Fast and robust image sequence mosaicking of nursery plug tray images

Suiyan Tan[1,2], Xu Ma[1*], Long Qi[1], Zehua Li[1]

(1. *College of Engineering, South China Agricultural University, Guangzhou 501642, China*;
2. *College of Electronic Engineering, South China Agricultural University, Guangzhou 501642, China*)

Abstract: A machine vision based system to estimate the nursery plug tray sowing quantity is necessary in the rice seedling nursery. Because the super hybrid rice is of small diameter, in order to obtain better estimation results of sowing quantity, only part of the tray scene is captured by machine vision system. Therefore, nursery tray image sequence mosaic algorithm is required so as to obtain the whole nursery tray sowing performance. In this paper, a fast nursery plug tray image mosaic technique based on Phase Correlation and Speeded up Robust Features (the PC-SURF) algorithm was introduced. To reduce huge computational complexity in feature extraction and avoid wrong match points from non-overlapping regions between two adjacent images, firstly the proposed method used Phase Correlation to approximately locate the overlapping regions between two sequential images so as to narrow down the processing area for further feature extraction. Then, the SURF algorithm was implemented in the overlapping region of the images to perform image registration and image blending. Image registration and blending were then performed using the RANSAC algorithm and transition smoothing method. Finally, sequential images were warped into a single frame to produce a panoramic tray image with high resolution and better quality. The test results showed that the PC-SURF algorithm greatly outperforms the SURF algorithm in point matching accuracy, time consumption, and image mosaic accuracy. The average feature point matching accuracy of the PC-SURF algorithm was improved by approximately 7.14%. The implementation time was almost three times faster. The Root Mean Squared Error (*RMSE*) of image blending in the R, G, B channels (*RMSEr*, *RMSEg*, and *RMSEb*) decreased by approximately 8.4%, 6.9%, 6.9% respectively. The *RMSE* of image registration *RMSEreg* decreased by 33.8%. In addition, the speed of the nursery tray image mosaic technology could meet the requirements for real-time implementation in super hybrid rice automated sowing machines.

Keywords: super hybrid rice, nursery plug tray, image mosaic, phase correlation, SURF algorithm
DOI: 10.25165/j.ijabe.20181103.2919

Citation: Tan S Y, Ma X, Qi L, Li Z H. Fast and robust image sequence mosaicking of nursery plug tray images. Int J Agric & Biol Eng, 2018; 11(3): 197–204.

1 Introduction

Super hybrid rice is an important grain crop in China and is mainly planted using nursery-transplanting techniques. To ensure high tiller ability and high yield production, super hybrid rice should be sowed at a stable and precise sowing rate of approximately 2-3 seeds per tray cell[1]. In practice, sowing performance of precision seeder is influenced by not only the operational parameters but also the physical properties of rice[2], and the sowing quantity varies during the sowing process. To solve the problem, it is necessary to precisely estimate the sowing quantity of each plug tray cell. When sowing performance gets worse, mechanical reseeding and automatic operation parameters adjustment of the seeder have to be implemented.

In the previous work, a machine vision based system was established to estimate the nursery tray sowing quantity[3], in which the method combined with feature extraction of one group of seeds and BP neural network was used to detect seed quantity. Because the super hybrid rice is of small diameter, whose average length, width and thickness were about 6 mm, 2 mm, 2 mm, respectively. The precision seeder sowed seeds onto a tray about 600 mm × 300 mm × 30 mm (length × width × height). Usually 40-50 g seeds, which were approximately 1400-1500 grains, were sown onto each tray. If one panoramic nursery tray image was captured by using an industrial camera, the image might include a lot of background information and the resolution of each seed would be poor. Consequently, the estimation result of sowing quantity got worse and the average accuracy was approximately 72.1%. Instead, the nursery tray images can be obtained in sequence, and each image captured part of the tray scene. In this way, high resolution of each seed could gain and the average estimation accuracy of sowing quantity for each image was 94.4%. But two problems might exist, one is that, if there is no overlapping region between adjacent images, some tray sowing quantity information might be lost. The other is that, if there are overlapping regions between adjacent images, there might be a lot of redundant information. It is hard to obtain the sowing performance of the whole nursery tray. So when sowing process got worse, it was inconvenient to locate the disqualified tray cells in both cases and it was hard to implement mechanical reseeding operation. Therefore, computer vision algorithms such as a fast nursery plug tray image sequence mosaic algorithm proposed in this paper are required to stitch consecutive images together to form a panoramic nursery tray image so as to obtain sowing performance of the whole nursery tray.

Received date: 2016-10-15 **Accepted date**: 2018-02-27
Biographies: Suiyan Tan, PhD, research interests: agricultural mechanization and precision agriculture, Email: tansuiyan@scau.edu.cn; **Long Qi**, PhD, Professor, research interests: modern agricultural machinery equipment, Email: qilong@scau.edu.cn; **Zehua Li,** PhD, Associate Professor, research interests: agricultural mechanization and precision agriculture, Email: 499927441@qq.com.
***Corresponding author: Xu Ma,** PhD, Professor, research interests: agricultural mechanization and automation. College of Engineering, South China Agricultural University, No.483, Tianhe District, Guangzhou 501642, China. Tel: +86-13560008927, Email: maxu1959@scau.edu.cn.

The image mosaic is a technique to combine multiple photographic images from one or more cameras with overlapping fields of view to produce a panoramic image and eliminate redundant information. Recently, there have been many applications in precision agriculture[4-8], 3D image reconstruction[9] and crop diagnostic and monitoring systems[10].

