

3D VISION-BASED DIETARY INSPECTION FOR THE CENTRAL KITCHEN AUTOMATION

Yue-Min Jiang¹, Ho-Hsin Lee¹, Cheng-Chang Lien², Chun-Feng Tai², Pi-Chun Chu² and Ting-Wei Yang²

¹Industrial Technology Research Institute, SSTC, Taiwan, ROC

²Department of CSIE, Chung Hua University, Taiwan, ROC

¹jongfat@itri.org.tw, ²cclien@chu.edu.tw

ABSTRACT

This paper proposes an intelligent and automatic dietary inspection system which can be applied to the dietary inspection for the application of central kitchen automation. The diet specifically designed for the patients are required with providing personalized diet such as low sodium intake or some necessary food. Hence, the proposed system can benefit the inspection process that is often performed manually. In the proposed system, firstly, the meal box can be detected and located automatically with the vision-based method and then all the food ingredients can be identified by using the color and LBP-HF texture features. Secondly, the quantity for each of food ingredient is estimated by using the image depth information. The experimental results show that the dietary inspection accuracy can approach 80%, dietary inspection efficiency can reach 1200ms, and the food quantity accuracy is about 90%. The proposed system is expected to increase the capacity of meal supply over 50% and be helpful to the dietician in the hospital for saving the time in the diet inspection process.

KEYWORDS

Dietary inspection, LBP-HF, Image depth

1. INTRODUCTION

In recent years, the food industry has been addressing the research on the food quality inspection for reducing the manpower and manual inspection error. To aim at this goal, in this study, the machine learning technologies are applied to develop the 3D vision-based inspection system [1,13] that can identify the meal categories and amount. In [2], the study indicated that the selected image features are crucial [14] to the detection of peel defects. In [3], the authors developed a vision-based method to improve the quality inspection of food products. In [4], Matsuda et al. proposed the food identification method by integrating several detectors and image features, e.g., color, gradient, texture, and SIFT features. Then, the multiple kernel learning (MKL) method is applied to identify the food quality. Yang et al. [5] proposed the pair wise local features to describe the texture distributions for eight basic food ingredients. However, the abovementioned methods do not address the quality inspection for the Chinese foods. In the Chinese food, several food ingredients are often mixed, e.g., the scrambled eggs with tomatoes, such that it is difficult to identify the food ingredients and quantity by using the conventional vision-based methods. In [8], Chen et al. proposed the diet ingredients inspection method by using the SIFT, Gabor texture, and depth camera to detect the diet ingredients. Based this method,

in this study, we apply the proposed the meal box detection and locating technology, LBP-HFtexture features, and depth images to construct a novel approach of the dietary inspection for the central kitchen automation. The system flowchart of the proposed 3D vision-based dietary inspection system is shown in Figure 1.

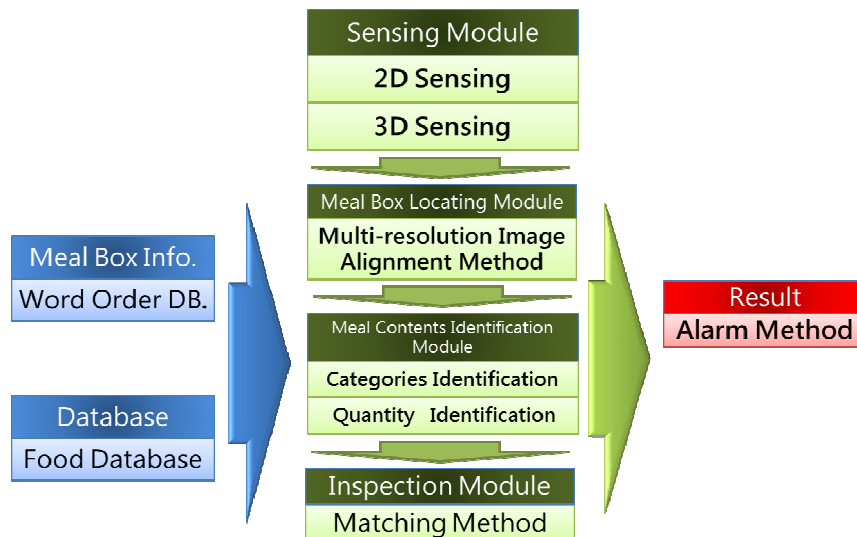


Figure 1. The system flowchart of the proposed 3D vision-based dietary inspection system.

