

Accurate 3D Surface Reconstruction for Smart Farming Application with a Shape from Focus System

¹ARPITA SWAIN, *Gandhi Institute of Excellent Technocrats, Bhubaneswar, India*

²SMRUTI CHANDAN PATI, *Gurukula Institute of Technology, Bhubaneswar, Odisha, India*

In precision agriculture, 3D vision systems are becoming increasingly important. By applying different optical 3D vision techniques, the acquired 3D data can provide information regarding the most important phenotype features in every agricultural scenario. However, most of these 3D vision systems are expensive, except some of the triangulation techniques. In this study, we focus on estimating accurate shapes using shape from focus (SFF), which is a triangulation technique. Typically, the SFF system incurs significant errors from images, including noise. As a solution to this problem, a simple low-pass filter such as the Gaussian filter has generally been used in most studies. However, when a low filter is applied, the noise is depressed but the signals are also blurred, which results in inaccuracies regarding the depth map. In this study, the noise is depressed independently without losing the original signals, and the edge components, which play important roles in finding a focused surface, are enhanced using the independent component analysis (ICA). The edge signals are amplified with a simple basis vector correction in the IC vector space. The experiments are implemented with simulated objects and real objects. The experimental results demonstrate that the obtained accuracy is comparable to that of existing methods.

1. Introduction

Precision agriculture requires sensing methodologies that provide information about individual crops and animals. During the last decade, 3D vision has become a key technology in precision agriculture for their extended capabilities compared to 2D. As 3D sensors are becoming smaller and smarter, the number of studies and applications related to agricultural 3D vision systems is increasing rapidly [1]. The 3D image generation techniques are mainly classified into three categories according to their principal measurement, namely, triangulation, time of flight, and interferometry. However, most of these 3D vision systems are of high cost owing to the expensive lasers or complicated scanning mirror systems involved, and hence, they are limited in their application, except some of triangulation techniques [2–4]. In the literature, many of the studies are aimed at improving the performance of triangulation methods in agricultural applications [4–7].

The cost can be reduced by applying passive optical methods. Shape from focus (SFF) system, a passive optical method that uses the image focus, is less expensive than other 3D solutions because of the simplicity of the system configuration and provides high accuracy. The expensive optical elements are not needed, except only one CCD camera. The purpose of SFF is to estimate the shape of objects by finding the exact focused position of each pixel from the scene. In the system, an image sequence of the scene is obtained from a fixed point of view, changing the focus consecutively with a predefined step. Then, SFF infers the 3D shape of the object from the image sequence. Many of the SFF studies have sought to improve the 3D shape in multiple ways. The research on SFF techniques is mainly divided into three categories: focus measure (FM), approximation, and optimisation techniques. FM is defined as the local sharpness measure of each pixel on the image. The initial depth map is obtained after applying the FM to the image to find the maximum response on every pixel from the image. The

sum of the modified Laplacian, Tenenbaum (TEN), and gray level variance are generally used in 3D shape recovery [8, 9]. After FM is applied, approximation techniques refine the initial depth map. Various focus curve fitting methods with planar and curved approximation for searching the exact focused image surface have been developed for better accuracy and speed. Different optimisation techniques have been applied, such as machine learning, and the hardware has been improved to correspond to the specific industrial field requirements and achieve better efficiency. Various algorithms have been used, such as neural network (SFF.NN), dynamic programming (SFF.DP), fuzzy logic (SFF.FL), and principal component analysis (SFF.PCA) [8–10].

Compared to other imaging systems, SFF has some inherent errors due to translation, magnification, and its discrete number of frames. CCD noise is the main cause of error in the SFF system. Therefore, preprocessing to reduce noise is an essential part of most SFF algorithms. Convolving the original image with a Gaussian filter is the typical way to reduce the noise effect [11]. However, this method is not effective with a severely noisy dataset because important signals are also averaged. Thus, the noise effect still remains, which results in failure to find the accurate shape of objects.

