Advance Journal of Food Science and Technology 5(4): 482-487, 2013
DOI:10.19026/ajfst.5.3295
ISSN: 2042-4868; e-ISSN: 2042-4876

Submitted: December 26, 2012 Accepted: February 08, 2013 Published: April 15, 2013

Research Article

Recognition of Artificial Ripening Tomato and Nature Mature Tomato Based on the Double Parallel Genetic Neural Network

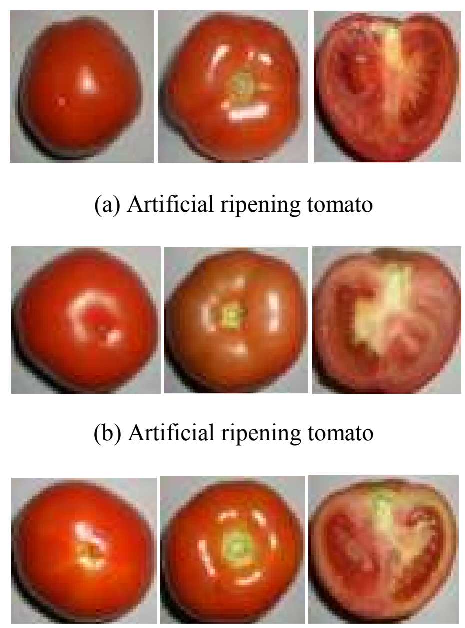
[1,2]Haibo Zhao and [3]Xianghong Zhou
[1]Engineering Technology Research Center of Optoelectronic Appliance, Anhui Province,
[2]Department of Electrical Engineering, Tongling University, Tongling Anhui, 244000, China
[3]No. 43 Research Institute, China Electronic Science and Technology Group Company, Hefei Anhui 230088, China

Abstract: In order to prevent artificial ripening tomato into markets to harm consumers' health, a double parallel genetic neural network identification system was designed. This system obtained tomato external color characteristic parameters (R, G, B) through the computer vision device and changed the RGB value into HIS value. Put tomato external color characteristic parameters as input, tomato maturity properties as output and verified the system with test samples. The test results show that, the correct recognition rate of the system is 93.8%, providing the reference for further research of artificial ripening tomato and natural mature tomato.

Keywords: Artificial ripening tomato, genetic algorithm, natural mature tomato, neural network

INTRODUCTION

There are many kinds of methods to identify artificial ripening tomatoes and natural mature tomatoes. We can see the shape and gently knead tomatoes with hand, to judge by feel (Ma, 2008). The color of natural mature tomato's appearance is orange red and artificial ripening tomato appearance's color is bright red; a natural mature tomato around pedicel shows green and artificial ripening tomato seldom (Cheng, 2008). From Fig. 1 we can see, in Fig. 1a, a big red tomato appearance, pedicel around shows khaki and after cutting the left a "hole" present, tell us that the tomato is not fresh and artificial ripening; in Fig. 1b, tomato appearance is big red, pedicel around shows few green and after cutting the right has little juice, the tomato is artificial ripening too; in Fig. 1c, tomato appearance is orange red, pedicel surrounding red and green color, after cutting juicy, red meat, no "hole", a natural mature tomato. In addition, although some tomatoes' appearance color and internal structure is normal, the color around pedicel is not normal, this kind of tomatoes is natural mature, not fresh, the consumer will not like it. This study will treat the stale tomato as artificial ripening tomato. Thus it can be seen that, from the tomato outside, appearance color and pedicel surrounding color is main basis to identify artificial ripening tomatoes and natural mature tomatoes. At present in China, mainly use artificial methods to identify artificial ripening tomato. Artificial recognition has not only heavy labor intensity, but also low work efficiency and accuracy. This study adopts automatic identification system with double parallel



(a) Artificial ripening tomato

(b) Artificial ripening tomato

(c) Nature mature tomato

Fig. 1: Comparison of artificial ripening tomato and nature mature tomato

Corresponding Author: Haibo Zhao, Engineering Technology Research Center of Optoelectronic Appliance, Anhui Province, Tongling Anhui 244000, China

genetic neural network structure, so as to realize the automatic recognition of artificial ripening tomato and natural mature tomato.

