

# 补偿模糊神经网络水果形状分级器分级误差

曹乐平<sup>1</sup>, 温芝元<sup>2</sup>

(1. 湖南生物机电职业技术学院, 长沙 410127; 2. 湖南农业大学理学院, 长沙 410128)

**摘要:** 针对神经网络对水果进行分级时精度有待提高的问题, 分析了补偿模糊神经网络柑桔形状分级器的分级误差。将柑桔图像前4个傅里叶描述子按期望输出模糊变量值大小排列成单调递增、单调递减、钟形分布和锯齿形分布4种训练样本, 分别训练同一补偿模糊神经网络水果形状分级器, 用递减排序后的同一测试样本检验分级器性能, 试验表明, 单调递减顺序训练样本所训练的分级器分级误差最小为1.875%, 钟形分布、单调递增顺序和锯齿形分布训练样本所训练的分级器分级误差依次增大, 分别为15%、63.125%、75%。分析分级误差与样本间顺序的对应关系, 建立分级误差模型, 结果表明, 同顺序的测试样本与训练样本间相关系数大, 分级误差小; 不同顺序的测试样本与训练样本间相关系数小, 分级误差大。因此, 测试样本与训练样本按水果同一品质特征同序排列, 提高样本间的相关程度, 将大幅度降低神经网络类分级器分级误差, 提高正确识别率。

**关键词:** 补偿模糊神经网络, 离散傅里叶变换, 谐波分量, 水果形状, 分级误差

**中图分类号:** TP183.S66

**文献标识码:** A

**文章编号:** 1002-6819(2008)-12-0102-05

曹乐平, 温芝元. 补偿模糊神经网络水果形状分级器分级误差[J]. 农业工程学报, 2008, 24(12): 102—106.

Cao Leping, Wen Zhiyuan. Error of the compensation fuzzy neural network fruit grader[J]. Transactions of the CSAE, 2008, 24(12): 102—106. (in Chinese with English abstract)

## 0 引言

目前应用机器视觉技术描述水果形状常用的方法有5种: 几何参数法、小波多尺度法、傅里叶描述子、主动形状模型和Zernike矩。通过这些方法提取水果形状特征参数后用多元统计、支持向量机、神经网络等方法进行水果形状等级映射。Majid Rashidi等<sup>[1]</sup>利用相垂直的两个剖面上椭圆长轴与短轴的比评判哈密瓜果形。D.P. Whitelock等<sup>[2]</sup>用含纵径和横径在内的6个几何参数识别苹果形状。桂江生等<sup>[3,4]</sup>通过小波多尺度法描述水果果形。M.Z. Abdullah等<sup>[5]</sup>用傅里叶描述子评定阳桃形状, 通过多层神经网络进行分级, 正确率为80%~89%。Ingrid Paulus等<sup>[6]</sup>借助傅里叶变换的前12个谐波分量评价6种苹果果形, 决定系数0.98。文献[7]用苹果离散傅里叶变换的前15个分量识别苹果形状, 分级精度在85%以上。文献[8]提出了基于主动形状模型的苹果形态分级方法, 判别准确度95%。文献[9, 10]将Zernike矩应用到苹果的识别中, 通过支持向量机进行分级, 正常果形、轻度畸形和严重畸形果识别率都较高。文献[11]设计了用3个方向的几何特征参数构造的苹果果形线性分类器, 100个苹果识别成功率为97%。以上文献虽应用了不同的水果形状等级映射方法, 但主要是从水果形状描述方面提高分级精度。

神经网络分级器的检测与分级精度并不稳定, 不仅与网络类型、网络结构、算法等有关, 还与学习样本数

量和顺序有关。文献[12]指出影响神经网络泛化能力的根本原因是训练样本质量、数量和代表性, 但未提及样本顺序问题。文献[13]比较多层感知器预测误差评价方法后得出: ‘bootstrap’方法最精确, 但就预测误差的降低未给出具体方法。水果形状分级误差来源于形状描述和等级映射两部分, 本文以补偿模糊神经网络水果形状分级器为研究对象, 区分同一学习样本集的4种不同顺序, 用同一测试样本集的顺序检验分级器的性能, 研究分级精度与学习样本顺序的对应关系, 为水果自动分级误差的降低提供依据。