In Fig. 1, firstly, the sensing module extracts 3D (depth) and 2D images. Secondly, the meal box locating module locates the position of the detected meal box and segment the regions for each food ingredient. Finally, the meal contents identification module identifies the food categories and amount for evaluate the food quality. The system operation procedures are described in Figure 2. The meal box is moving on the conveyor and the sensing module extracts 3D (depth) and 2D images continuously. Once the meal box is located with the meal box locating module, the food quality can be inspected with the color, texture, and depth image features. The experimental results show that the dietary inspection accuracy can approach 80%, dietary inspection efficiency can reach 1200 ms, and the food quantity accuracy is about 90%. The proposed system is expected to increase the capacity of meal supply over 50% and be helpful to the dietician in the hospital for saving the time in the diet inspection process.

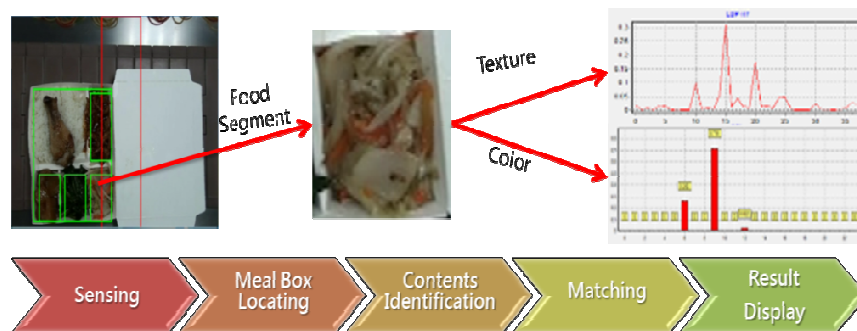


Figure 2. The system operation procedures of the proposed 3D vision-based dietary inspection system.

2. AUTOMATIC MEAL BOX INSPECTION

The baseline of the system design is based on the domain knowledge of dietician. In this section, the methodologies of the meal box locating and the meal contents identification are described. In the diet content and amount identification process, the 2D/3D image features, e.g., depth, color [7] and textures[6], are used to train the dietary inspection system, and then the system can identify the diet categories and amounts. By using the novel automatically foods recognition and amount identification system, the manual operations can be reduced significantly and the accuracy and efficiency of food arrangement can be improved significantly.

2.1. Meal Box Detection and Locating with Multi-resolution Image Alignment

To develop a real-time vision-based dietary inspection system, the analyses of the video content captured from the camera are crucial to detect and locate the meal box. In Fig. 3, we can see that the meal box is moving continuously on the meal dispatch conveyor at central kitchen. Then, how to detect the meal box and locate the position of meal box in real-time is a problem. Here, we proposed a novel meal box locating method by using the multi-resolution image alignment method to match the meal box template shown in Fig. 4-(b) to the captured images from low resolution to high resolution within the region of interest (ROI) shown in Fig. 3.

