Australian Journal of Crop Science AJCS
AJCS 7(5):699-705 (2013) ISSN:1835-2707

Development of an expert system based on wavelet transform and artificial neural networks for the ripe tomato harvesting robot

Arman Arefi* **and Asad Modarres Motlagh**

Department of Agricultural machinery Engineering, Agricultural Faculty, Urmia university, P.O.box 57155-1849, Urmia, Iran

***Corresponding author: arefi.arman@yahoo.com**

Abstract

Robots can reduce production costs in large-scale greenhouses. For automating the harvest operation, robot should be able to recognize the fruit. This study was conducted with a view to develop a robotic harvesting system that performs recognition of ripe tomato. For this aim, several color images of tomato plants were taken in RGB color model. After analyzing data colors, detection of ripe tomato was done in two steps, first removing background of image and then recognizing ripe tomato from unripe tomato. Removing background was done based on threshold method using R-G equation. The identification of ripe tomato from unripe tomato was performed by integrating image processing and artificial neural network. Image processing was done based on wavelet technique. Energy and correlation were defined as two wavelet features and totally 90 wavelet features were extracted from each tomato. Artificial neural network was used for classification of ripe tomato from unripe tomato. A feed-forward neural network with two hidden layers was developed. The 63% of samples were used in the training stage of neural network, 12% of samples for validation, and testing of the network was performed with 25% of samples. The proposed algorithm could classify ripe tomato and unripe tomato with acceptable accuracies of 95.45% and 90% respectively.

Keywords: tomato, robot, image processing, wavelet, greenhouse.
Abbreviations: ANN- artificial neural network; CWT- continuous wavelet transform; DWT- discrete wavelet transform; FFT- fast Fourier transform; MLP- multi-layer perception RT- ripe tomato; UT- unripe tomato.

Introduction

Tomato is considered as one of the main agricultural crops in the world due to its high nutritive value, and consumed daily by millions of people from diverse cultural backgrounds. It contains lots of vitamins and nutrients that vitamin C is the best known of them. Due to its important nutrition, the worldwide production of tomato has been increased in the recent years. As its global production in 2009 was more than 150 million tons, its significance has increased in comparison with preceding decades. Iran has been ranked 7th in the world in 2009 (FAO, 2009). Existence of greenhouses is one of the main reasons of increasing the production of tomato because greenhouses allow tomato plants to be grown in different weather conditions and to continue producing crops throughout the year. The harvesting by workers is one of the major obstacles in increasing fruit production. Sarig's studies in 2005 showed that the cost of harvesting by labors is expensive and time-consuming. In addition, picking fruits by hand is very tedious. To overcome these problems, robots seem suitable instruments to automate harvesting operations. Harvesting costs can be reduced by robots, so use of image processing techniques for fruit automatic harvesting has gained interest in recent years (Wang et al., 2008). Over the past years, many studies have been carried out for automatic harvesting of fruits, such as orange and citrus (Hannan et al., 2009; Okamoto and Lee, 2009), strawberry (Feng et al., 2008), apple (Bulanon et al., 2004; Beaten et al., 2007), cherry (Tanigaki et al., 2008), eggplant (Hayashi et al., 2002) and mushroom (Reed et al., 2001). A machine vision was developed to recognize the red strawberries in greenhouse. For this aim, an algorithm was written in Nrgb color model. If strawberries satisfied equation 0.58<Nr<1, they would be considered as red strawberries (Hayashi et al., 2005). A robotic harvesting system was developed for the automatic recognition of 'Fuji' apples on the tree (Bulanon et al., 2002). The color properties used in the model were luminance (Y), red color difference (C_r), green color difference (C_g) and blue color difference (C_b). Since the color of 'Fuji' apple was red, the red color difference (C_r) was used. Results of the segmentation showed a success rate of over 88%. In 2008, recognition of cotton was performed by Wang et al. They defined R-B equation in RGB color model that could successfully detect cotton from the background. Since there are UTs with RTs on a plant, it is necessary to develop an algorithm that recognizes RTs and keep UTs for later. Until now, several studies for recognition of RT have been done. Hayashi and Sakaue (1996) proposed an algorithm based on color features for automatic recognition of tomato. The defined algorithm was not able to recognize tomatoes one by one and considered several tomatoes as a big tomato. Later another system equipped to a stereoscopic vision system was developed that could recognize tomatoes individually (Hayashi et al., 2005). Although in Hayashi's study, recognizing tomatoes from the background was successfully done, but it did not mention the recognition of red tomatoes from UTs. In case of cluster harvesting of tomato, there is a research that was performed by Kondo et al. (2009). They introduced a machine vision system based on color in HSI model for autonomous tomato harvesting. The system had to detect main stem and peduncle. Since peduncle was not clearly isolated from other plant parts, identifying peduncle was not easy and sometimes the system encountered difficulty. All developed algorithms for recognizing RT had been defined based on color features. Recently, a new technique named wavelet is gaining popularity in image processing as an effective and useful technique. For the first time, ability of wavelet technique in detecting the RT was evaluated in this paper. After removing background, the developed algorithm first separated the glued tomatoes and then extracted wavelet features in two color spaces RGB and HSI. Finally, extracted features were defined as input ANN.