## 1 补偿模糊神经网络水果形状分级器与试验材料

### 1.1 补偿模糊神经网络水果形状分级器

神经网络所具有的强大的学习能力为刻画难于用数学解析表达式描述的复杂非线性系统和非线性关系提供了一种方便的工具<sup>[14]</sup>。不少学者用神经网络作为分级器实现水果大小、形状、颜色、缺陷等外观品质及糖度、酸度等内部品质的检测与分级<sup>[15-19]</sup>。本文用傅里叶描述子识别柑桔的形状, 借助补偿模糊神经网络进行形状分级。通过图像边界求水果形心, 处于边界上的中心角等分点构成半径序列, 将其归一化后, 作离散傅立叶变换, 取离散傅立叶变换的前4个谐波分量 $F(1)$ 、 $F(2)$ 、 $F(3)$ 和 $F(4)$ 作为水果的形状参数。定义:  $x_1=F(1)$ 为水果曲线弯曲程度模糊变量, 模糊子集为 $A_1^k$ , 语言值{强, 中}={H,M};  $x_2=F(2)$ 为水果细长比程度模糊变量, 模糊子集 $A_2^k$ , 语言值{小, 较小}={S,SE};  $x_3=F(3)$ 为水果三角形状模糊变量, 模糊子集为 $A_3^k$ , 语言值{较尖, 较圆滑}={J,JS};  $x_4=F(4)$ 为水果方形度模糊变量, 模糊子集为 $A_4^k$ , 语言值{较圆, 圆}={SEC,SC};  $y$ 为水果形状等级模糊变量,  $B^k$ 为水果等级模糊子集, 语言值{优秀, 标准,

收稿日期: 2007-05-22 修订日期: 2008-01-04

基金项目: 湖南省教育厅科学研究项目 (06D059)

作者简介: 曹乐平 (1964—), 男, 湖南益阳人, 副教授, 研究方向: 农产品智能检测。长沙 湖南生物机电职业技术学院, 410127。

Email: clp4218@126.com

合格品, 次品}={BO,SQ,HG,TA}。模糊 IF-THEN 规则  $FR^{(k)}$ : [IF  $x_1$  is  $A_1^k$  and ... $x_4$  is  $A_4^k$ ] THEN  $y$  is  $B^k$  ( $k=1,2,\cdots 16$ ) 列于表 1。

表 1 模糊控制规则  
Table 1 Fuzzy control rule

序号	$x_1$	$x_2$	$x_3$	$x_4$	$y$
1	H	SE	JS	SC	SQ
2	M	SE	JS	SC	SQ
3	H	S	JS	SC	HG
4	M	S	JS	SC	BO
5	M	S	J	SC	SQ
6	H	S	J	SC	HG
7	M	SE	J	SC	SQ
8	H	SE	J	SC	SQ
9	H	SE	JS	SEC	SQ
10	M	SE	JS	SEC	SQ
11	H	S	JS	SEC	HG
12	M	S	JS	SEC	SQ
13	M	S	J	SEC	SQ
14	H	S	J	SEC	HG
15	M	SE	J	SEC	SQ
16	H	SE	J	SEC	TA

建立如图 1 所示的 5 层补偿模糊神经网络水果形状分级器 ( $4 \times 64 \times 8 \times 4 \times 1$ )。输入层连接各输入向量; 模糊化层每个结点代表一个语言变量值, 计算各输入向量属于各语言变量值模糊集合的隶属函数; 模糊推理层采

用乘积推理规则, 每个结点代表一条模糊规则, 计算各条规则的适应度; 补偿运算层采用消积与积极运算; 反模糊化层用改进型重心函数解模糊。

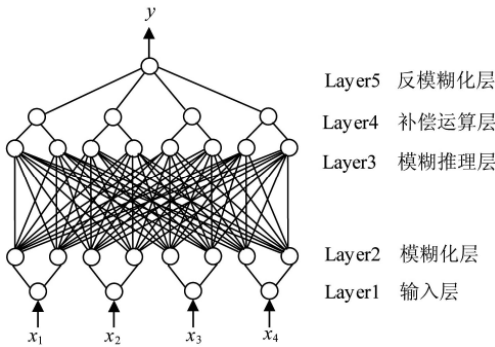


图 1 水果形状分级器  
Fig.1 Fruit shape grader

1.2 试验材料

市场采购湘西自治州椪柑 432 个, 将其分为 272 个训练样本和 160 个测试样本, 其分布如表 2, 离散傅立叶变换后前 4 个谐波分量统计如表 3。

表 2 训练样本与测试样本分布  
Table 2 Training sample and test sample distribution

训练样本/个				测试样本/个			
次品	优质	合格	标准	次品	优质	合格	标准
24	24	64	160	10	10	40	100

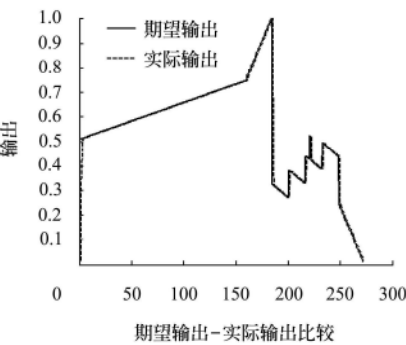
表 3 样本统计  
Table 3 Test sample statistics