2. Materials and Methods

The independent component analysis (ICA) is a useful tool for the separation of a set of signals from the original signal. The ICA decorrelates the mixed signal y assuming that the signal is composed of an independent signal vector s and a mixing matrix A . It can be expressed by the following equation:

$$y = As; \tag{1}$$

This equation can also be expressed as follows, with the assumption of a square matrix of A :

$$s = Wy; \tag{2}$$

The goal of the ICA algorithm is to find the separation matrix W .

Preprocessing

- (1) Vector population: the input image $I(x_i, y_j, z_k)$, composed of Z frames, is converted by the vector $v(i, j, k)$ for each pixel, where $i = 0, 1, 2, 3, \dots, X, j = 0, 1, 2, 3, \dots, Y$, and $k = 0, 1, 2, 3, \dots, Z$, in each of the $X \times Y$ dimensions. The vector population is defined by Equation (3)

$$v(i, j, k) = [p(i-1, j, k) \ p(i, j-1, k) \ p(i, j, k-1) \ p(i, j, k) \ p(i, j, k+1) \ p(i+1, j, k) \ p(i, j, k+1)]^T; \tag{3}$$

Here, $v(i, j, k)$ consists of seven neighbouring pixels. Figure 1 shows $v(i, j, k)$ in the Cartesian coordinate.

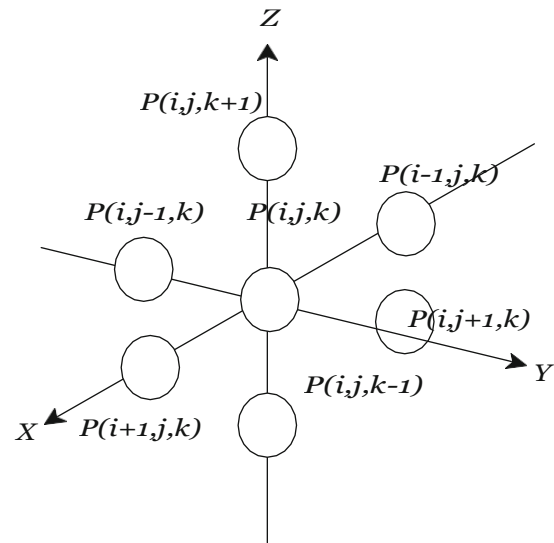


Figure 1: Vector population in the Cartesian coordinate.

- (2) Zero mean and whitening: v is normalised considering the first- and second-order statistics. The equations are as follows:

$$\begin{aligned} v &= v - \bar{v}, \\ v' &= Mv = D^{-1/2} E^T v; \end{aligned} \tag{4}$$

Here, \bar{v} is the mean of v and v' is the covariance matrix, D is the eigenvalue matrix, and E is the orthogonal eigenvector.

FastICA. To divide the independent components from the original vector, the separation matrix W is calculated. In this study, a fixed-point FastICA is used for its generality. W is induced by minimisation of mutual information. The mutual information is defined using negentropy, which is the index of non-Gaussianity. The negentropy $J(y)$ is defined by the following equation:

$$J(y) = H(y_{\text{Gauss}}) - H(y); \tag{5}$$

Here, $H(y)$ is the entropy for y , and y_{Gauss} is a Gaussian with unit variance. The minimisation of mutual information is equal to a maximisation of the negentropy. Here, the FastICA algorithm yields the solution as the following equation [12]:

$$\begin{aligned} W^+ &= E \left\{ \frac{g'(W^T x)}{g(W^T x)} \right\} W_{v'} = Mv = D^{-1/2} E^T v, \\ W^* &= \frac{W^+}{\|W^+\|}; \end{aligned} \tag{6}$$

Here, g is the first derivative of a nonquadratic nonlinear function, and E means the averaging over all column

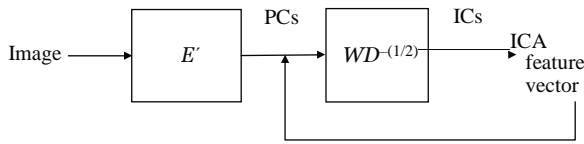


Figure 2: Transition stages of the proposed process.

vectors of matrix x . The calculation of W^* is repeated until convergence.