Results and discussions

Removing the background

One of the easiest, fastest and the most efficient methods in image processing is image segmentation based on threshold. In this method high contrast between objects in an image is the most important agent (Zheng and Sun, 2006). A high contrast between tomatoes and background could be produced by the D_{rg} and D_{rb} equations after considering color data of objects. Although the D_{rb} equation could be acceptable, comparison between D_{rg} and D_{rb} equations signified that D_{rg} equation was more effective because of producing higher contrast (Fig 5). Finally, D_{rg} equation was utilized to remove the background. For presenting more details, an image sample and its histogram corresponding to the D_{rg} is shown in Fig 1. It can be seen that there is a deep valley in histogram that has successfully separated the tomato with more intensity value in the right side from the background in the other side. Reflecting the light by fruit is the most common problem in harvesting and sorting automatic operation reported by other researchers (Hayashi et al., 2002; Bulanon et al., 2004). Those parts of tomato that exposed to direct sun light were not extracted. Because the color of those parts was close to white color (intensity value R, G and B components are 255 in white color) and the difference between R and G components was close to zero.

Separation of touching tomatoes

Proposed algorithm showed acceptable results separating touching tomatoes. After separation of tomatoes, location of each individual tomato can be measured. An end-effector was developed for tomato harvesting robot by Monta et al. (1998). It is thought that the proposed algorithm can work well with that end-effector.

Sensitivity analysis

The results of sensitivity analysis obtained for 90 primary features are shown in Fig 2. During the training and validation stage, output values were examined and compared with desired ones to determine sensitivity value. According to the results, total features of RT had higher sensitivity index than the UT. It's concluded that the classifier would easily recognize RT than UT. In order to reduce the size of input vector of ANN, value 0.15 was determined as sensitivity threshold. In this sensitivity level, only those features were selected which their sensitivity indexes were higher than the sensitivity threshold of RT and UT. The bigger the sensitivity level, the lower the number of features. The results showed that classification accuracy decreased when sensitivity threshold was higher than 0.15. According to the Fig 2, correlation features had higher sensitivity index than the energy ones and this shows that correlation features carry more information than energy features. The 16 top wavelet features are represented in Table 1.

Table 1. The top wavelet features based on their effect in classification accuracy of tomato.

Wavelet features	
Energy	Correlation
V’B’3	C’RG’1
D’G’3	C’HS’1
D’S’3	C’HI’1
C’S’1	C’SI’1
C’S’2	C’RG’2
	C’BG’2
	C’HS’2
	C’RG’3
	C’BG’3
	C’HS’3
	C’SI’3

Where the first index C, H, V and D show rotationally invariants, horizontal, vertical and diagonal respectively. Second index show band color (R, G, B, H, S and I show red, green, blue, hue, saturation and intensity, respectively). Numbers 1, 2, and 3show the 1^{th}, 2^{th} and 3^{th} level of decomposition.

Table 2. Confusion matrix and classification accuracy of ANN models

Output/Desired	RT	UT
Ripe Tomato	21	3
Unripe Tomato	1	27
Accuracy (%)	95.45	90.00

Determination of ANN topology

The number of hidden layers and the number of neurons in hidden layers are considered as two important factors that help to design the best topology for an MLP network. Generalization ability of the MLP depends on the number of hidden layers. It is reported that MLP may not have a good accuracy when the number of hidden layers are not enough (too few). On the other hand, if the number of hidden layers is too many, over fitting may occur. Therefore, determining the optimal number of hidden layers is so necessary. For this aim, trial-and-error is known as an accepted method, which has been suggested by several researchers. In the current research, when the number of hidden layers was increased from one to two layers, the MLP’s performance was increased and classification accuracy decreased when the number of hidden layers was increased to three layers. Therefore, ANN architecture with two hidden layers was selected for further investigation. The best number of neurons in the hidden layer was determined based on trial-and-error method too.