MATERIALS AND METHODS

Hardware of recognition system: Recognition system is as shown in Fig. 2. Computer has IBM compatibles (Intel Celeron 420), 2G memory, 60G hard disk and 128M video memory. Image acquisition card uses CA-CPE-3000 (Zhang *et al*., 2001; Xie, 2002) from technology group of the Zhongzi, the maximum resolution of the image acquisition display is 768×576×32 bit, image transfer as much as 60 MB/S. The CCD camera uses Panasonic WV-CP480, 1/3 inch CCD, resolution 752 (H) ×582 (V), the minimum illumination 0.001Lux, level clear 540 TVL. When testing, around tomato pedicel contact to loophole closely and ensure without gap between tomato and loophole. On both sides of the tomatoes install two mirrors angled 45° with horizontal plane. Then a CCD camera can absorb three sides image information appearance of each tomato and basically guarantee the comprehensive requirements of color detection.

Extraction of color feature: There are a variety of color models when describing color of an object. In the practical application, the commonly used models are RGB and HIS color model (Li *et al*., 2008; Jau *et al*., 2008; Zhao *et al*., 2009; Qin *et al*., 2008). RGB model, based on the display device, can accurately say color composition on screen. But the component of RGB model has no direct contact with humans' sense of colors; HIS model, based on humans' mental feeling of color, is in line with people's visual feeling and also the main use of color model in computer vision technology. The RGB model to HIS model conversion formula is:

$$\begin{cases} H = \cos^{-1}\left(\dfrac{1/2((R-G)+(R-B))}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right) \\ I = (R+G+B)/3 \\ S = 1 - 3\min(R,G,B)/(R+G+B) \end{cases} \tag{1}$$

where,
R : Red
G : Green
B : Blue
H : Hue
I : Brightness (intensity)
S : Saturation

In order to make up the shortages of a single color space representation to color characteristics, this study put a total of 24 variables as artificial ripening tomato quantitative description: the average of tomato external image R, G, B, H, I, S color component (μ_R, μ_G, μ_B , μ_H, μ_I, μ_S) and (μ'_R, μ'_G, μ'_B, μ'_H, μ'_I, μ'_S) and standard

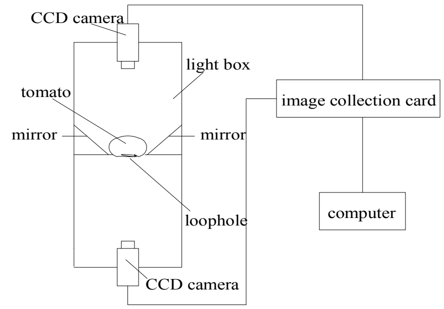


Fig. 2: Recognition system

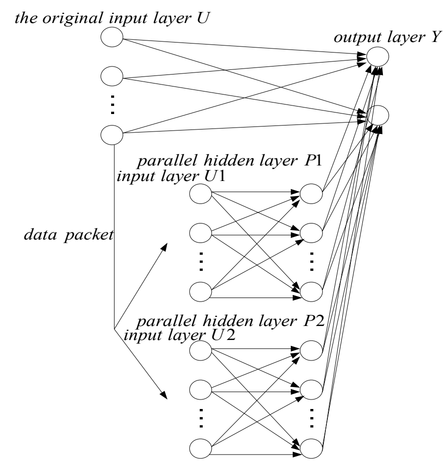


Fig. 3: Structure of double parallel feed forward neural network

deviation (σ_R, σ_G, σ_B, σ_H, σ_I, σ_S) and (σ'_R, σ'_G, σ'_B, σ'_H, σ'_I, σ'_S). Among them, (μ_R, μ_G, μ_B , μ_H, μ_I, μ_S) and (σ_R, σ_G, σ_B , σ_H, σ_I, σ_S) mean tomato appearance color's feature vector, (μ'_R, μ'_G, μ'_B, μ'_H, μ'_I, μ'_S) and (σ'_R, σ'_G,σ'_B, σ'_H, σ'_I, σ'_S) mean surrounding tomato pedicel color's feature vector.