	F(1)				F(2)				F(3)				F(4)			
	平均值	最大值	最小值	标准差	平均值	最大值	最小值	标准差	平均值	最大值	最小值	标准差	平均值	最大值	最小值	标准差
训练样本	0.0158	0.0612	0.0010	0.0246	0.0056	0.0218	0.0005	0.0083	0.0066	0.0241	0.0005	0.0101	0.0162	0.0608	0.0010	0.0253
测试样本	0.0172	0.0610	0.0011	0.0217	0.0060	0.0215	0.0005	0.0087	0.0093	0.0240	0.0003	0.0101	0.0097	0.0530	0.0011	0.0146

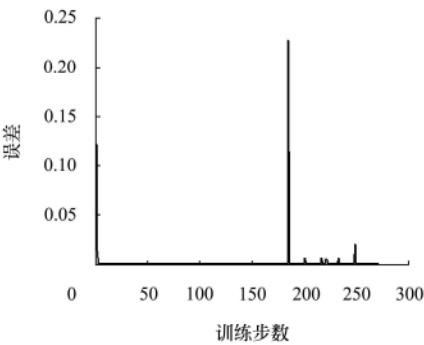
2 训练样本顺序对分级器分级误差的影响

形状等级模糊区间按固定区间划分法均匀分割为:  $0 \sim 0.25$  为次品,  $0.25 \sim 0.5$  为合格果,  $0.5 \sim 0.75$  为标准果,  $0.75 \sim 1.0$  为优质果。将 272 个训练样本以期

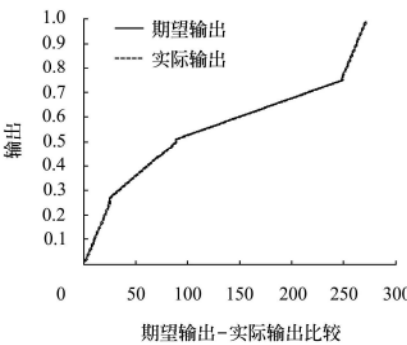
望输出大小排序, 形成单调递增顺序、单调递减顺序、钟形分布和锯齿形分布。在同一精度 0.0005 下, 分别训练同一补偿模糊神经网络水果形状分级器, 输出及训练误差如图 2。