Classification accuracy

The wavelet as a new technique was used for recognition of RT and UT. Energy and correlation as two common features were extracted and considered. Results showed that energy of RT tomato was more than UT tomato (Fig 3). This shows that an UT had more color uniformity than a RT. This can be due to gradual ripping of tomatoes because there are yellow-red pixels in a RT. So it can be concluded that energy and correlation are two effective features of recognizing RT from UT. In addition, it observed that energy and correlation values of images increased when division of an image was increased from level1 to level 3 (Fig 3). Because when an image is divided to higher levels, the values with higher frequency (higher variance value) are selected. The higher the frequency value, the higher the energy value. As it can be observed from Fig 3, the difference between energy value of

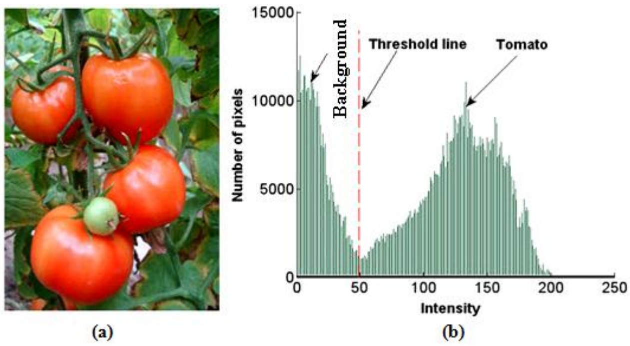

Fig 1. The corresponding Histogram to Drg relation: **a** original image. **b** corresponding histogram, it can be observed that threshold line has successfully separated tomatoes (placed in the right of the valley) from the background (placed in the left of the valley) and so this issue ensure removing the background.

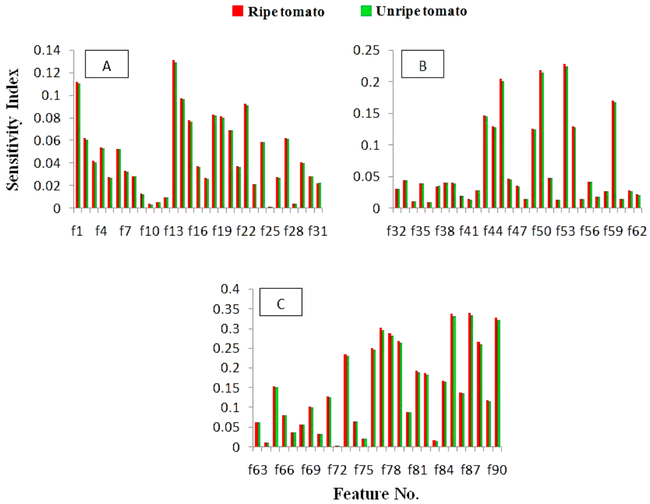

Fig 2. Sensitivity analysis for extracted features of tomato images: A, B and C show sensitivity value for features 1 to 90. Those features had sensitivity value more than 0.15 were defined as ANN inputs

RT and UT increased when level of divisions was increased. So, sub-division of image could help to better recognition RT from UT. The wavelet features gave the acceptable classification accuracies for RT and UT. Table 2 shows the obtained confusion matrix and accuracy of ANN. The classification accuracy of UT was lower as compared to RT. This is perhaps because UT included wide range of colors from yellow color to light red color. Another reason can be due to that images were acquired without any artificial lighting system. In parts of the greenhouse where light was enough, images were prepared with good quality and UT was detected more easily. On the other hand, in the parts of greenhouse where light was not enough, images were prepared with inappropriate quality and the color of UT was darker and was incorrectly recognized as RT. It is guessed that accurate classification of UT tomato can be increased by using an artificial lighting system. Obtaining better results can be attributed to the number of color models too. Since only two color models, the RGB and HIS, were used, it is guessed that several color models may be better. Requiring a square box around the objects can be another obstacle in increasing performance of wavelet features because the square box added extra pixels (Choudhary et al., 2008).

Materials and methods

Outline of developed algorithm

The developed algorithm for detection of RT consisted of seven image processing steps: a) collecting data, b) removing background, c) removing noises, d) separation of touching tomatoes e) feature extraction, f) feature selection and g) intelligence classification. This process has been shown in Fig 4 and has been brought below.

Collecting data

The experimental setup was composed of a 3 charged coupled device (3CCD) camera (Sony Cyber Shot w200, Japan, resolution 1944×2592 pixels) and a personal computer with 2.20 GHz processor and 2.00GB RAM. A total of 200 images of the tomatoes were taken as samples from greenhouses in natural conditions of greenhouse without any artificial lighting system. The distance of camera from tomatoes was about 20 cm. There were various objects in a color image that it was necessary to define them before anything. The objects were classified into three categories including background, UT and RT. Background included all green tomatoes, branches, leaves and greenhouse space. American standards were used to define RT and UT (Code of Federal Regulations United States standards for grades of fresh tomatoes, 1991). According to this standard, RT is those tomatoes that more than 90% of their surface includes red color, and the other color tomatoes such as light red (more than 60% but not more than 90% of the surface shows pinkish red or red color), pink (more than 30% but not more than 60% of the surface shows pink or red color), turning (more than 10% but not more than 30% of the surface shows a definite change in color from green to yellow, pink or red color.) and breakers (there is a definite break in color from