Double parallel neural network structure: This study adopts Double Parallel Feed forward Neural Network (DPFNN), Paralleled a single forward Network and a multilayer Feed forward Network. In DPFNN, the output node receives not only the information of hidden unit, but also the information of input layer node directly. So DPFNN is a linear-nonlinear coordinately mathematical model (Guo, 2009; Zhao *et al*., 2010), as shown in Fig. 3.

Considering the recognition system in this study acquires tomato appearance color and surrounding pedicel color at the same time, neural network uses the parallel hidden layer structure. It groups the collected data, sends the tomato appearance color data to input layer U1 and the data of surrounding pedicel color to input layer U2. Due to the tomato external color feature vector (μ_R, μ_G, μ_B, σ_R, σ_G, σ_B, μ_H, μ_I, μ_S, σ'_H, σ'_I, σ'_S, μ'_R, μ'_G, μ'_B, μ'_H, μ'_I, μ'_S, σ'_R, σ'_G, σ'_B, σ'_H, σ'_I, σ'_S) having different dimension and order of magnitude, to avoid the characteristics of high dynamic range submerging

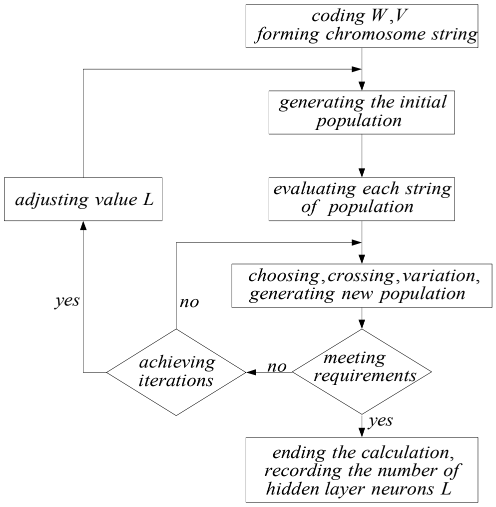


Fig. 4: Flow chart of determining neuron number of hidden layer

the features of low dynamic range, all sample data participating in the analysis will be under normalization (Chen *et al*., 2009):

$$c_n = \frac{c - \min(c)}{\max(c) - \min(c)} \tag{2}$$

where,

- c : Tomato appearance color's feature
- max(c) : Maximum value of c
- min(c) : Minimum value of c
- c_n : The characteristic value after the normalization

Put new vector (c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8 c_9 c_{10} c_{11} c_{12} c_{13} c_{14} c_{15} c_{16} c_{17} c_{18} c_{19} c_{20} c_{21} c_{22} c_{23} c_{24}) as the input of original input layer in neural network, after data grouping, (c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8 c_9 c_{10} c_{11} c_{12}) as the input of input layer U1, (c_{13} c_{14} c_{15} c_{16} c_{17} c_{18} c_{19} c_{20} c_{21} c_{22} c_{23} c_{24}) as the input of input layer U2. The output of the neural network based on tomato mature property is divided into two kinds of situations: artificial ripening or natural mature. It is expressed as (0, 1) and (1, 0) with the thermometer method. In hidden layer neurons number L is determined by trial and error method and the process shows in Fig. 4. In the diagram W is for the weight array from input layer to hidden layer, V for the weight array from hidden layer to output layer.

