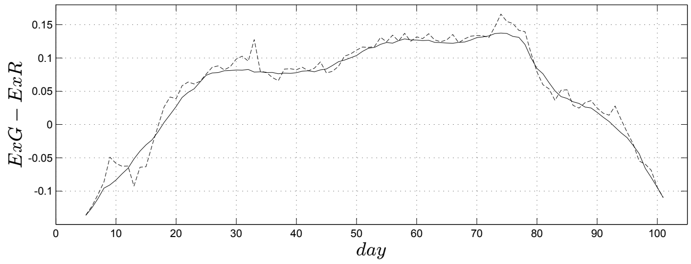
Barrow, H. and Tenenbaum, J. (1978). recovering intrinsic scene characteristics from images. In Hanson, A. and Riseman, E., editors, *Computer Vision Systems*, pages 3–26. Academic Press.

Guijarro, M., Pajares, G., Riomoros, I., Herrera, P., Burgos-Artizzu, X., and Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture*, 75(1):75 – 83.

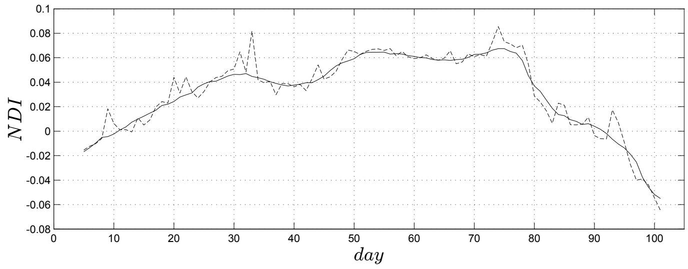
Jiang, X., Schofield, A., and Wyatt, J. (2010). Correlation-based intrinsic image extraction from a single image. In Daniilidis, K., Maragos, P., and Paragios, N., editors, *Computer Vision ECCV 2010*, volume 6314 of *Lecture Notes in Computer Science*, pages 58–71.



(a) Excess Green (ExG)



(b) Excess Green minus Excess Red (ExG - ExR)



(c) Normalized Difference Index (NDI))

Figure 6: Vegetation index time series computed from images acquired by field server installed in Roi-Et province. The time series obtained from image sequence without any pre-processing is plotted with dash line. The time series computed from decomposed reflectance is plotted with solid line.

Laffont, P.-Y., Bousseau, A., Paris, S., Durand, F., and Drettakis, G. (2012). Coherent intrinsic images from photo collections. *ACM Transactions on Graphics (SIGGRAPH Asia Conference Proceedings)*, 31.

Rother, C., Kiefel, M., Zhang, L., Schölkopf, B., and Gehler, P. V. (2011). Recovering intrinsic images with a global sparsity prior on reflectance. In Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., and Weinberger, K., editors, *Advances in Neural Information Processing Systems 24*, pages 765–773. Curran Associates, Inc.

Schowengerdt, R. A. (2006). *Remote Sensing, Third Edition: Models and Methods for Image Processing*. Academic Press, Inc., Orlando, FL, USA.

Shen, L., Tan, P., and Lin, S. (2008). Intrinsic image decomposition with non-local texture cues. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–7.

Weiss, Y. (2001). Deriving intrinsic images from image sequences. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 68–75.

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., and Mortensen, D. A. (1993). Plant species identification, size, and enumeration using machine vision techniques on near-binary images. In DeShazer, J. A. and Meyer, G. E., editors, *SPIE Conf. on Optics in Agriculture and Forestry*, volume 1836, pages 208–219. SPIE-The International Society for Optical Engineering.