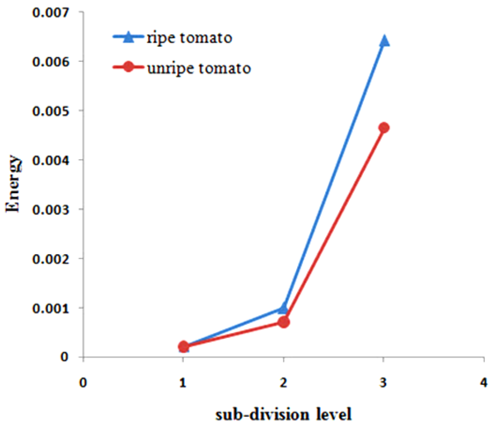

Fig 3. Comparing energy value of ripe and unripe tomato at three sub-divisions levels: the difference between the extracted energy for ripe and unripe tomato increased with increasing division of image so that the most difference was obtained for level 3.

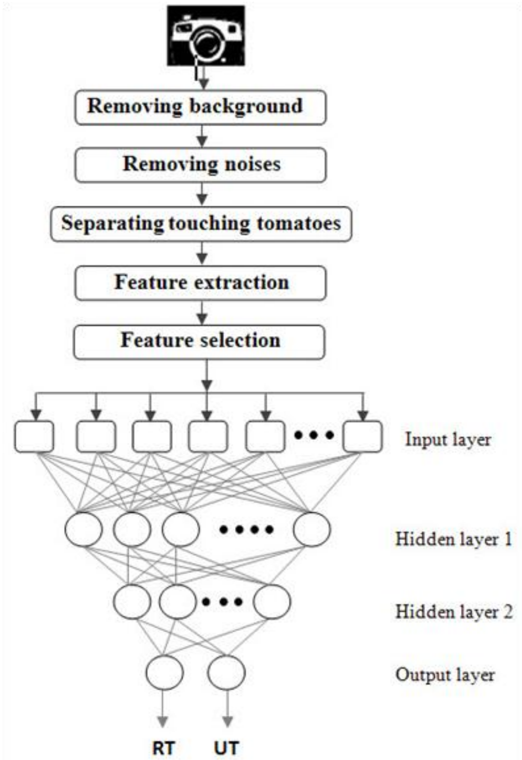

Fig 4. Flowchart of process of ripe tomato recognition: first, color images of tomato were aquired. Background of the image was removed by threshold segmentation. All noises were removed and glued tomatoes were separated based on morphology operation. Wavlet features were extracted for each tomato and then the most effective features were selected. The selected features were defined as ANN input and finally ripe and unripe tomato were recognized using trained ANN

green to yellow, pink or red on not more than 10% of the surface) were considered as UT.

Removing background

Before processing of tomato images, it was necessary to remove background. Since the background seemed in a different color from RT and UT, it was thought that segmentation based on threshold may be effective. Removing background was done in several steps, as follows:

Color data of RT, UT and background objects were extracted from images in RGB color space. The extracted data was analyzed and two effective equations D_{rg}= R-G and D_{rb}= R-B were defined that color distribution of them is shown in Fig 5. As it can be understood from diagram, the tomatoes had higher R-G value in comparison with the other objects, so D_{rg} equation was applied on color image to obtain a gray image. The color image and gray image are represented in Fig 6-a and Fig 6-b. The gray image was converted to a binary image (Fig 6-c). Tomatoes and background have the value 1 and 0 respectively. Finally, the binary image was multiplied to each R, G and B color bands to reconstruct color image (Fig 6-d).

Removing noises

Since tomato plants were cultivated at parallel lines, tomatoes at behind rows were extracted with ones at front row. To overcome this problem and to avoid robot mistake, removing all extra tomatoes was necessary. Extra tomatoes seemed smaller because their distance to machine vision system was more than near tomatoes. On this basis, removing the extra tomatoes was done based on size. To reach this aim, *opening by reconstruction operation*[1] was used.

Separation of touching tomatoes

Robot travels between ridges and stops in front of a plant. A machine vision system measures fruit maturity and finally robot will pick up RT one by one. To reach this aim, it was necessary to first separate touching tomatoes. Separating touching objects was performed in several steps as follows:

1. The two-dimensional Euclidean distance transform of the binary images was calculated (Gonzalez & Woods, 2004). As shown in Fig 6-e, the distance transform calculated the distance between a nonzero pixel (tomato pixel) and its nearest pixel with value of zero (the nearest background pixel).
2. The resulted image from former step was converted to binary image. Objects are now all fully separated (Fig 6-f).
3. Several cycles of non-merging dilation were carried out (Fig 6-g).
4. Dilated image multiplied with original one to restore object boundaries (Fig 6-h).