Genetic algorithm:

The determination of coding scheme: In order to overcome the shortcomings of binary code, the real number coding will be used. The real expression can be directly genetically operated on the phenotype of solution, instead of converting the numerical system. This study encode W and V as a chromosome string at the same time. The specific coding mode is: first gradate W and then gradate V.

The generation of initial population: Random number generator produces initial population containing N chromosome string. If the size of the initial population value of N is too small, it will be easy to fall into local optimal solution; the value of N is too big, it will reduce the efficiency. When practical application, we can only determine the value of N based on experience or experimental, generally selected in (20, 100) (Rudolph, 1994).

The determination of evaluation function: Evaluation function is defined as:

$$f_j = \frac{1}{\sum_{s=1}^{s}\sum_{i=1}^{n}(d_{si} - r_{si})^2} (j = 1, 2, \cdots, N) \tag{3}$$

where,

- j : The j^{th} evaluation value of chromosome list
- s : The number of training sample
- n : The number of neural network output nodes
- d_{si} : The expected network output of training sample
- r_{si} : The real network output of training sample

Selecting operation: In the fitness proportion method, the selected probability of each individual is proportional to its fitness value, namely:

$$p_{si} = \frac{f_i}{\sum_{i=1}^{N} f_i} \tag{4}$$

where,

- p_{si} : For the selected probability of the i^{th} chromosome list
- $\sum_{i=1}^{N} f_i$: For the sum value of each individual fitness

Adaptive crossover and mutation operation: In this study two-point-cross method is used for cross operation, bitwise variation method for variation operation (Liao *et al*., 2010). In order to maintain the diversity of population, use adaptive crossover probability P_c and mutation probability P_m. Computation formula is as follows:

$$P_c = \begin{cases} P_{c1} - \dfrac{(P_{c1} - P_{c2})(f' - f_{av})}{f_{\max} - f_{av}}, & f' > f_{av} \\ P_{c1}, & f' \le f_{av} \end{cases} \tag{5}$$

$$P_m = \begin{cases} P_{m1} - \dfrac{(P_{m1} - P_{m2})(f'_{\max} - f)}{f_{\max} - f_{av}}, & f' > f_{av} \\ P_{m1}, & f \le f_{av} \end{cases} \tag{6}$$

where,

- f' : The fitness value of the bigger one between two intersecting individuals
- f_{av} : The average fitness value of each generation group

f_{max} : Largest fitness value of the group
f : The fitness value of variation individual

Generally, $P_{c1} = 0.9$, $P_{c2} = 0.6$, $P_{m1} = 0.1$, $P_{m2} = 0.001$.

Weight adjustment of hidden layer output layer: This study adopt recursive least square (Zhao and Shan, 2007) and its basic idea is: the exact weight solution of iteration every moment, obtained by recursion of covariance matrix which is formed by input training sample, is the solution when gradient of error is zero. The characteristic of this algorithm is strong directional and fast convergence speed of iteration etc. The error objective function is defined as:

$$G(k) = \frac{1}{2}\sum_{t=1}^{k}\lambda^{k-t}(d(t) - f(t))^2 \tag{7}$$

where,
λ : Weighted forgetting factor and $0 \le \lambda \le 1$
d(t) : The desired output of the output node
f(t) : The actual output of the output node

The weight update process which using the method of recursive least square as follows:

$$w_l(k) = w_l(k-1) + g(k)[d(k) - h(k)w_l(k-1)] \tag{8}$$

$$g(k) = \frac{P(k-1)h^T(k)}{\lambda + h(k)P(k-1)h^T(k)} \tag{9}$$

$$P(k) = [P(k-1) - g(k)h(k)P(k-1)]/\lambda \tag{10}$$

RESULTS AND DISCUSSION

The process of system test as follows. Collecting tomato's external image in the test equipment, extracting color characteristic parameters (R, G, B, H, I, S) after image processing, figuring out the average value and standard deviation of color characteristic parameters and normalizing them. After that, cutting tomatoes, judging its mature properties (output values of neural network) from internal quality and then corresponding the data after normalization process with the output values of neural network, as the sample data. At last, training the double parallel feed forward neural network with genetic algorithm and the trained network model can use to forecast the test sample.