Denoising. Assuming that the original signal is contaminated by white Gaussian noise, the shrinkage function is used to find the noise components. Here, the probability distribution of the function is assumed as Laplacian [13]. The shrinkage function $f_{\delta s}$ is followed by

$$f_{\delta s} = \frac{1}{1 - \sigma^2} \text{sign}(\delta s) \max(0, |s| - \sigma^2), \quad \delta s$$

where σ^2 denotes the covariance of the Gaussian noise and the noise level of the image.

Edge Enhancement. Typically, the focus is on the surface of objects, that is, the edge components in the images. Thus, enhancing the edge component is one of the strategies to improve the depth map. From the independent signal s , the edge components are enhanced by eliminating the components that maximise the mutual information. The component involves common information such as the average images between two consecutive frames. Here, the components are removed by simply replacing the value of the basis vector of W to zero.

Transition to Principal Components. Finally, the signal is changed to principal components from the independent components. After the transition, the final depth map is estimated using the first principal component, which has maximal covariance. Figure 2 shows the transition stages of the proposed method. The final depth map $D_{\delta j, j}$ is calculated by the following equation:

$$D_{\delta i, j} = \max_{\delta abs} \delta f_{k1} \quad \delta \delta$$

Here, f_{k1} is the first principal component of the input signal.

3. Results and Discussion

Experimental Setup. For the experiments, a simulated cone, which was created by a virtual program, is used. The simulated cone consists of 97 frames sized 360×360 . Figure 3 shows the real objects used for the experiments, which are a real cone object composed of 30 frames sized 200×200 , a TFT-LCD cell composed of 60 frames sized 300×300 , and the Lincoln head on a US penny composed of 60 frames sized 300×300 . Figure 4 shows the simulated cone with 30 and 60 frames. It has a runtime of only 1 min with dual core Intel i3-2100 processors running at 3.1 GHz and 8 GB RAM in MATLAB.

3.2. Quantitative Analysis. For the quantitative analysis of the proposed algorithm, the root mean square error (RMSE) and correlation are calculated for each algorithm. The RMSE is an indicator of the amount of error in the images. It is denoted as

$$RMSE = \sqrt{\frac{1}{X \cdot Y} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} (f_{\delta x, y} - g_{\delta x, y})^2} \quad \delta \delta$$

where X and Y are the number of pixels in the horizontal and vertical directions in the image, respectively, $f_{\delta x, y}$ is the reference image, and $g_{\delta x, y}$ is the resulting image of the algorithm. Correlation is the similarity of the two images. It is denoted as

$$Correlation = \frac{\sum_m \sum_n (F_{mm} - \bar{F})(G_{mm} - \bar{G})}{\sqrt{\sum_m \sum_n (F_{mm} - \bar{F})^2 \sum_m \sum_n (G_{mm} - \bar{G})^2}} \quad \delta \delta$$

where F and G are the reference image and the resulting image of the algorithm, respectively, and \bar{F} and \bar{G} are their means.

In the case of experiments for synthetic object, it was possible to calculate RMSE and correlation as their true depth maps are known. For evaluating the performance of the proposed method, we experimented with three different Gaussian noise situations by changing the variance. The first situation is noise free, i.e., $\sigma^2 = 0$, the second is $\sigma^2 = 0.05$, and the third is $\sigma^2 = 0.1$. Figure 5 shows the simulated cone with additive Gaussian noise.

The traditional SFF, which simply uses a Gaussian filter, the SFF.PCA, and the proposed method were applied to remove these noise effects, recover the 3D shape, and compare it with respect to the above two matrices. Table 1 presents the robustness of the proposed algorithm compared with that of the others. In the table, it is clear that our approach significantly improved the results in comparison with the other methods. In a severe noise situation, when $\sigma^2 = 0.1$, the RMSE is improved by 30% and 38% compared to the traditional SFF and SFF.PCA, and the correlation is improved by 99% and 139% compared to the traditional SFF and SFF.PCA, respectively.