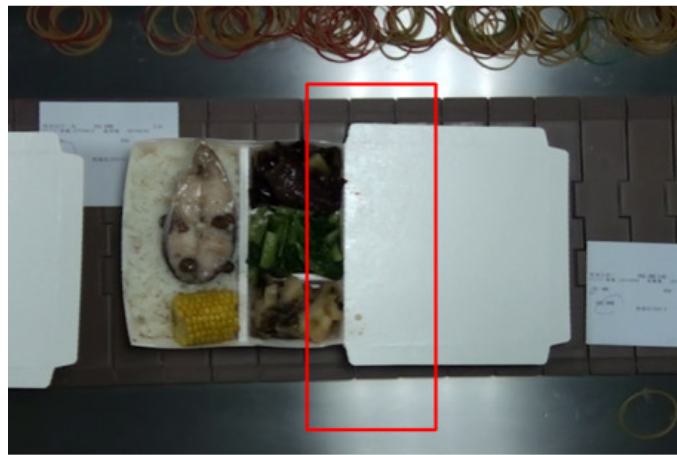


Figure3.The ROI setting in the meal box image.

Based on the careful observations, the image template of the meal box is difficult to generate because that the foods can cover the left side of meal box and no texture information exist on the right side of meal box. Hence, we extract the middle image in the meal box image shown in Fig. 4-(b) as the image template of meal box to detect and locate the position of meal box.

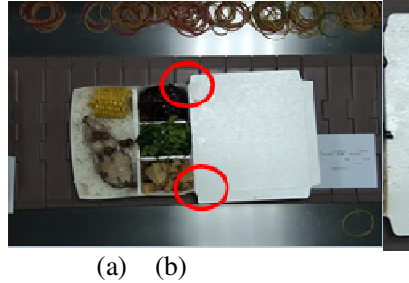


Figure 4. Selection of image template of meal box to detect and locate the position of meal box. (a) Meal box image on the meal dispatch conveyor. (b) Image template of meal box.

The algorithm of meal box locating is described as follows.

1. Image template and meal box image are decomposed into specified multi-resolution levels (pyramid image representation) shown in Fig. 5.
2. Perform the pixel-based template matching (correlation matching) in the lower level (lower resolution). Then, some candidate regions are extracted.
3. Perform the pixel-based template matching in the higher resolution images within the neighbouring region obtained from the candidate regions in step 2.

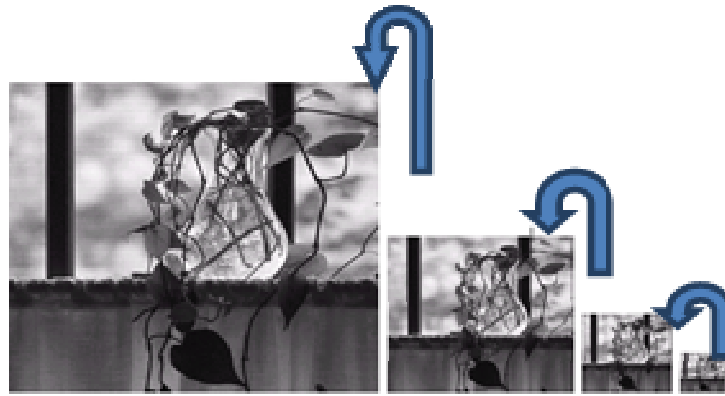


Figure 5. Image are decomposed into multi resolution levels (pyramid image representation).

To speed up the calculation efficiency, all the integer multiplication operations defined in Eq. (1) for the computation of correlation are calculated in advance and stored as a look-up table.

$$\text{GrayTable}(A, B) = \{(A - \bar{A}) \times (B - \bar{B}), 0 \leq A \leq 255, 0 \leq B \leq 255\} \quad (1)$$

where \bar{A} and \bar{B} are the mean values of the image A and B respectively. The system operation flowchart of the multi-resolution meal box locating module is shown in Figure 6.