a. 钟形分布样本训练输出



b. 钟形分布样本训练误差



c. 单调递增顺序样本训练输出

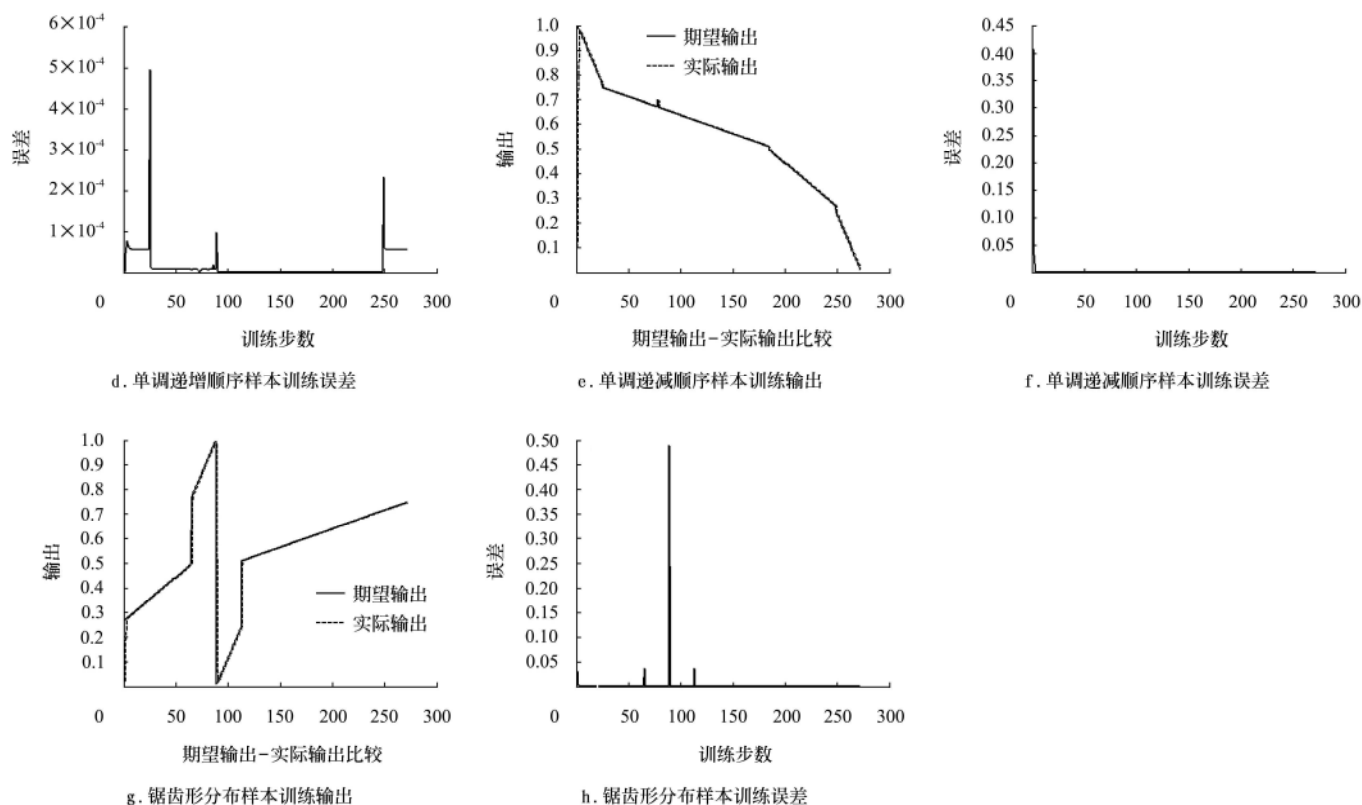


图 2 补偿模糊神经网络水果形状分级器训练

Fig.2 Compensation fuzzy neural network fruit grader training

160 个测试样本按期望输出值大小递减排序, 测试用 4 种不同顺序训练样本学习好了的补偿模糊神经网络水果形状分级器, 结果如表 4~7。

表 4 分级器的测试性能 (单调递增顺序样本训练)

Table 4 Test performance of the grader (sample training of the monotone increasing order)

优质			标准			合格			次品		
测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %
10	10	100	100	41	41	40	40	100	10	10	100

表 5 分级器的测试性能 (单调递减顺序样本训练)

Table 5 Test performance of the grader (sample training of the monotone decreasing order)

优质			标准			合格			次品		
测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %
10	1	10	100	0	0	40	1	2.5	10	1	10

表 6 分级器的测试性能 (钟形分布样本训练)

Table 6 Test performance of the grader (sample training of the bell-shape distribution)

优质			标准			合格			次品		
测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %
10	10	100	100	11	11	40	2	5	10	1	10

表 7 分级器的测试性能 (锯齿形分布样本训练)

Table 7 Test performance of the grader (sample training of the zigzag distribution)

优质			标准			合格			次品		
测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %	测试样本 / 个	误判 / 个	误差 / %
10	10	100	100	60	60	40	40	100	10	10	100

### 3 测试误差分析与讨论

160 个测试样本检验 4 种不同顺序训练样本学习好了的水果形状分级器, 总分级误差如表 8。

表 8 分级模型性能测试

Table 8 Test performance of the model

样本顺序	分级误差 / %			
	单调递增 学习样本	单调递减 学习样本	锯齿分布 学习样本	钟形分布 学习样本
单调递减测试样本	63.125	1.875	75	15

测定训练样本  $x_i (i=1,2,3,4)$  与测试样本  $x'$  间的相关

系数  $r_i$ , 如表 9, 平均相关系数  $\bar{r} = (\prod_{i=1}^4 absr_i)^{0.25}$ 。

作总分级误差与平均相关系数的 2 次拟合, 如图 3。平均相关系数为  $\bar{r}=0.0967$  时, 总分级误差出现峰值  $e=81.69\%$ , 经多次训练与测试, 峰值点左右, 虽总分级误差都有减小的趋势, 但左侧的减小空间不大, 总是偏离 0 一定距离, 其大小由神经网络泛化能力的高低决定,