Qualitative Analysis. Figure 6 shows the restored 3D shapes of the simulated cone. The first row is the traditional SFF, the second is SFF.PCA, and the third one is the proposed method. The first column is the noise-free situation, the second column is $\sigma^2 = 0.05$, and the third one is $\sigma^2 = 0.1$. In all the cases, it is shown that the results of the proposed algorithm have a smoother surface compared to that of others. Figure 7 shows the restored 3D shapes of the real objects with noise variance of 0.05. The first column is the traditional SFF, the second is SFF.PCA, and the third one is the proposed method. In all the cases, the proposed method excels in comparison with the others. It is shown that

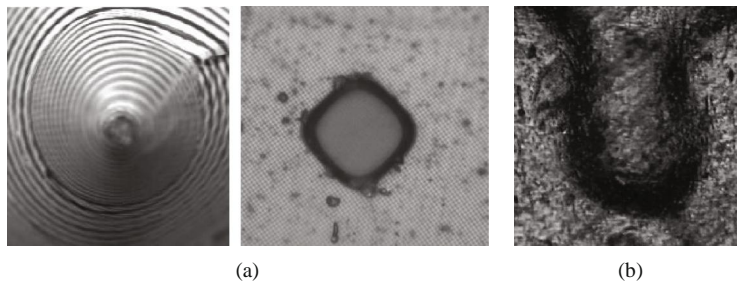


Figure 3: Real objects. (a) Real cone object and TFT-LCD cell. (b) Lincoln head on a US penny.

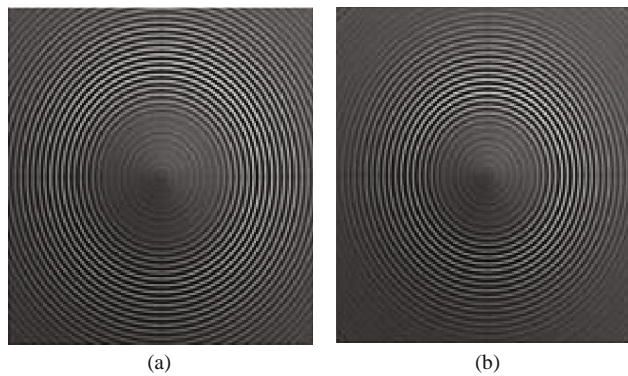


Figure 4: The simulated cone. (a) 30 frames, (b) 60 frames.

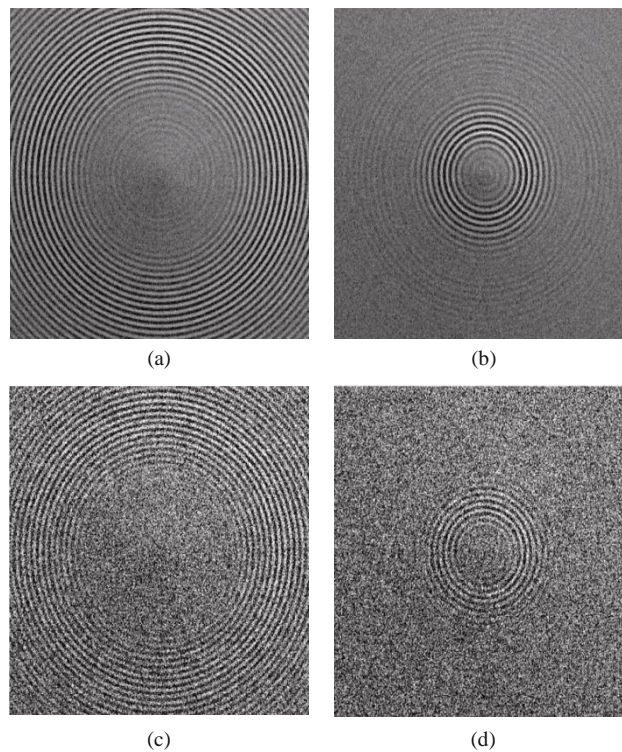


Figure 5: The simulated cone with additive Gaussian noise $\sigma^2 = 0:05$ (a, b) and $\sigma^2 = 0:1$ (c, d), with 30 frames (a, c) and 60 frames (b, d).