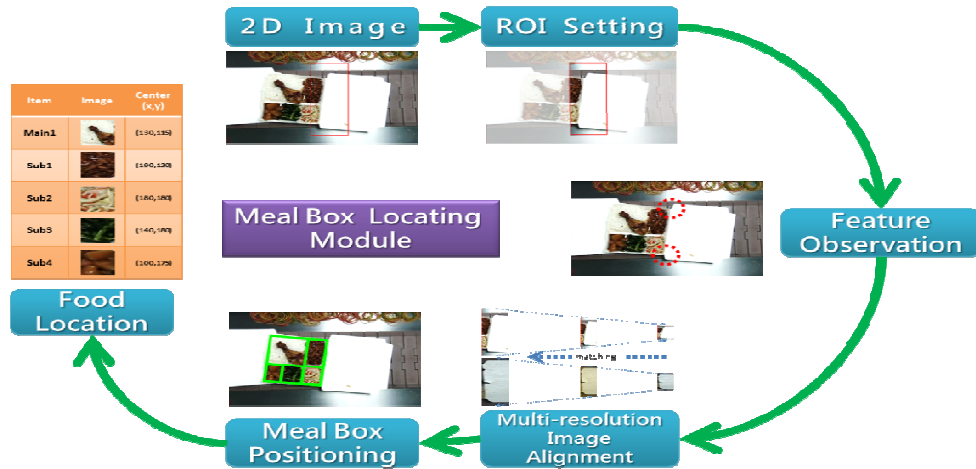


Figure6.The system operation flowchart of the multi resolution meal box locating module.

2.2. Meal Box Contents Identification

This section will describe the identification methods for the contents in the meal box, i.e., the processes in the "meal box location content identification module". The extracted features include the color distribution (polor histogram)and texture(LBP-HF) feature within the ROI. Finally, the dietary inspection is designed with the similarity measure between the online input image and the trained image features in the database. Figure 7 illustrates the flowchart of food quality inspection system.



Figure7.Flowchart of food quality inspection.

2.2.1.Color Polar Histogram

Once the meal box is aligned, we can segment the regions for each food ingredient to extract the color distribution feature for identifying the food ingredient color. Here, we transfer the color space of the image of each food ingredient into YCbCr color space and use the CbCr color channels to establish the color polar histogram [7,9] with angle range from -127° to 127° . Figure 8 shows the polar coordinates.

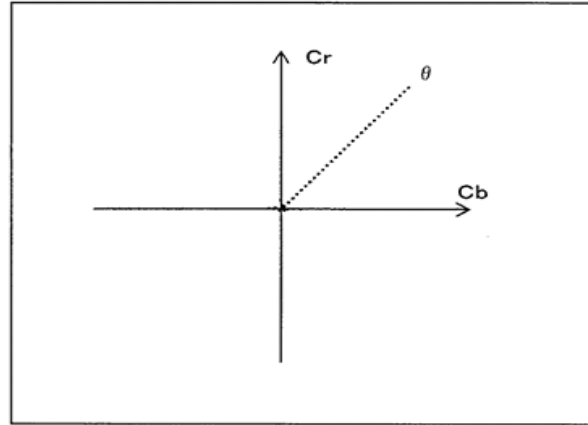


Figure 8. Polar coordinates for Cr and Cb color space.

$$h_l = \sum_{i=1}^N \delta[\theta(C_b, C_r) - l] \quad (2)$$

Here, the color polar histogram is represented as $H = \{h_1, \dots, h_m\}$. The value for each histogram bin can be calculated as the formula in Eq. (2), where the θ is the degree of coordinate. The extracted color polar histogram feature for the image of the food ingredient of the sample image is illustrated in Fig. 9.

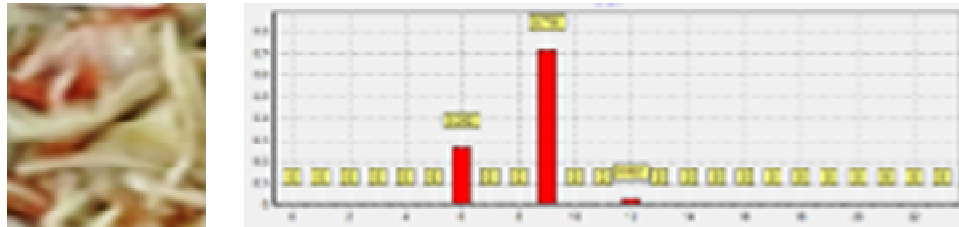


Figure 9. The extracted color polar histogram feature for the image of the food ingredient of the sample image.