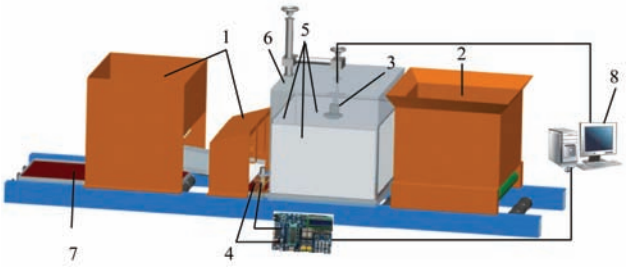
In general, image mosaic method can be divided into two steps: image registration and image blending. Currently, based on the image information used, the image registration methods can be divided into two categories: frequency domain methods and spatial domain methods. The Phase Correlation algorithm is a frequency domain method. It calculates the spatial displacement between consecutive images based on the information from their 2D Fourier transforms. It is also a simple one-step registration process that makes fast implementation possible. The Phase Correlation algorithm has been applied to many registration applications[11]. However, in the Phase Correlation algorithm, when the two images to be analyzed are not identical (except for a translative shift), multiple peak values may result from this, which approximately matches the maxima in the δ function array of the inverse Fourier transformation. To obtain a precise displacement between images, multiple iterating calculations are required[12]. Furthermore, Phase Correlation algorithm can only be implemented on two images with the same size. Therefore, drawbacks of Phase Correlation are its poor performance and strict requirements on the size of the stitched images. There are other spatial domain methods that can also be used to tackle the image registration problem, such as the template-matching algorithms and feature-based algorithms[13,14]. The Speeded Up Robust Features (SURF) algorithm proposed by Bay[15] is a feature-based algorithms. It has recently gained popularity because of its promising performance and its rotation, scale and illumination-invariant properties[16]. The SURF algorithm makes use of integral images to efficiently compute an approximation of the Hessian matrix. It is widely applied in the field of precision agriculture[10]. However, the detectors in the SURF algorithm need to work at different scales, and the search for correspondences often requires the comparison of images at different scales, which is time consuming. Additionally, SURF's descriptors are formed and described by a 64-dimensional vector for each feature, which means the more SURF feature points detected, the longer the computation time consumed. Therefore, the drawback of the SURF method is the huge computational complexity in the feature extraction stage because of the large number of features present in one entire image. Considering that the two adjacent images are only partially overlapped, if the SURF features are extracted and matched from a limited region instead of the entire image, the computational complexity can be greatly reduced[17]. The limited region can be detected using Phase Correlation algorithm.

In this paper, a fast image mosaic technique that combines Phase Correlation and the SURF algorithm (PC-SURF) was proposed, which significantly improved the performance compared with the SURF mosaic algorithm. In the PC-SURF algorithm, firstly the Phase Correlation algorithm was applied to two consecutive images to approximately locate the overlapping regions between them; then, the SURF algorithm was implemented in the overlapping region of the images to perform image registration. The test results showed that the proposed PC-SURF algorithm outperformed the SURF algorithm in feature point matching accuracy, time consumption, and image blending and registration accuracy.

2 Materials and methods

2.1 Experimental systems

Using a digital camera, sequential images with overlapping regions of interest can be stitched together to generate high-resolution panoramic images. In this study, our previously developed machine vision system[3] is used to investigate a fast nursery plug tray image mosaic technique. The system consists of five main components: a digital camera, a computer, an infrared photoelectric sensor module, illuminants and illuminant cabinet. It is placed between the precision seeder and soil coverer device, as shown in Figure 1.



1. Precision seeder 2. Soil coverer 3. Digital camera 4. Infrared photoelectric sensor module 5. Illuminants 6. Illuminant cabinet 7. Nursery plug tray 8. Computer

Figure 1 Schematic diagram of machine vision-based system

The digital camera is a Logitech C920, and its resolution is 720×960 pixels. Logitech C920 has automatic focus which made it more convenient to use in the fieldwork. It is mounted on the illuminant cabinet and aimed vertically downward to capture the top view of the nursery tray. The computer is configured with a 2.3 GHz CPU and 4 GB RAM. An infrared photoelectric sensor module is used to detect whether the nursery tray reach the capture area of the camera. In the image acquisition process, the nursery trays move from left to right across the automated sowing machine. When a nursery tray reach the capture area, the infrared photoelectric sensor detect it and feed the signal back to the connected computer, which trigger the camera to capture the tray image sequence at a fixed interval. To acquire high-quality tray images, an illuminant cabinet with four LED illuminants is installed in the capture area.

The size of the nursery tray is approximately 600 mm×280 mm, with 30×14 cells. The camera is set to capture approximately 1/3 of the size of the nursery tray, as shown in Figure 2a. To ensure that there is approximately 30% to 50% overlapping region between adjacent images, five images are captured for each nursery tray at a fixed interval. Sequential images are shown in Figures 2b-2f. The software program platform for the nursery plug tray image mosaic system is Matlab (The Math Works, Inc., USA, and Version2010).