Color is perceived by humans as a combination of R (red), G (green), and B (blue) which is usually called three primary colors. Although RGB has been known as a powerful model, but there is high correlation among the R, G and B components. To overcome this problem, several color spaces have been derived from RGB color space by linear or non-linear relations. HSI has been known as one of the most potent color spaces that is represented by a non-linear relation from RGB as follows:

$$H = \arctan\left[\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right],$$

$$S = 1 - \frac{3}{(R+G+B)}(\min(R,G,B)), \quad I = \frac{1}{3}(R+G+B)$$

[1] Find more information in digital and image processing book by Gonzalez and Woods (2004).

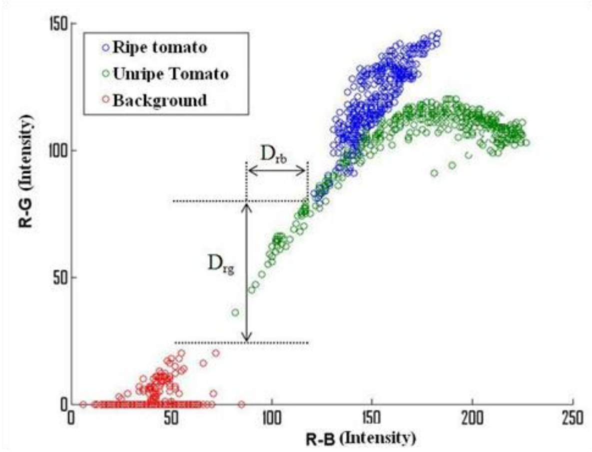

Fig 5. Pixel distribution diagram of the ripe tomato, unripe tomato, and background (Arefi et al., 2011); D_{rb} show difference between red and blue components, and subtraction of red from green component is shown by D_{rg}.

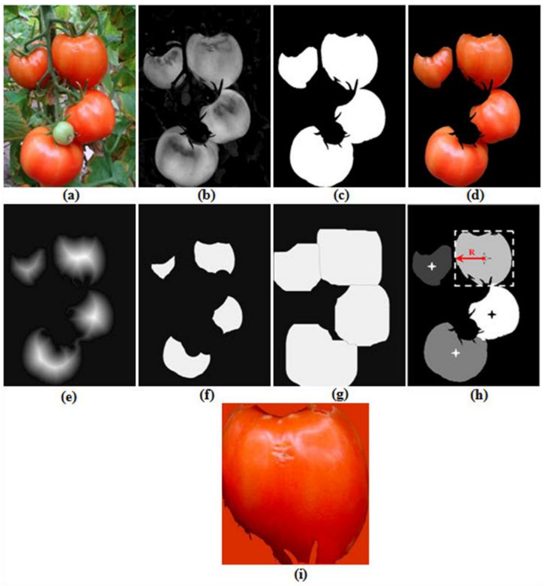

Fig 6. Image preparing operations before recognition of ripe tomato from unripe tomet. (**a**) original color image (**b**) gray image (**c**) binary iamge (**d**) image after removing background (**e**) The two-dimensional Euclidean distance transform (**f**) separated objects (**g**) non-merging dilation (**h**) drawing a squre box around each tomato (**i**) removing all pixles out of bounding box and focusing on object.

In HSI color model, color information (hue and saturation components) has been separated from intensity (Cheng et al., 2001). On this basis, two color models RGB and HSI were used for extracting wavelet features.

Preview on wavelet

Wavelet as a new method for analyzing images at multiple resolutions is gaining popularity in image processing field. Beside wavelet, there are several methods for image and signal analysis that Fourier transform is known the best. Fourier analysis breaks down a signal into multiple continuous sinusoids of different frequencies. In cases that frequency is the most important feature to distinguish a signal, the Fourier transform can be useful. However Fourier analysis has numerous serious drawbacks that missing time information is one of them. When considering a Fourier transform of a signal, it is impossible to predict when a particular event takes place. Moreover the Fourier transform is not able to perform local analysis or show other characteristics of signal as trend, drift, abrupt changing and beginning and ending of events. In effort to correct these disadvantages, wavelet analysis was proposed and progressed. The wavelet transform provides a progressive or "pyramidal" encoding of the image at various scales, which is more flexible than conventional windowed approaches like the FFT (Chui, 1992b). A wavelet is a waveform with limit duration that has an average value of zero. Wavelet analysis decomposes a signal into shifted and scaled versions of the original wavelet. Mathematically, the process of wavelet analysis is represented by the CWT for input signal f (t) (Misiti et al., 2002):