TABLE 1: Results of RMSE and correlation.

Method		Variance of Gaussian noise		
		0	0.05	0.1
Traditional SFF	RMSE	8.9320	15.607	35.384
	Correlation	0.8629	0.6131	0.0875
SFF.PCA	RMSE	8.1961	8.4943	31.220
	Correlation	0.9453	0.9073	0.1097
Proposed	RMSE	7.0132	8.0385	21.912
	Correlation	0.9691	0.9435	0.2625

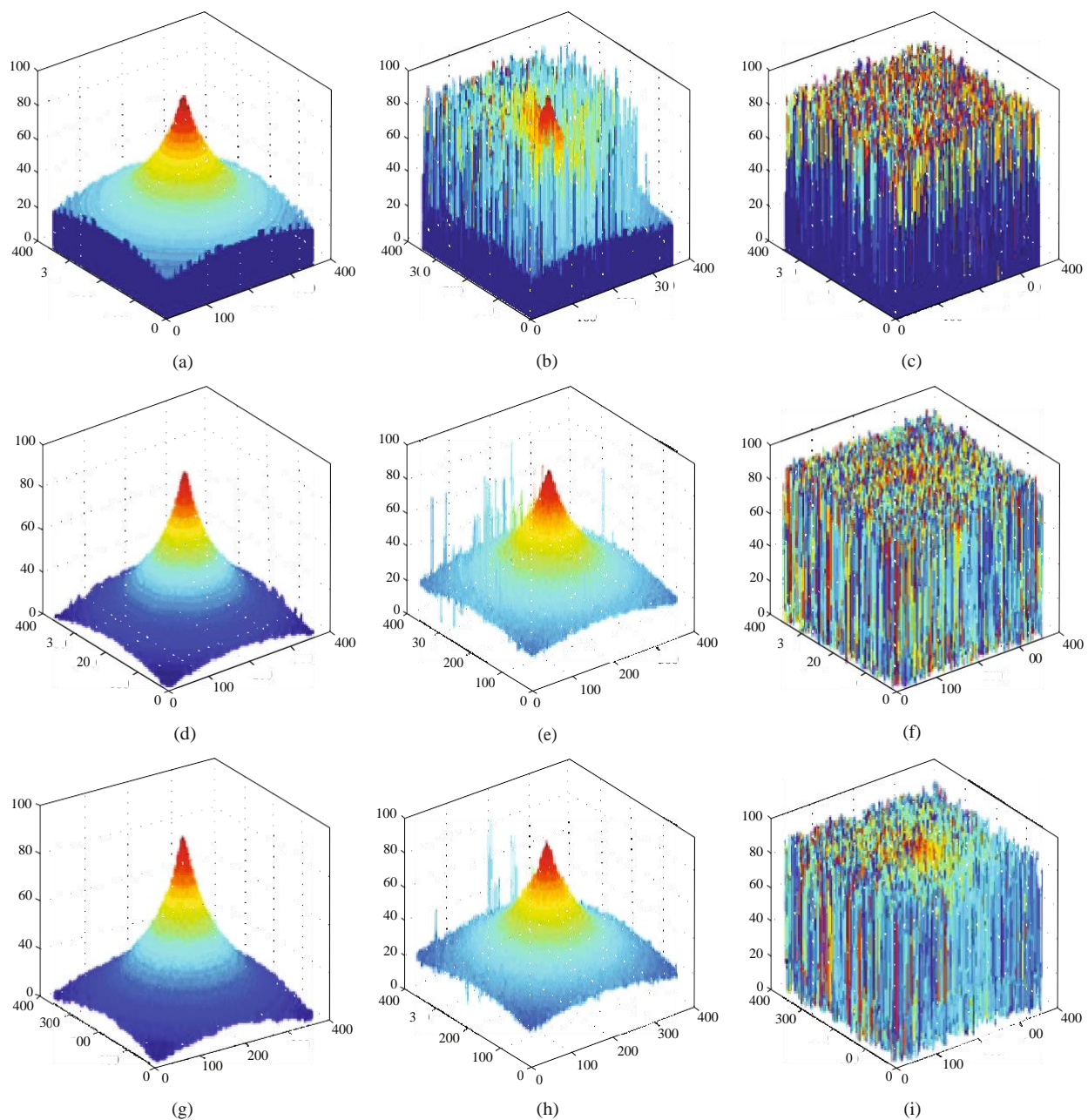


FIGURE 6: Restored 3D shape of the simulated cone. The variances are 0 (a, d, g), 0.05 (b, e, h), and 0.1 (c, f, i). Traditional method (a–c), SFF.PCA (d–f), and proposed method (g–i).