2.2.2 Local Binary Pattern-Histogram Fourier (LBP-HF)

Local binary pattern-histogram Fourier (LBP-HF)[10] is based on the LBP method for rotation-invariant. The LBP operator is powerful for texture description. It labels the image pixels by thresholding the surrounding pixels with comparing the value of center pixel and summing the thresholded values weighted by powers of two. The LBPlabel can be obtained with the formula in Eq. (3).

$$LBP_{P,R}(x, y) = \sum_{p=0}^{P-1} s(f(x, y) - f(x_p, y_p)) 2^p \quad (3)$$

where $f(x, y)$ is the center pixel (red dot) of image f shown as Figure 10. P is the number of surrounding points, R is sampling radius, and $s(z)$ is the thresholding function shown as Eq.(4).

$$s(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases} \quad (4)$$

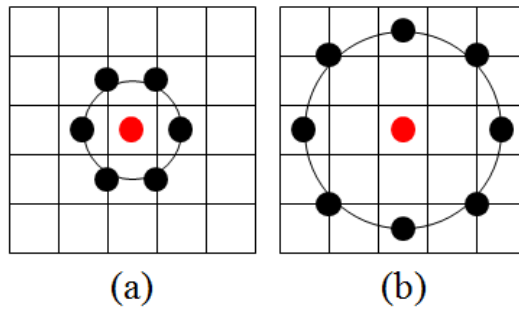


Figure 10. LBP sampling radius. (a) (P, R) = (6, 1). (b) (P, R) = (8, 2).

Furthermore, an extended LBP operator called uniform LBP [11] is proposed to describe the region texture distribution more precisely. The uniform LBP operator is constructed by considering if the binary pattern contains at most two bitwise transitions from 0 to 1 or 1 to 0 when the bit pattern is considered circular[6]. For computing the uniform LBP histogram, each uniform pattern is assigned to a specified bin and all non-uniform patterns are assigned into a single bin. The 58 possible uniform patterns (all zeros, all ones, non-uniform) of (P, R) = (8, R) are shown in Figure 11.

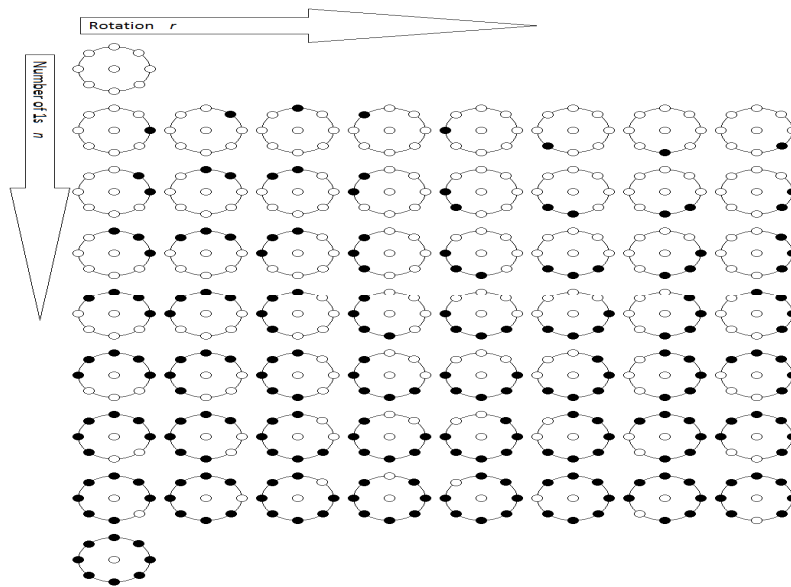


Figure 11. 58 possible uniform patterns of (P, R) = (8, R).