2.2 Image mosaic methodologies

2.2.1 The Phase Correlation algorithm

The Phase Correlation algorithm exhibits a distinct maximum at the location where two images match[18]. It can be described as follows: let f_1 and f_2 be two adjacent images that are displaced by (x_0, y_0) pixels

$$f_2(x,y) = f_1(x-x_0, y-y_0) \quad (1)$$

where, x and y are any pixel locations within the image f_2; x_0 and y_0 are the displacements in x and y directions. Their corresponding Fourier transforms F_1 and F_2 are related by

$$F_2(u,v) = F_1(u,v)e^{-j(ux_0+vy_0)} \quad (2)$$

where, u and v are the x and y pixel locations, as expressed in the

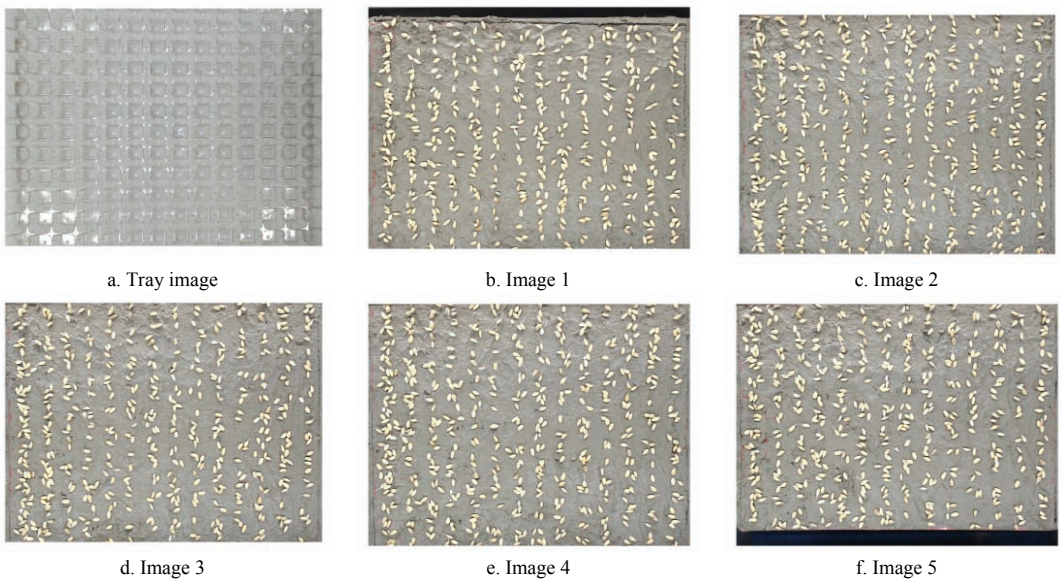
frequency domain. Rearranging Equation (2) and multiplying the numerator and denominator of the left term by the conjugate, and the normalized cross power spectrum of the two images is given as

$$\frac{F_2(u,v)F_1(u,v)^*}{\left|F_2(u,v)F_1(u,v)^*\right|}=e^{-j(ux_0+vy_0)} \tag{3}$$

The inverse Fourier transform can be applied to Equation (3), and a Phase Correlation function δ is given as

$$\delta(x_0,y_0)=F^{-1}\left[\frac{F_2(u,v)F_1(u,v)^*}{\left|F_2(u,v)F_1(u,v)^*\right|}\right]=F^{-1}[e^{-j(ux_0+vy_0)}] \tag{4}$$

where, δ is a two-dimensional function array. The δ function array has a peak at the point of the maximum correlation at (x_0, y_0) between images f_1 and f_2, and values of approximately zero elsewhere. To improve the robustness of the Phase Correlation algorithm[19], the Canny operator was applied to extract the contours of the binary sowing tray image, then, Phase Correlation was applied to the contour-images instead of the original images. Two adjacent images were captured from the machine vision-based system, and their contour-images and phase correlation map are shown in Figures 3a-3c.



a. Tray image b. Image 1 c. Image 2

d. Image 3 e. Image 4 f. Image 5

Figure 2 Nursery tray image sequence acquisition

a. Contour image 1 of nursery tray b. Contour image 2 of nursery tray c. Phase correlation map of image 1 and image 2

Figure 3 Displacement parameters estimation using Phase Correlation

In an ideal case that two images are exactly the same, there is only one maximum peak. However, two adjacent images are only partially the same, which results in multiple peak values close to the maximum in the δ function array of the inverse Fourier transformation, and there is an edge effect caused by the periodicity of the Fourier transformation, as shown in Figure 4c. Although there are multiple peaks, they are quite concentrated. Hence, we can extract one of the maximum peaks to approximately determine the displacement and locate the overlapping region of two adjacent images. If the displacement of the two images f_1 and f_2 is (a, b) and the size of the seeded tray image is m/pixels×n/pixels, the overlapping regions of the two images can be described as

Overlapping region of f_1: [a:m, b:n] (5)

Overlapping region of f_2: [1:m−a+1, 1:n−b+1] (6)

2.2.2 SURF algorithm

The SURF algorithm[15] is an emerging feature extraction method that consists of a feature detector and feature descriptor. The SURF feature is invariant to rotation, scale and illumination.

1) Fast-Hessian SURF Detector

SURF uses the determinant of the Hessian matrix as a discriminant to identify local maximum values, and it is based on scale space theory[20]. Given a pixel X=(x, y) in image $f(X)$, its Hessian matrix $H(X, \sigma)$ at scale σ is defined as

$$H(X,\sigma)=\begin{bmatrix}L_{xx}(X,\sigma) & L_{xy}(X,\sigma)\\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma)\end{bmatrix} \tag{7}$$

where, $L_{xx}(X, \sigma)$ represents the convolution result of the Gaussian second-order filter $\frac{\partial^2}{\partial x^2}g(\sigma)$ with image function $I(X)$ at point X=(x, y), which is similarly for $L_{xy}(X, \sigma)$ and $L_{yy}(X, \sigma)$. By computing the determinant of the Hessian Matrix for each pixel in the image, the feature point value is obtained.