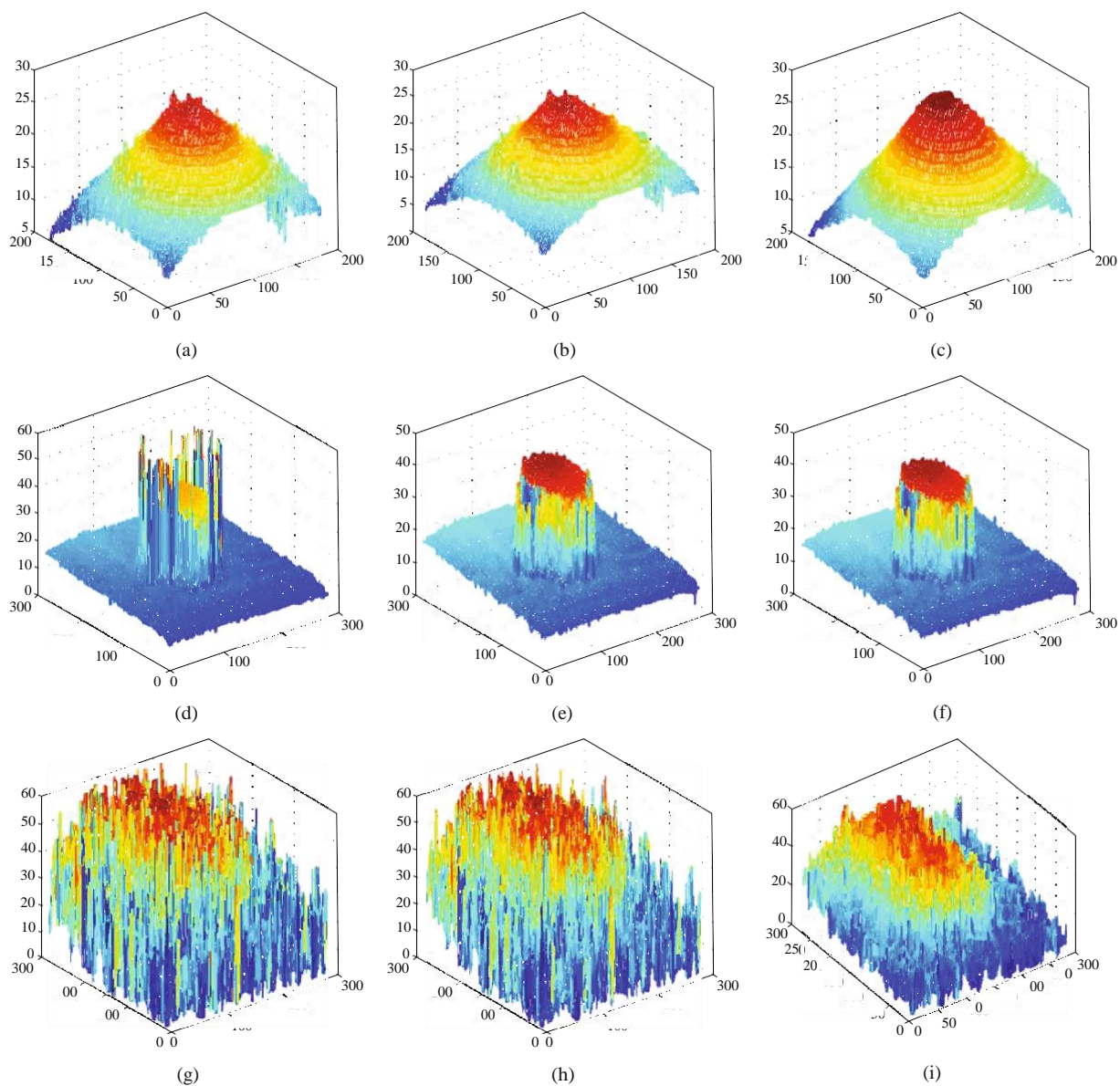


Figure 7: Restored 3D shape of real objects with noise variance 0.05. Real cone object (a–c), TFT-LCD cell (d–f), and Lincoln head on US (g–i). Traditional method (a, d, g), SFF.PCA (b, e, h), and the proposed method (c, f, i).

the results of the proposed algorithm have a smoother surface compared to that of the others.

4. Conclusions

For application in precision agriculture, a low-cost SFF system to recover accurate 3D information is developed. The conventional SFF system is weak for noise. In this study, a novel technique to improve the shape accuracy, which employs the ICA to reduce noise and enhance the edge components, has been proposed. First, the raw image passed through vector population and normalisation. Next, with the proposed approach, the signal was decomposed to independent components based on the fixed-point FastICA, and the noise components of the signal were reduced by the

shrinkage function. Then, the edge components of the signal were enhanced by removing the common component from the basis vector. Finally, the depth map was restored using the first principal component of the signal. The experiments were conducted on a simulated cone object and three real objects, using three different noise situations, thus ensuring the robustness of our algorithm. The experimental results proved that the proposed method improves the accuracy compared with the previous SFF.

References

- [1] M. Vázquez-Arellano, H. Griepentrog, D. Reiser, and D. Paraforos, "3-D imaging systems for agricultural applications—a review," *Sensors*, vol. 16, no. 5, p. 618, 2016.
- [2] F. Narvaez, G. Reina, M. Torres-Torriti, G. Kantor, and F. Auat Cheein, "A survey of ranging and imaging techniques for precision agriculture phenotyping," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 6, pp. 2428–2439, 2017.