The uniform LBP owns a rotation invariant property. The rotation of uniform LBP is just as a horizontal shift in Figure 11 and shown in Figure 12. Based on this property, the LBP-HF [6] image feature is proposed. The LBP-HF image feature is generated by performing Fourier transform to every row in the uniform LBP histogram (except the first and the last row) to Discrete Fourier Transform to construct these features, and let $H(n, \cdot)$ be the DFT of n -th row of the histogram $h_l(U_P(n, r))$, which is shown as Eq. (5).

$$H(n, u) = \sum_{r=0}^{P-1} h_l(U_P(n, r)) e^{-i2\pi ur/P} \tag{5}$$

In Eq. (5), $H(n, u)$ is the Fourier transformed histogram, n is the number of “ I ”, u is the frequency, h_l is the uniform LBP histogram of the image I , $U_P(n, r)$ is the uniform LBP operator, and r denotes the row index. We apply the feature vectors consisting of three LBP histogram values (all zeros, all ones, non-uniform) and Fourier magnitude spectrum values of LBP-HF in Eq. (6) to describe the texture distribution of the food ingredient image.

$$fv_{\text{LBP-HF}} = [|H(1,0)|, \dots, |H(1, \frac{P}{2})|, \dots, |H(P-1,0)|, \dots, |H(P-1, \frac{P}{2})|, h(U_P(0,0)), h(U_P(P,0)), h(U_P(P+1,0))] \quad (6)$$

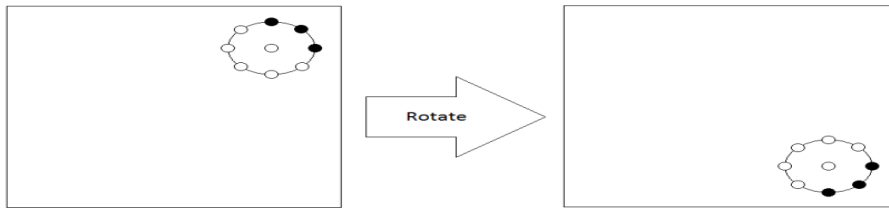


Figure12. Rotation doesn't change the row it belongs to in Figure11

2.2.3.Data Fusion and Matching

In this study, we utilize Bhattacharyya distance[12]to measure the similarity between the trained and test patterns that are described with the LBP-HF texture description and polar color histogram. Bhattacharyya distance measurement shown in Eq. (7) can be used to measure the similarity between two discrete probability distributions.

$$d_B(y) = \frac{1}{S} \sum_{b=1}^S \sqrt{1 - \rho_B[H_b, P_b(y)]}, \quad (7)$$

where, $\rho_B[H_b, P_b(y)] = \sum_{i=1}^m \sqrt{\frac{h_i \cdot p_i(y)}{\sum_{i=1}^m h_i \cdot \sum_{i=1}^m p_i(y)}}$.

2.3.Food Quantity Measurement

For the inspection of amount of food ingredient, we use depth information obtained from the depth sensor to evaluate the amount of each food ingredient. Figure 13 illustrates the captured depth information used to determine the amount of food ingredient.