In actual calculation, Bay[17] proposed a box filter approximation instead of a Gaussian second-order filter. L_{xx}, L_{xy} and L_{yy} are simplified as D_{xx}, D_{xy} and D_{yy}. These approximate second-order Gaussian derivatives can be evaluated very rapidly using integral images, independently of image size. Therefore, the approximate determinant of the Hessian matrix H_{approx} can be described as

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9Dxy)^2 \quad (8)$$

The construction of a scale image pyramid in the SURF algorithm is divided into octaves, and there are scale levels in each octave, which represent a series of filter response maps obtained by convolving the same input image with a box filter of increasing size. To localize feature points in the image and over scales, non-maximum suppression in a 3×3×3 neighborhood is applied. The maxima of the determinant of the Hessian matrix are then interpolated in scale and image space.

2) SURF descriptor

The SURF descriptor is obtained by calculating the main direction and feature vector[21]. First, the feature point is set as the center, and the Harr wavelet responses in the x and y directions within a circular neighborhood of radius 6 s around the feature point are calculated, where s is the scale at which the feature points are detected. The Harr wavelet responses are represented as vectors. Then, all of the vectors in the x and y directions of the Harr wavelet responses within a sliding orientation window covering an angle of size $\pi/3$ around the feature point are summed up. The two summed responses yield a new vector. Finally, the longest vector is the dominant orientation of the feature point.

For extraction of the descriptor, a square region with a size of 20 s is constructed, and the region is split into 4×4 square sub-regions with 5×5 regularly spaced sample points inside. This preserves important spatial information. For each sub-region, the Harr wavelet responses d_x in the x direction and d_y in the y direction are computed. The responses are weighed with a Gaussian kernel centered at the feature point. The responses over each sub-region for d_x and d_y are summed up separately. At the same time, the sum of the absolute values of the responses, $|d_x|$ and $|d_y|$ are extracted. Hence, each sub-region has a 4-dimensional descriptor vector v for its underlying intensity structure

$$v = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|) \quad (9)$$

Concatenating this for all 4×4 square sub-regions will result in a descriptor vector of length 64.

2.2.3 Image registration

After generating feature descriptors with 64 dimensions, the Euclidean distance similarity judgment method matches the SURF descriptors

$$D = \sum_{i=1}^{64} (v_2(i) - v_1(i))^2 \quad (10)$$

where, the minimum distance $D_{\min}$ and the second-minimum distance $D_{scn\min}$ are calculated by Equation (10). The ratio is $R=D_{\min}/D_{scn\min}$, where R is less than a given threshold. In our tests, it was defined as 0.49. After feature matching, there will be many coarse matched points. Random Sample Consensus (RANSAC)[22] was employed to refine the coarse matched points. The RANSAC algorithm estimates a transformation model (homography) according to all of the matched points. It keeps the points which comply with the models as inliers and discards the other points which fail to comply with the model as outliers. With the accurate matching points, the projective transformation matrix H_{pro}[23] between two images f_1 and f_2 is as follows

$$\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \quad (11)$$

$$H_{pro} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \quad (12)$$

where, (x_1, y_1) in image f_1 and (x_2, y_2) in image f_2 are matching points after RANSAC. With at least 4 pairs of matching points, H_{pro} can be calculated.

2.2.4 Image blending

After the projective transformation matrix H_{pro} is obtained, the overlapping region can be accurately calculated. Image blending is applied to smooth the overlapping regions and combine the images into a panoramic image. Transition smoothing methods[20], also known as alpha blending methods, attempt to minimize the visibility of the seams by smoothing the common overlapping regions of the stitched images. A slow transition weight ω from the first image to the second image in the overlapping region is implemented, which is proportional to its distance from the overlapping region.

If $f_1(x, y)$ and $f_2(x, y)$ are two adjacent images, and $f(x, y)$ is the blending image

$$f(x,y) = \begin{cases} f_1(x,y) & (x,y) \subset f_1 \\ \omega_1(x,y) f_1(x,y) + \omega_2(x,y) f_2(x,y) & (x,y) \subset (f_1 \cap f_2) \\ f_2(x,y) & (x,y) \subset f_2 \end{cases} \quad (13)$$

The smooth transition can be achieved in the x direction, y direction or in both directions. In our experiment, the smooth transition needs to be achieved in the y direction

$$\omega_1(y) = \frac{y_D - y}{y_D - y_U}, \omega_2(y) = \frac{y - y_U}{y_D - y_U} \quad (14)$$

where, y_U and y_D are the upper and lower boundaries of the overlapping region.

2.2.5 The proposed PC-SURF algorithm

The SURF algorithm need to work at different scales, and the search for correspondences often requires the comparison of images at different scales, which is time consuming. In addition, SURF's descriptors are formed and described by a 64-dimensional vector for each feature. To improve the huge computational complexity in SURF algorithm, the PC-SURF image mosaic algorithm was proposed based on Phase Correlation and SUFR algorithm. It is divided into two main steps. First, the overlapping region between two adjacent images is determined using a Phase Correlation algorithm, which aims to reduce the processing area for further feature extraction. Second, the SURF features are extracted from the overlapping region for feature point matching. These features are matched to determine the correspondence between images. Image registration and blending are then performed using the RANSAC algorithm[24] and transition smoothing method[25]. Finally, the image sequence is warped into a single frame to produce an entire plug tray panoramic image.

Before the Phase Correlation algorithm, some image preprocessing have to be done. The images are firstly segmented to rice seeds from background. The Otsu threshold is used to produce a complete binary image segmenting the seeds against the background. Then the Canny operator is applied to extract the contours of the binary sowing tray image. The operating procedure is outlined in Figure 4.