- [3] S. Liu, L. M. Acosta-Gamboa, X. Huang, and A. Lorence, "Novel low cost 3D surface model reconstruction system for plant phenotyping," *Journal of Imaging*, vol. 3, no. 3, p. 39, 2017.
- [4] N. Fahlgren, M. Gehan, and I. Baxter, "Lights, camera, action: high-throughput plant phenotyping is ready for a close-up," *Current Opinion in Plant Biology*, vol. 24, pp. 93–99, 2015.
- [5] B. Billiot, F. Cointault, L. Journaux, J. C. Simon, and P. Gouton, "3D image acquisition system based on shape from focus technique," *Sensors*, vol. 13, no. 4, pp. 5040–5053, 2013.
- [6] D. Li, L. Xu, X. S. Tang, S. Sun, X. Cai, and P. Zhang, "3D imaging of greenhouse plants with an inexpensive binocular stereo vision system," *Remote Sensing*, vol. 9, no. 5, p. 508, 2017.
- [7] G. Polder and J. W. Hofstee, "Phenotyping large tomato plants in the greenhouse using a 3D light-field camera," in *2014 ASABE Annual International Meeting*, vol. 1, pp. 153–159, Montreal, Quebec Canada, 2014.
- [8] S. K. Nayar, "Shape from focus system," in *Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 15–18, pp. 302–308, Champaign, IL, USA, 1992.
- [9] T. S. Choi and A. S. Malik, *Vision and Shape*, Sejong Publishing Company, Korea, 2008.
- [10] M. T. Mahmood, W. J. Choi, and T. S. Choi, "PCA-based method for 3-D shape recovery of microscopic objects from image focus using discrete cosine transform," *Microscopy Research and Technique*, vol. 71, no. 12, pp. 897–907, 2008.
- [11] A. S. Malik and T. S. Choi, "Effect of noise and source illumination on 3D shape recovery," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 22, pp. 945–958, 2008.