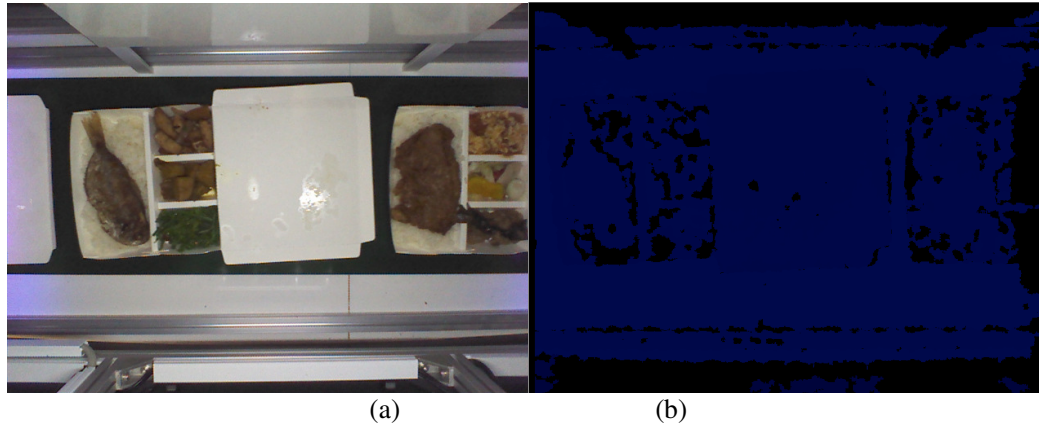


Figure 13. The captured depth information used to determine the amount of food ingredient. (a) Meal box image. (b) Depth image for (a).

3. EXPERIMENTAL RESULTS

In this section, we apply the automatic meal box detection/locating module and automatic food ingredients identification module to construct a food quality inspection system. The operation scenario is shown in Figure 14.

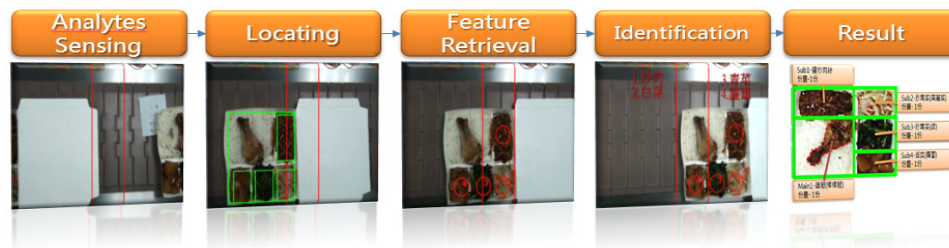


Figure 14. The system operation scenario.

The proposed food quality inspection system is implemented on the meal box dispatch conveyor of the Chinese food central kitchen. It automatically check compliance of the meal box content between customer-made meals orders of the dietician designed. This automated inspection system's operation module used Intel i5 2.2GHz CPU to analysis the contents of meal box, and it used the Microsoft Kinect camera in capture module.

The performance of food quality inspection is evaluated with two different meal boxes that are three and four food ingredients' partitions. Figure 15 shows two different meal boxes. There are 9 dishes types in the meal box including the one main dishes and 3 or 4 vice-dishes, which are shown in Figure 15. The efficiency of the meal box location detection and locating module and food ingredients identification module is listed in Table 1.



Figure 15.(a) 4 vice-dish meal box. (b) 3 vice-dish meal box.

Table 1.The efficiency analysis of the meal box detection and locating.

Test Video	Meal box location detect module(1 box)	food ingredients identification module(1 box)	AVG. time for one meal box
Video 1 (4 vice-dish)	10.16 ms	116.44ms	126.60 ms
Video 2 (3 vice-dish)	9.8 ms	95 ms	114.8 ms.

Table 2 illustrates the accuracy analysis of the proposed food quality inspection system. The accuracy of meal box detection and locating is higher than 85% and the accuracy of food ingredients identification can approach 85%.

Table 2.Accuracy analysis for the food quality inspection system.

Test Video	Meal box location detect module	food ingredients identification module
Video 1 (4 vice-dish)	85.3 %	82.1 %
Video 2 (3 vice-dish)	89.6 %	89.3 %

For the amount inspection of each food ingredient, we apply the depth information to evaluate the amount of each food ingredient. The efficiencies for the meal box location detection module, food ingredients identification, and quantity estimated module are shown in Table 3.The complete average processing time of each meal box is about 1.4 second. In Table 4, the accuracy analysis is listed.