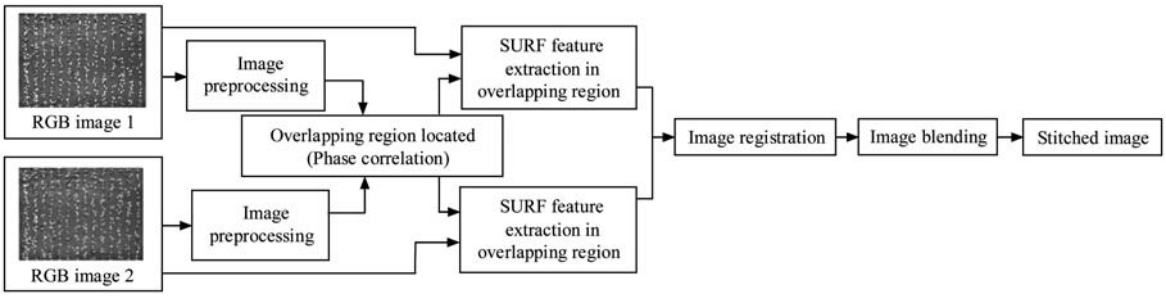


Figure 4 Flow chart of the PC-SURF algorithm framework

3 Experiments and results

3.1 Displacement estimation of two adjacent images based on Phase Correlation

Nursery trays were transported by a conveyor belt with an adjustable speed ranging from 0.062 m/s to 0.13 m/s, corresponding to 450 trays/h to 800 trays/h. The displacements of two adjacent images captured from the camera at a fixed interval of 1 s, 1.5 s and 2 s were estimated respectively. Figure 5 showed the results of the displacement of two adjacent images corresponding to different conveyor speeds, where *a* represented the vertical displacement and *b* represented the horizontal displacement. This figure was plotted by averaging 10 independent test samples for each conveyor speed.

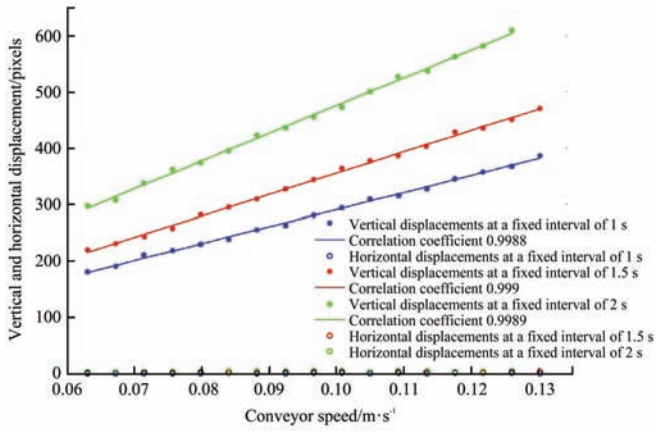


Figure 5 Vertical and horizontal displacement averaged over 10 samples with different conveyor speeds

A careful examination of Figure 5 revealed that there is a linear relationship between the conveyor speed v and vertical displacement a with different captured intervals. Equations (15), (16) and (17) expressed the linear relationship with v>0.062 m/s when the captured intervals were set 1 s, 1.5 s and 2 s respectively. Meanwhile, there was no great change in the horizontal displacement b at different conveyor speeds.

$$a = 300.5v - 9.31 \quad (15)$$

$$a = 378.2v - 22.35 \quad (16)$$

$$a = 492.2v - 16.24 \quad (17)$$

When two images with different size is about to be mosaic, Equation (15), (16) and (17) can be applied to determine the overlapping region of the images.

3.2 Nursery tray image mosaic quality estimation

Five images were captured for each nursery tray, numbered as 1, 2, 3, 4, and 5. The nursery tray sequential images shown in Figs. 2b-2f were subjected to the proposed framework shown in Figure 3. First, image1 and image 2 were stitched as image 1-2, and image 3 and image4 were stitched as image 3-4. Then, image 1-2 and image 3-4 were stitched as image 1-2-3-4. Finally, image 1-2-3-4 and image 5 were stitched to form the panoramic image. The image mosaic result was presented in Figure 6. It was evident that the panoramic view of the nursery tray was very well reconstructed, as the seams of the constituent images were hardly noticeable. Because there is slight distortion in the lens, the image sequence is transformed by perspective collineation several times and the deformation is accumulated and magnified, so there is a little distortion in the panoramic image. Furthermore, to prove the practicality of the proposed work, the SURF algorithm and the PC-SURF algorithm were tested on a nursery tray image sequence and compared with respect to the number of matching points, time consumption and image mosaic accuracy specification. In the experiments, the conveyor belt speed was set to 0.084 m/s, corresponding to 500 trays/h.

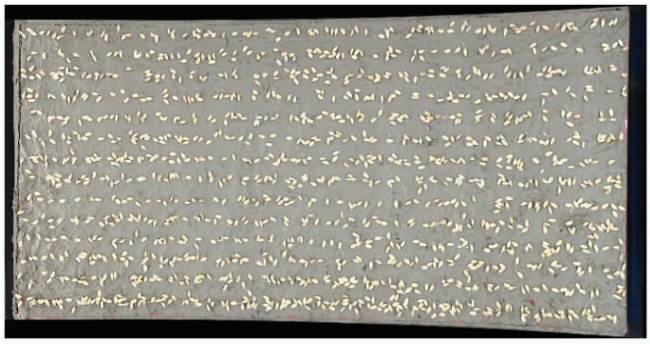
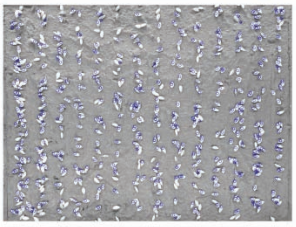
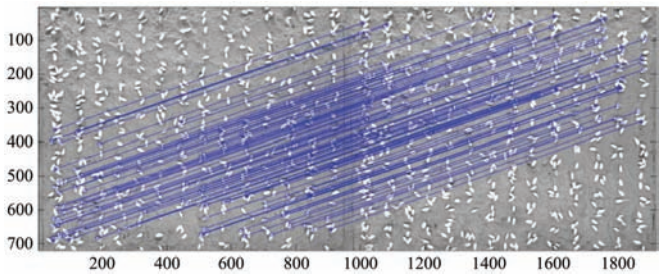