Selecting 160 tomatoes, artificial ripening and natural mature every 80, as the training sample. Another selecting 80 tomatoes, artificial ripening and natural mature every 40, as the testing sample. The selection principle of training and testing sample is the sample with enough representative and comprehensive. And the population size takes N = 80 and two hidden neurons number take L = 9 are the training conditions. All sorts of color characteristic value of training and testing samples, as shown in Table 1 and 2.

Table 1: Tomato color characteristic value of training samples

Color characteristic parameters	Mature properties	Sample number	Mean	Max.	Min.
μ_R	(0, 1)	80	212	225	208
	(1, 0)	80	204	210	197
μ'_R	(0, 1)	80	186	193	172
	(1, 0)	80	168	173	164
μ_G	(0, 1)	80	46.9000	52.7000	42.8000
	(1, 0)	80	60.6000	68.1000	55.2000
μ'_G	(0, 1)	80	58.3000	63.5000	42.7000
	(1, 0)	80	72.7000	82.4000	65.3000
μ_B	(0, 1)	80	22.6000	32.8000	13.7000
	(1, 0)	80	50.6000	52.1000	48.3000
μ'_B	(0, 1)	80	40.0000	43.8000	32.9000
	(1, 0)	80	100.4000	120.9000	90.7000
μ_H	(0, 1)	80	16.7400	18.7200	12.8200
	(1, 0)	80	12.2700	14.9300	10.8100
μ'_H	(0, 1)	80	11.2900	13.2100	10.9500
	(1, 0)	80	9.2300	10.2400	8.7800
μ_I	(0, 1)	80	0.2510	0.2990	0.1870
	(1, 0)	80	0.4750	0.5070	0.3920
μ'_I	(0, 1)	80	0.3160	0.3470	0.2580
	(1, 0)	80	0.5040	0.6290	0.4040
μ_S	(0, 1)	80	0.5810	0.6100	0.5090
	(1, 0)	80	0.5060	0.5770	0.4320
μ'_S	(0, 1)	80	0.6370	0.7930	0.5380
	(1, 0)	80	0.6160	0.7240	0.5190
σ_R	(0, 1)	80	0.0803	0.0880	0.0702
	(1, 0)	80	0.0608	0.0649	0.0551
σ'_R	(0, 1)	80	0.0662	0.0742	0.0536
	(1, 0)	80	0.0504	0.0537	0.0495
σ_G	(0, 1)	80	0.0637	0.0806	0.0240
	(1, 0)	80	0.0475	0.0704	0.0239
σ'_G	(0, 1)	80	0.0537	0.0838	0.0293
	(1, 0)	80	0.0484	0.0632	0.0355
σ_B	(0, 1)	80	0.0273	0.0385	0.0148
	(1, 0)	80	0.0317	0.0472	0.0233
σ'_B	(0, 1)	80	0.0274	0.0440	0.0181
	(1, 0)	80	0.0248	0.0303	0.0116
σ_H	(0, 1)	80	0.3740	0.4470	0.1540
	(1, 0)	80	0.2730	0.5630	0.0280
σ'_H	(0, 1)	80	0.3850	0.5630	0.2860
	(1, 0)	80	0.2950	0.4480	0.1830
σ_I	(0, 1)	80	0.000567	0.000832	0.000397
	(1, 0)	80	0.000486	0.000643	0.000295
σ'_I	(0, 1)	80	0.000543	0.000630	0.000296
	(1, 0)	80	0.000682	0.000782	0.000406
σ_S	(0, 1)	80	0.000473	0.000509	0.000327
	(1, 0)	80	0.000307	0.000468	0.000259
σ'_S	(0, 1)	80	0.000448	0.000627	0.000290
	(1, 0)	80	0.000482	0.000539	0.000237