$$C\ (s,\ shift) = \int_{-\infty}^{+\infty} f(t)\psi(s, shift, t)dt\,, \tag{1}$$

where, C (scale, shift) is the CWT coefficients of f (t), shift is the translation parameter, measure of location, and s is the scale parameter. Scaling a wavelet simply means stretching or compressing it. Notice that there is an inverse relation between wavelet scales and frequency; the higher scales, the lower frequency. Calculating coefficients of CWT at every possible scale is not desirable and produces a lot of data. If we can choose effective scales, the analysis will be much more efficient and accurate. For many signals, the low-frequency content gives signal individuality and is considered as an important feature. On the other hand, high-frequency includes details of signal. On this basis, DWT was developed. In the DWT method two filter banks are used to separate low-frequency and high-frequency. Approximation and details are two sets of coefficient generated at each level of decomposition at one-dimensional DWT (Fig 7). At each level of decomposition a down-sampling is used to remove extra samples as shown in Fig 7 (Misiti et al., 2002). Unlike signals, images are two-dimensional and consequently two-dimensional DWT should be used. For this purpose, rows of image are first passed through low-pass and high-pass filters that followed by 2:1 down-sampling (for N input samples, there are N/2 low-pass output samples and N/2 high-pass ones). Former step is repeated for columns of resulting image too (Gonzalez and Woods, 2004); see Fig 8.

Extraction of wavelet features

In order to reduce sharp edges at the object boundary, it was better to limit an image to its objects (Garcia-Sevilla and Petrou, 2001). For this aim, a program was written using Matlab. In this program, first area and center of the tomato were computed. Then using area, the radius value (R) of each tomato was computed and a square R×R was drawn around it (Fig 6-h). The pixels outside this bounding box were removed and its background pixels were filled by the mean gray value of the tomato (Fig 6-i). The resulting image was converted to HSI spaces. According to Fig 9, Decomposition of resulting images was done at 3 levels using fourth-order Daubechies wavelet (Db4). The coefficients of all three levels of details were extracted for horizontal, vertical and diagonal orientations. The normalized energy, as one of the most effective wavelet features, was computed for details by Eq. 1:

$$E = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i,j)\right)^2 \quad (2)$$

Wavelet energy shows distribution of energy along frequency axis over scales and orientations (Van De Wouwer et al., 1999).

Color covariance, calculated from the energies of the detail images, is known as another powerful wavelet feature. Using following equation, color covariance was calculated by multiplying the corresponding detail coefficients of each two color bands at an orientation of decomposition (Van De Wouwer et al., 1999).

$$C_{pq} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(p(i,j)\times q(i,j)\right) \quad (3)$$

Where C_{pq} is the covariance feature between the p(i, j) and q(i, j) bands, at a particular orientation.

p and q are color bands including red (R), green (G), blue (B), hue (H), saturation (S) and intensity (I).

Since covariance features are dependent on the energies of the corresponding color bands, it is better to normalize them (Livens et al., 1997). For this aim, another feature named correlation features was defined by Eq. 3:

$$\overline{C}_{pq} = \frac{C_{pq}}{\sqrt{C_{pp}}\times\sqrt{C_{qq}}} \quad (4)$$

Where $\overline{C}_{pq}$ is the color correlation feature between the color bands, p and q, and C_{pq} is the color covariance feature between the color bands, p and q.

Energy feature of the nine decomposed images was extracted for each color band of RGB and HSI models (overall 54 energy features). In addition, by adding details (pixel-by-pixel) of three orientations three rotationally invariant details were obtained for each color band (Kim et al., 1998). Three energy features and three correlation features were calculated for each invariant detail. Totally, 36 energy and correlation features for invariant details were extracted and total of extracted features reached 90 features.

Feature selection

In pattern recognition and machine learning problems, feature selection is known as an important step. Feature selection step eliminates features with little or no predictive information and holds on the most significant subset of input variables. With reducing the magnitude of input vector, machine learning will able to do classification problems easily. During the last decade, researchers have proposed several feature selection procedures, such as principal component analysis (Omid et al., 2010), decision trees (Mollazade et al., 2009), genetic algorithms (Lu et al., 2008), support vector machines (Yang et al., 2007), etc. NeuroSolutions has provided a strong tool, Sensitivity analysis, for feature selection (NeuroSolutions for Excel, 2005). Sensitivity analysis was done to select most effective features among the primary 90 extracted features. Sensitivity analysis indicates the significance of input variables in a particular neural network, i.e. a network is analyzed and an index of importance assigned to each input. Using results of sensitivity analysis, system designer can give decision on extracted features and select those that have higher importance index.