Figure 6 Nursery tray image sequence mosaic result

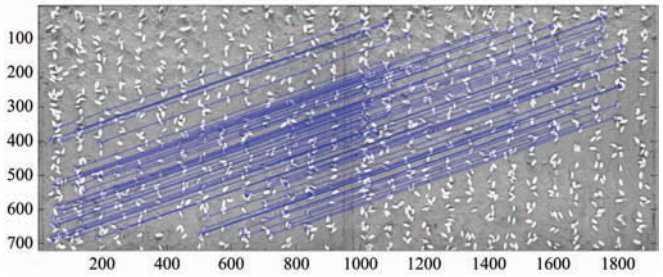
First, feature point matching quality experiments were carried out. Figure 7 showed results of the numbers of feature points, initial matching point pairs and precision matching point pairs after RANSAC when using the SURF algorithm. 638 and 693 feature points were detected in image3 and image4 respectively, as shown in Figure 7a and Figure 7b. Thereafter, the number of initial matching point pairs were 122 (shown in Figure 7c), and the number of precision matching point pairs after RANSAC were 86 (shown in Figure 7d). On the other hand, Figure 8 shows the results of the PC-SURF algorithm. Because the SURF algorithm was restricted to the overlapping region detected by Phase Correlation, only 378 and 377 feature points were detected in image 3 and image4 respectively, as shown in Figure 8a and Figure 8b. The number of initial matching point pairs were 132 (shown in Figure 8c), and the number of precision matching point pairs after RANSAC were 98 (shown in Figure 8d). The specific experimental data were shown in Table 1. Although much fewer feature points were detected using the PC-SURF algorithm than using the SURF algorithm, the matching accuracy was higher, which implied better feature point matching quality.



a. Feature points detected in image 3 b. Feature points detected in image 4

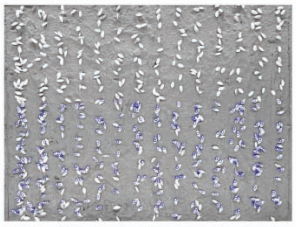


c. Initial matching point pairs of SURF algorithm

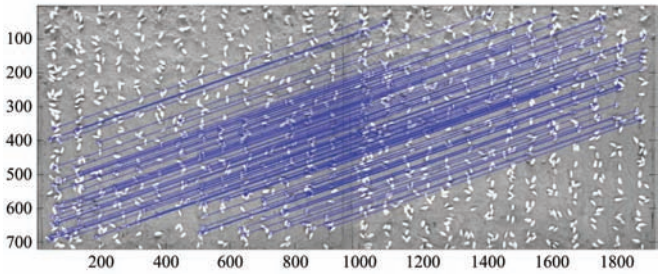


d. Precision matching point pairs after RANSAC

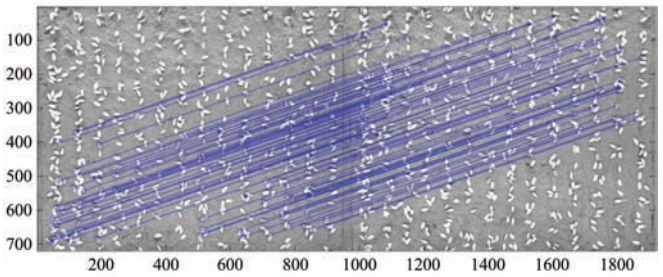
Figure 7 Feature points, initial matching point pairs, and matching point pairs after the RANSAC of the SURF algorithm



a. Feature points detected in image 3 b. Feature points detected in image 4



c. Initial matching point pairs of the PC-SURF algorithm



d. Precision matching point pairs after RANSAC

Figure 8 Feature points, initial matching point pairs, and matching point pairs after the RANSAC of the PC-SURF algorithm

Table 1 Feature point pairs matching accuracy results

	SURF algorithm					PC-SURF algorithm				
	SURF feature points in image 3	SURF feature points in image 4	Matching point pairs	Matching point pairs after RANSAC	Matching accuracy	SURF feature points in image 3	SURF feature points in image 4	Matching point pairs	Matching point pairs after RANSAC	Matching accuracy
1st mosaic	617	601	110	85	0.77	348	326	118	91	0.77
2nd mosaic	638	693	122	86	0.70	378	377	98	132	0.74
3rd mosaic	840	929	77	43	0.56	328	343	50	79	0.63
4th mosaic	1374	670	68	40	0.58	319	321	41	61	0.67

Next, the mosaic time consumption was analyzed and compared. Each test result in Table 2 was the average of 10 independent test sample results. The time consumption increased dramatically with stitching frequency, when the SURF algorithm was applied. However, the time consumption of each mosaic process was less and remained stable when the PC-SURF algorithm was used. The mosaic time consumption using the proposed algorithm was much less than that using SURF. The total time using the SURF algorithm for forming the nursery tray panoramic image was 19.37 s, while it was 6.63 s using the PC-SURF algorithm. The rice seedling automated sowing machines usually operate at a speed of 500 trays/h, which is approximately 7.2 s/tray. Thus, 6.63 s met the online testing requirements.