Min.: Minimum; Max.: Maximum

Table 2: Tomato color characteristic value of test samples

Color characteristic parameters	Mature properties	Sample number	Mean	Max.	Min.
μ_R	(0, 1)	80	219	232	214
	(1, 0)	80	201	204	192
μ'_R	(0, 1)	80	182	187	175
	(1, 0)	80	173	178	169
μ_G	(0, 1)	80	52.3000	59.9000	49.1000
	(1, 0)	80	58.9000	62.7000	56.8000
μ'_G	(0, 1)	80	62.3000	65.8000	45.0000
	(1, 0)	80	78.3000	80.3000	59.4000
μ_B	(0, 1)	80	22.6000	35.5000	16.9000
	(1, 0)	80	55.1000	57.4000	52.4000
μ'_B	(0, 1)	80	38.7000	42.2000	33.2000
	(1, 0)	80	105.7000	118.8000	88.7000
μ_H	(0, 1)	80	15.7200	17.6400	11.3800
	(1, 0)	80	13.7300	15.0300	11.8200
μ'_H	(0, 1)	80	12.5700	14.2800	9.9500
	(1, 0)	80	10.2000	11.6700	7.6900
μ_I	(0, 1)	80	0.2490	0.3180	0.1730
	(1, 0)	80	0.4090	0.5120	0.3840
μ'_I	(0, 1)	80	0.3060	0.3250	0.2630
	(1, 0)	80	0.6730	0.7730	0.5300

Table 2: Continue

Color characteristic parameters	Mature properties	Sample number	Mean	Max.	Min.
μ_S	(0, 1)	80	0.5940	0.6270	0.5180
	(1, 0)	80	0.5480	0.6630	0.5490
μ'_S	(0, 1)	80	0.6970	0.7360	0.5340
	(1, 0)	80	0.7070	0.8200	0.6390
σ_R	(0, 1)	80	0.0829	0.0838	0.0747
	(1, 0)	80	0.0639	0.0746	0.0527
σ'_R	(0, 1)	80	0.0669	0.0728	0.0529
	(1, 0)	80	0.0550	0.0639	0.0518
σ_G	(0, 1)	80	0.0649	0.0938	0.0227
	(1, 0)	80	0.0482	0.0738	0.0254
σ'_G	(0, 1)	80	0.0509	0.0839	0.0210
	(1, 0)	80	0.0427	0.0645	0.0338
σ_B	(0, 1)	80	0.0299	0.0450	0.0197
	(1, 0)	80	0.0399	0.0465	0.0303
σ'_B	(0, 1)	80	0.0286	0.0439	0.0145
	(1, 0)	80	0.0268	0.0323	0.0146
σ_H	(0, 1)	80	0.3890	0.4530	0.1760
	(1, 0)	80	0.3730	0.5480	0.0190
σ'_H	(0, 1)	80	0.4280	0.5250	0.3100
	(1, 0)	80	0.3120	0.4770	0.2060
σ_I	(0, 1)	80	0.000579	0.000734	0.000363
	(1, 0)	80	0.000404	0.000626	0.000275
σ'_I	(0, 1)	80	0.000559	0.000732	0.000286
	(1, 0)	80	0.000538	0.000707	0.000436
σ_S	(0, 1)	80	0.000462	0.000527	0.000318
	(1, 0)	80	0.000319	0.000485	0.000220
σ'_S	(0, 1)	80	0.000467	0.000663	0.000238
	(1, 0)	80	0.000540	0.000617	0.000283

Min.: Minimum; Max.: Maximum

Table 3: Recognition results of genetic neural network

Mature properties	Sample number	Identify results		Correct recognition rate/%
		Artificial ripening	Natural mature	
Artificial ripening	40	37	3	92.5
Natural mature	40	2	38	95

Table 4: Comparison between two algorithms

Class of algorithm	Correct recognition rate /%	
	Training sample	Test sample
Genatic alogrithm	100	93.8
BP alogrithm	95.6	87.4

Identifying different varieties of tomatoes with the trained network and the results shows as Table 3. The test showed that the system whose average correct identification rate was 93.8% identifies better.