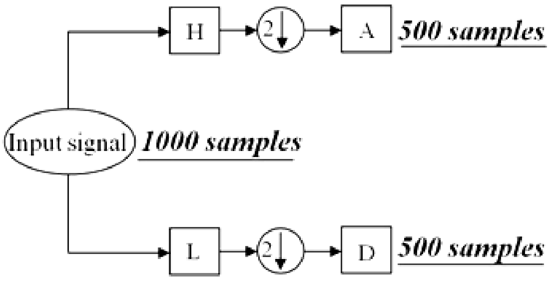

Fig 7. Decomposition of signal by two filters: Input signal is broken down into two frequencies by two low-pass and high-pass filters. In continuous, down-sampling reduces samples of each out-put signals to half (Misiti et al., 2002).

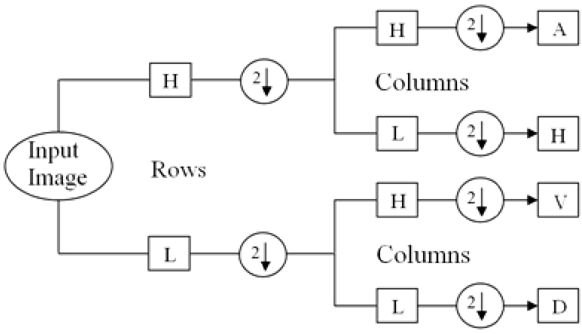

Fig 8. Decomposition of image with a filter bank: Input image breaks down into one approximation (A) and three details including horizontal (H), vertical (V) and diagonal (D) (Gonzalez and woods, 2004).

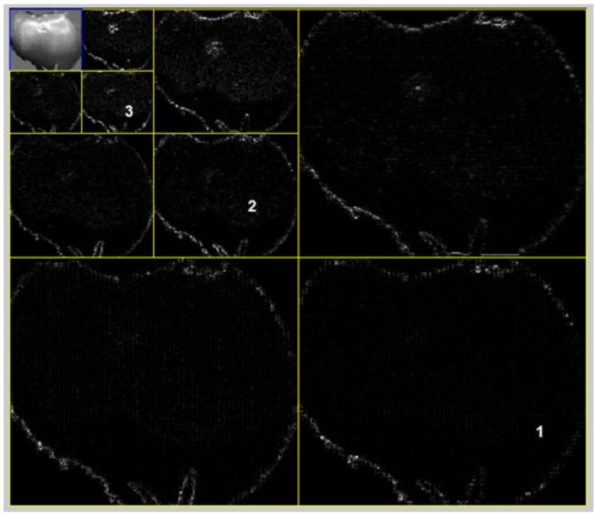

Fig 9. Each gray image was broken down into three bands by the discret wavlet transform. The first band is shown by 1, second and third bands are shown by 2 and 3, respectively.

Structure of ANN

ANN as a strong tool was used to classify RT from UT based on wavelet features. Using NeuroSolutions, a MLP network which is commonly used to classification problems was designed. MLPs often have one or more hidden layers of linear or non-linear neurons followed by an output layer.

Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. Based on trial-error method, a network with two hidden layers including 10 and 10 neurons was developed. The data set on 210 samples was split into three categories: 63% for training, 12% for cross-validation and the remaining data points (25%) for testing ANN. After adequate training, the network weights are adapted and employed for validation. The networks were trained three times and the average values were recorded for each parameter. The tangent sigmoid which is a non-linear transfer function was used for all of the neurons in the hidden and the output layers. Training was performed to minimize the mean square error (MSE) between targets and outputs.

Conclusions

The paper proposes a vision-based system for recognizing red tomatoes in a greenhouse setting. Fruit is located in images via color. After considering color data of objects, an algorithm based on threshold segmentation was defined to remove image background and could successfully remove background. Afterward touching fruits were separated using combination of erosion and non-merging dilation operations. The ninety wavelet features were extracted for each ripe and unripe to determine their effect on classification accuracy. Extracted features were reduced to the 16 effective features and used as ANN inputs. Results showed that combination of image processing and ANN could recognize ripe and unripe tomato with high accuracy.

References

Arefi A, Modarres Motlagh A, Mollazade K, Farrokhi Teimourlou R (2011) Recognition and localization of ripen tomato based on machine vision. Aust J Crop Sci. 5(10): 1144-1149.

Beaten J, Donne K, Boedrij S, Beckers W, Claesen E (2007) Autonomous fruit picking machine: A robotic apple harvester. Paper presented at the 6th International symposium on Field and Service Robotics, Chamonix, France, 1-7 December 2007.

Bulanon D-M, Kataoka T, Okamoto H, Hata S (2004) Development of a real-time machine vision system for the apple harvesting robot. Paper presented at SICE Annual Conference, Sapporo, Japan, 4-6 August 2004.

Bulanon DM, Kataoka T, Ota Y, Hiroma T (2002) A segmentation algorithm for the automatic recognition of fuji apples at harvest. Biosyst Eng. 83(4): 405-412.

Cheng HD, Jiang XH, Sun Y, Wang J (2001) Color image segmentation: advances and prospects. Pattern Recogn. 34: 2259- 2281.