Table 2 Mosaic time consumption results

Time consumption/s	1st mosaic	2nd mosaic	3rd mosaic	4th mosaic	Total time
SURF algorithm	3.28	3.35	5.91	6.83	19.37
PC-SURF algorithm	1.69	1.74	1.75	1.45	6.63

Finally, the Root Mean Squared Error (RMSE) was calculated to quantitatively describe the image mosaic accuracy specification since the RMSE was unaffected by the individual error points observed. The smaller the RMSE was, the higher the matching accuracy was. The quality of the image mosaic mainly depended on the accuracy of the image registration and image blending. In our experiments, Equations (18)-(20) defined the RMSE of image blending in the R channel *RMSEr*, the G channel *RMSEg*, and the B channel *RMSEb*, respectively. Equation (21) defined the Root Mean Squared Error of image registration, *RMSEreg*.

$$RMSEr = \sqrt{\frac{\sum_{i=1}^{row}\sum_{j=1}^{col}(f_{1R}(i,j)-(f_R(i,j)))^2}{row\times col}} \quad (18)$$

$$RMSEg = \sqrt{\frac{\sum_{i=1}^{row}\sum_{j=1}^{col}(f_{1G}(i,j)-(f_G(i,j)))^2}{row\times col}} \quad (19)$$

$$RMSEb = \sqrt{\frac{\sum_{i=1}^{row}\sum_{j=1}^{col}(f_{1B}(i,j)-(f_B(i,j)))^2}{row\times col}} \quad (20)$$

$$RMSEreg = \sqrt{1/n\sum_{i=1}^{n}[(h_1x'_i + h_2y'_i + h_3 - x_i)^2 + (h_4x'_i + h_5y'_i + h_6 - y_i)^2]} \quad (21)$$

where, $f_{1R}(i, j)$, $f_{1G}(i, j)$, and $f_{1B}(i, j)$ represented the intensity values for the R, G, B channels of the base image f_1; $f_R(i, j)$, $f_G(i, j)$, and $f_B(i, j)$ represented the intensity values for the R, G, B channels of the blending image f; *row* and *col* represented the size of the overlapping region of adjacent images f_1 and f_2, which can be calculated after image registration. h_1, h_2, h_3, h_4, h_5 and h_6 were parameters of the projective transformation matrix H_{pro}, which were estimated by the RANSAC algorithm. (x_i, y_i) was feature point in image f_1, and (x'_i, y'_i) was the corresponding matching point in image f_2. The specific experimental data of the above were shown in Table 3. Each group of test results was the average of 10 independent test sample results.

Table 3 Root mean squared error (RMSE) results of image registration and blending

	SURF algorithm				PC-SURF algorithm			
	RMSEr	*RMSEg*	*RMSEb*	*RMSEreg*	*RMSEr*	*RMSEg*	*RMSEb*	*RMSEreg*
1st mosaic	13.34	13.61	14.02	4.65	12.94	13.08	13.49	4.1
2nd mosaic	15.5	15.81	15.83	4.12	15.5	15.77	15.85	3.69
3rd mosaic	21.5	21.52	20.62	17.99	20.26	20.33	19.43	6.68
4th mosaic	29.87	32.04	32.3	42.49	26.24	28.11	28.3	31.37
Average error	20.05	20.75	20.69	17.31	18.73	19.32	19.27	11.46

3.3 Analysis

Because the camera is directed vertically downward to capture the top view of the nursery tray, there shows a linear relationship between the conveyor speed and the vertical displacement, while the horizontal displacement is almost constant across different conveyor speeds. Based on the experimental results from Tables 1-3, the PC-SURF algorithm has a much better mosaic result than the SURF algorithm, in terms of the number of matching points, time consumption and the values of *RMSEr*, *RMSEg*, *RMSEb*, and *RMSEreg*. The average feature point pairs matching accuracy of our proposed algorithm is 0.70 compared to 0.65 for the SURF, an increase of approximately 7.14%, and the implementation time of our proposed algorithm is almost 3 times faster than that of the SURF algorithm. Moreover, the average image blending and registration errors of the PC-SURF algorithm, *RMSEr*, *RMSEg*, *RMSEb*, and *RMSEreg*, decrease by approximately 8.4%, 6.9%, 6.9% and 33.8%, respectively. According to the experimental data, the reasons that the proposed algorithm is more effective than the SURF algorithm can be attributed to the following factors: (1) the SURF algorithm suffers from computational complexity, as the features are extracted from the entire image instead of only the matched region. So feature point extraction from the non-overlapping region of mosaic images is useless for feature point matching. Moreover, when the mosaic sequence increases, the previously stitched image size increases as well. The time consumption for the feature point extraction from the stitched image increases dramatically. In Table 2, the 3rd and the 4th image mosaics consume much more time than the 1st and 2nd image mosaics. To resolve the problem, Phase Correlation is applied to approximately locate the overlapping region of two adjacent images, and the SURF algorithm and feature point extraction are restricted to the overlapping region, which greatly improved the splicing time. (2) Generating and matching feature points are restricted in the overlapping region, which can avoid many wrong match points from non-overlapping regions and enable avoidance and elimination of incorrect matching point pairs. Therefore, compared to SURF algorithm, the PC-SURF algorithm can improve the feature point matching accuracy and the image blending and registration accuracy.