Table 3.Detection efficiency of the automated optical inspection system

Test Video	Meal box location detect module(1 box)	food ingredients identification module(1 box)	food quantity estimated module(1 box)	AVG. time for one meal box
Video 1	21.36 ms	94.1 ms	25.2ms	140.66 ms

Table 4.Detection accuracy of the automated optical inspection system

Test Video	Meal box location detect module	food ingredients identification module	food quantity estimated module
Video 1	85.3 %	82.1 %	74.2%

4. CONCLUSIONS

In the proposed system, firstly, the meal box can be located automatically with the vision-based method and then all the food ingredients can be identified by using the colour and LBP-HF texture features. Secondly, the quantity for each of food ingredient is estimated by using the image depth information. The experimental results show that the dietary inspection accuracy can approach 80%, dietary inspection efficiency can reach 1200ms, and the food quantity accuracy is about 90%. The proposed system is expected to increase the capacity of meal supply over 50% and be helpful to the dietician in the hospital for saving the time in the diet identification process.

REFERENCES

- [1] Chetima, M.M. &Payeur, p. (2008) "Feature selection for a real-time vision-based food inspection system", Proc. of the IEEE Intl Workshop on Robotic and Sensors Environments, pp 120-125.
- [2] Blasco, J. Aleixos, N. Molto, E. (2007) "Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm", Journal of Food Engineering, Vol. 81, No. 3, pp535-543.
- [3] Brosnan, T. & Sun, D.-W.(2004) "Improving quality inspection of food products by computer vision – a review", Journal of Food Engineering, Vol. 61, pp. 3-16.
- [4] Matsuda, Y. & Hoashi, H. (2012) "Recognition of multiple-food images by detecting candidate regions." In Proc. of IEEE International Conference on Multimedia and Expo, pp 1554–1564.
- [5] Yang, S. Chen, M. Pomerleau, D. Sukhankar. (2010)"Food recognition using statistics of pairwise local features."International Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, pp. 2249–2256
- [6] Zhao, G. Ahonen, T. Matas, J. Pietikainen, M. (2012) "Rotation-invariant image and video description with local binary pattern features," IEEE Trans. Image Processing, Vol. 21, No. 4, pp. 1465–1477.
- [7] Suau, P., Rizo, R. Pujol, M. (2004) "Image Recognition Using Polar Histograms." <http://www.dccia.ua.es/~pablo/papers/ukrobraz2004.pdf>
- [8] Chen, M.Y., Yang, Y.H., Ho, C.J., Wang, S.H., Liu, S.M., Chang, E., Yeh, C.H., Ouhyoung, M. (2012) "Automatic Chinese food identification and quantity estimation." SIGGRAPH Asia 2012 Technical Briefs
- [9] Chang, P., Krumm, J.(1999) "Object Recognition with Color Cooccurrence Histograms", IEEE International Conference on Computer Vision and Pattern Recognition,
- [10] Ren, J., Jiang, X., Yuan, J., Wang, G., (2014)"Optimizing LBP Structure For Visual Recognition Using Binary Quadratic Programming", Signal Processing Letters, IEEE, On page(s): 1346 - 1350 Volume: 21, Issue: 11.
- [11] Ojala, T., Pietikainen, M., Mäenpää, T.(2002)"Multi resolution gray-scale and rotation invariant texture classification with local binary patterns." IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 971–987
- [12] F. Aherne, N. Thacker and P. Rockett,(1998)"The Bhattacharyya Metric as an Absolute Similarity Measure for Frequency Coded Data," Kybernetika, vol. 34, no. 4, pp. 363-368.
- [13] M.R. Chandraratne, D. Kulasiri, and S. Samarasinghe,(2007) "Classification of Lamb Carcass Using Machine Vision: Comparison of Statistical and Neural Network Analyses", Journal of Food Engineering, vol. 82, no. 1, pp. 26-34.
- [14] Aleixos, N., Blasco, J., &Moltó, E. (1999). "Design of a vision system for real-time inspection of oranges."In VIII national symposium on pattern recognition and image analysis. Bilbao, Spain.pp. 387–394