Table 4 shows the comparative result of correct recognition rate of identification test to ripening tomato which is comparing the neural network of same structures trained by BP algorithms with trained by genetic algorithms. According to Table 4, the correct recognition rate, the automatic identification of artificial ripening tomato with the neural network trained by genetic algorithms, is obviously higher than the same structure neural network which is trained with BP algorithm.

CONCLUSION

Based on the extraction to image information of tomato's external color, with the average value and standard deviation of the color component of image R, G, B, H, I, S surrounding tomato's appearance and pedicel as the input of the neural network and the multilayer feed forward neural network adopting genetic algorithm training realized the automatic identification of artificial ripening tomato. Tests showed that the correct recognition rate of the system is 93.8, which laid a certain theoretical and practical basis for further recognition studies of artificial ripening tomato.

ACKNOWLEDGMENT

The authors thank the Engineering Technology Research Center of Optoelectronic Appliance, Anhui Province of China for support.

REFERENCES

Chen, K.J., X. Sun and Q.Y. Lu, 2009. Automatic color grading of beef lean tissue based on BP neural network and computer vision. Trans. CSAE, 40(4): 174-178.

Cheng, R.Z., 2008. How to identify artificial ripening tomato. Shan Xi Lao Nian, 4: 34.

Guo, J.P., 2009. Convergence of batch gradient descent algorithm for high order double parallel neural networks. M.S. Thesis, Dalian University of Science and Technology.

Jau, U.L., S.T. Chee and W.N. Giap, 2008. A comparison of RGB and HSI color segmentation in real-time video images: A preliminary study on road sign detection. Proceeding of International Symposium on Information Technology, 3: 344-349.

Li, G., B. Li and Y. Wang, 2008. Method for classification of pearl color based on HIS model. Trans. CSAE, 24(8): 284-287.

Liao, B.X., Y.M. Yang and X.X. Zhang, 2010. Application of adaptive genetic algorithm and RBF network in passing ball. Comp. Simul., 27(9): 169-172.

Ma, X.Q., 2008. How to identify artificial ripening tomato and nature mature tomato. Nong Cun Bai Shi Tong, 11: 56.

Qin, J.W., T.F. Burks, D.G. Kim and D.M. Bulanon, 2008. Classification of citrus peel diseases using color texture feature analysis. Proceeding of American Society of Agricultural and Biological Engineers-Food Processing Automation Conference. Providence, Rhode Island, June 28-29, pp: 170-178.

Rudolph, G., 1994. Convergence analysis of canonical genetic algorithms. IEEE T. Neural Networ., 5(1): 96-101.

Xie, F.Y., 2002. Reaearching on color sortor for raisin. M.S. Thesis, Changchun University of Science and Technology.

Zhang, C.L., J.L. Fang and W. Pan, 2001. Automated identification of tomato maturation using multilayer feedward neural network with genetic algorithm. Trans. CSAE, 17(3): 153-156.

Zhao, Z.G. and X.H. Shan, 2007. Optimization approach based on genetic algorithm for RBF neural network. Comp. Eng., 33(6): 211-212.

Zhao, H.Q., J.Q. Qi and J. Wang, 2010. Concentration estimation of gas mixture using a wavelet-based DPFNN. Chinese J. Sens. Actuat., 23(5): 744-747.

Zhao, X., T.F. Burks, J. Qin and M.A. Ritenour, 2009. Digital microscopic imaging for citrus peel disease classification using color texture features. Appl. Eng. Agric., 25(5): 769-776.