4 Conclusions

A quick nursery tray image mosaic algorithm the PC-SURF based on Phase Correlation and SURF was proposed and evaluated in terms of the number of matching point accuracy, time consumption, and image blending and registration accuracy. From this research, the following conclusions can be obtained:

(1) The PC-SURF algorithm can approximately locate the overlapping region of two sequential images by using Phase Correlation. The generation and matching of SURF feature points are restricted in the overlapping region, which shortens the computation time and avoids incorrect matching points. The test results show that the average feature point matching accuracy of the PC-SURF algorithm is improved by approximately 7.14%. The implementation time is almost three times faster. The RMSE of image blending in the R, G, B channels (*RMSEr*, *RMSEg*, and *RMSEb*) decreased by approximately 8.4%, 6.9%, 6.9% respectively. The RMSE of image registration *RMSEreg* decreased by 33.8%.

(2) This nursery tray mosaic technology can meet the realtime requirements of super hybrid rice seedling automated sowing machines. The PC-SURF algorithm achieves an average nursery tray mosaic speed of 6.63 s, and rice seedling automated sowing machines usually operate at a speed of 500 trays/h, which is approximately 7.2 s/tray. The results show great practical value for nursery tray sowing quantity estimation.

Acknowledgements

The authors gratefully acknowledge the financial support from the National Key Research and Development Program of China (Grant No. 2017YFD0700802), the National Natural Science Foundation of China (Grant No. 51505156), the Earmarked Fund for Modern Agro-industry Technology Research System (Grant No. CARS-01-43).

[References]

[1] Yi C C, Din S F, Gang J W. Development of an automatic pallet handling system for seeded trays. Biosystem Engineering, 2006; 93: 123–138.

[2] Zhao Z, Wu Y F, Yin J J, Tang Z. Monitoring method of rice seeds mass in vibrating tray for vacuum-panel precision seeder. Computers and Electronics in Agriculture, 2015; 114: 25–31.

[3] Tan S Y, Ma X, Wu L L, Li Z H, Liang Z W. Estimation on hole seeding quantity of super hybrid rice based on machine vision and BP neural network. Transactions of the CSAE, 2014; 30: 201–208. (in Chinese)

[4] David S, Murat K. Design and implementation of a computer vision-guided greenhouse crop diagnostics system. Machine vision and

applications, 2015; 26: 495–506.
[5] Chen H P, Shen X J, Li X F, Jin Y S. Bionic mosaic method of panoramic image based on compound eye of fly. Journal of Bionic Engineering, 2011; 8: 440–448.
[6] Xiang H T, Tian L. Method for automatic georeferencing aerial remote sensing (RS) image from an unmanned aerial vehicle (UAV) platform. Biosystems Engineering, 2011; 108: 104–113.
[7] João Valente, David Sanz, Jaime D C, Antonio. Barrientos. Near-optimal coverage trajectories for image mosaicing using a mini quad-rotor over irregular-shaped fields. Precision Agriculture, 2013; 14: 115–132.
[8] D Gómez-Candón, A I De Castro, F López-Granados. Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat. Precision Agriculture, 2014; 15: 44–56.
[9] Kise M, Zhang Q. Creating a panoramic field image using multi-spectral stereovision system. Computers and Electronics in Agriculture, 2008; 60, 67–75.
[10] Yu H F Y, Tsung C L, Ting Y L, Chang C L, Wei C C, Ta T L. An automated growth measurement system for leafy vegetables. Biosystems Engineering, 2014; 117: 43–50.
[11] Hector E, Glenn J F. Automated registration of hyperspectral images for precision agriculture. Computers and Electronics in Agriculture, 2005; 47: 103–119.
[12] Kim D H, Yoon Y I, Choi J S. An efficient method to build panoramic image mosaics. Pattern Recognition Letters, 2003; 24: 2421–2429.
[13] Tang L, Tian L F. Real-time crop row image reconstruction for automatic emerged corn plant spacing measurement. Transactions of the ASABE, 2008; 51: 1079–1087.
[14] Nakarmi A D, Tang L. Automatic inter-plant spacing sensing at early growth stages using a 3D vision sensor. Computers and Electronics in Agriculture, 2012; 82: 23–31.
[15] Bay H, Andreas E, Tinne T, Luc V G. Speeded-up robust feature (SURF). Computer vision and image understanding, 2008; 110: 346–359.
[16] Pang Y W, Li W, Yuan Y, Pan J. Fully affine invariant SURF for image matching. Neurocomputing, 2012; 85: 6–10.
[17] Ren X, Sun M, Zhang X F. A simplified method for UAV multispectral image mosaicking. Remote Sensing, 2017; 9: 962–983.
[18] Reddy B S, Chatterji B N. An FFT-based technique for translation, rotation, and scale-invariant image registration. IEEE Transactions on Image Processing, 1996; 5: 1266–1271.
[19] Chen Y, Xu M, Liu H L, Huang W N, Xing J. An improved image mosaic based on canny edge and an 18-dimensional descriptor. Optik, 2014; 125: 4745–4750.
[20] Li J, Nigel M A. A comprehensive review of current local features for computer vision. Neurocomputing, 2008; 71: 1771–1787.
[21] Yang Z L, Shen D G, Yap P T. Image mosaicking using SURF features of line segments. PloS One, 2017; 12: 1–15.
[22] Cheng L, Li M C, Liu Y X, Cai W T, Chen Y M, Yang K. Remote sensing image matching by integrating affine invariant feature extraction and RANSAC. Computers and Electrical Engineering, 2012; 38: 1023–1032.
[23] Fu Z X, Wang L M. Optimized design of automatic image mosaic. Multimedia tools and applications, 2014; 72: 503–514.
[24] Fischler M, Bolles R. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communication of the ACM, 1981; 24: 381–395.
[25] Richard P, Rafael G, László Neumann. Image blending techniques and their application in underwater mosaicking, 2014; Springer.