Choudhary R, Paliwl J, Jayas DS (2008) Classification of cereal grains using wavelet, morphological, colour, and textural features of non-touching kernel images. Biosyst Eng. 99: 330-337.

Chui CK, (1992b) an introduction to wavelets, Academic Press.

Code of Federal Regulations United States standards for grades of fresh tomatoes (1991) USDA Agricultural Marketing Service, Washington, D.C, USA.

FAO (2009) FAOSTAT database. http://faostat.fao.org/faostat/.

Feng G, Qixin C, Masateru N (2008) Fruit detachment and classification method for strawberry harvesting robot. Int J Adv Robotic Syst. 5:41-48.

Garcia Sevilla P, Petrou M (2001) Analysis of irregularly shaped Texture regions. Comput Vis Image Und. 84: 62–76.

Gonzalez CR, Woods ER (2004) Digital image processing. In: Prentice-Hall, 2rd edn. Upper Saddle River, New Jersey, USA.

Hannan M-W, Burks T-F & Bulanon D-M (2009) A machine vision algorithm combining adaptive segmentation and shape analysis for orange fruit detection. Agricultural Engineering International: CIGR Ejournal, Manuscript 1281, Vol. XI.

Hayashi S, Sakaue O (1996) Harvesting by robotic system basic mechanism and performance. ASABE Paper No. 963067.

Hayashi S, Ganno K, Ishii Y, Tanaka I (2002) Robotic harvesting system for eggplants. Jpn Agri Res Quart. 36(3): 163-168.

Hayashi S, Ota T, Kubota K, Ganno K, Kondo N (2005) Robotic harvesting technology for fruit vegetables in protected horticultural production. In: Information and technology for sustainable fruit and vegetable production, Frutic 05, Motpellier, France.

Kim N, Wilson D, Rouse G, Udpa S (1998) Ultrasonic image texture analysis for characterizing intramuscular fat content of live beef cattle. Ultrasonic Imaging. 20: 191–205.

Kondo N, Yamamoto K, Shimizu H, Yata K, Kurita M, Shiigi T, Monta M, Nishizu T (2009) A machine vision system for tomato cluster harvesting robot. Am Soc Agri Biol Eng. 2(2): 60-65.

Livens S, Scheunders P, Van De Wouwer G, Van Dyck D (1997) Wavelets for texture analysis, an overview. Paper presented at the 6th International Conference on Image Processing and its Applications, 2: 581–585

Lu J, Zhao T, Zhang Y (2008) Feature selection based-on genetic algorithm for image annotation. Know-Based Syst. 21(8): 887-891.

Misiti M, Misiti Y, Oppenheim G, Poggi JM (2002) Wavelet Toolbox User's Guide. By The MathWorks Inc, USA.

Mollazade K, Ahmadi H, Omid M, Alimardani R (2009) An intelligent model based on data mining and fuzzy logic for fault diagnosis of external gear hydraulic pumps. Insight. 51(11): 594-600.

Monta M, Kondo N, Ting KC (1998) End-effectors for tomato harvesting robot. Artif Intell Rev. 12: 11–25.

NeuroSolutions for Excel (2005) NeuroDimension Inc, Gainesville, Florida, USA.

Okamoto H, Lee WS (2009) Green citrus detection using hyperspectral imaging. Comput Electron Agr. 66: 201-208.

Omid M, Mahmoudi A, Omid MH (2010) Development of pistachio sorting system using principal component analysis (PCA) assisted artificial neural network (ANN) of impact acoustics. Expert Syst Appl. 37(10): 7205-7212.

Reed JN, Miles SJ, Butler J, Baldwin M, Noble R (2001) Automatic mushroom harvester development. J Agri Eng Res. 78 (1): 15-23.

Sarig Y (2005) Mechanized fruit harvesting-site specific solutions. Information and technology for sustainable fruit and vegetable production, Frutic. 05: 237-247.

Tanigaki K, Fujiura T, Akase A, Imagawa J (2008) Cherry-harvesting robot. Comput Electron Agr. 63: 65–72.

Van D, Wouwer G, Scheunders P, Van Dyck D (1999) Statistical texture characterization from discrete wavelet representations. IEEE Transactions on Image Processing. 8(4): 592–598.

Wang Y, Zhu X, Ji C (2008) Machine vision based cotton recognition for cotton harvesting robot. Computer Computing Tech Agri. 2: 1421-1425.

Yang HC, Zhang SB, Deng KZ, Du, PJ (2007) Research into a feature selection method for hyperspectral imagery using PSO and SVM. J China Univ Mining Technol. 17(4): 473-478.

Zheng C, Sun DW, Zheng L (2006) Recent developments and applications of image features for food quality evaluation and inspection: A Review. Trends Food Sci Tech. 17: 642-655.