表 9 相关性测定

Table 9 Relevant determinations

钟形分布					递增分布					递减分布					锯齿形分布				
$r_1$	$r_2$	$r_3$	$r_4$	$\bar{r}$	$r_1$	$r_2$	$r_3$	$r_4$	$\bar{r}$	$r_1$	$r_2$	$r_3$	$r_4$	$\bar{r}$	$r_1$	$r_2$	$r_3$	$r_4$	$\bar{r}$
-0.2524	0.0832	-0.1283	-0.1172	0.1333	-0.0613	-0.0403	0.0720	-0.2019	0.0774	-0.1818	-0.0321	0.2206	0.2480	0.1337	-0.0548	0.2379	-0.0798	-0.1320	0.1083

表明通过训练样本与测试样本弱的相关性即两样本的乱序排列使总分级误差降低无法达到低误差状态。测试样本的递减分布与训练样本的递增分布及锯齿形分布相关系数平均值最小，水果形状分级误差最大。而峰值点右侧，存在接近于 0 的大幅度下降空间，表明通过训练样本与测试样本强的相关性即两样本的同序排列使总分级误差降低接近于 0，测试中，当训练样本与测试样本都为相同的递减顺序排列时，平均相关系数值最大，水果形状分级误差最小为 1.875%，此时，分级器识别正确率最高，比文献[5，7-10]高出 3% 以上。神经网络的泛化能力具有选择相关性，不仅与训练样本数量有关，还与训练样本的分布有关，低的误差要求训练样本与检测样本有高的相关性。

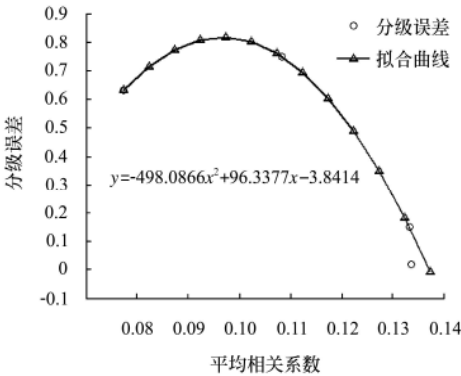


图 3 总分级误差拟合曲线  
Fig.3 Fitted curve of the total graduation error

对于神经网络分级器，提高分级正确率，训练样本尽可能数量多而全面，覆盖所有等级，避免神经网络泛化能力不强而分级中出现多的泛化检测样本；在目前增殖神经网络仍处于研究中的情况下，分级器训练样本不具备高的通用性，须区别不同品种，不同产地，不同采摘期，实现测试样本与训练样本同品种，同产地，同采摘期，同批次化，极大程度地保证与检测样本高的相关性。待分级的在线水果，按训练样本分布形式以同特征排序，即待分级的水果以收购批次为单位，按生成训练样本的特征顺序制定分级秩序，不能交叉，否则大的分级误差不可避免。

为了进一步验证分级误差与样本顺序的相关性，采用柑橘等级规格（NY/T1190-2006）分级标准，横径区分椪柑大小，并划分为 6 组；傅里叶描述子识别椪柑的形状，并划分为 3 组，对以上 160 个测试样本进行大小和形状 4 等级综合分级，分级误差如表 10。由表 10 可知，综合分级同样以两类样本顺序相同的情况分级误差最小。因此，对于多指标综合分级中，可采用如下操作提

高分级正确率：

- 1) 配置滚筒式大小分级机，由该分级机预先分级一定数量的水果，形成集与集之间大小不一但同集间大小均匀的多个训练样本集，训练样本集的数量与滚筒式大小分级机所分级的水果等级数对应。
- 2) 大小均匀的同一训练样本集训练一条水果自动化分级生产线的神经网络分级器。样本数量与自动化分级生产线数量对应。
- 3) 进入正式分级阶段，滚筒式大小分级机分级的不同大小的水果分别由神经网络分级器训练中对应水果大小的自动化分级生产线实施分级，一条水果自动化分级生产线分级一种大小的水果。

表 10 综合分级误差

Table 10 Integrative grading error

样本顺序	分级误差/%			
	单调递增 学习样本	单调递减 学习样本	锯齿分布 学习样本	钟形分布 学习样本
单调递减测试样本	64.375	2.5	76.875	16.25

4 结 论

试验表明，补偿模糊神经网络水果形状分级器分级误差随训练样本与测试样本间相关系数的变化而变化，同排列顺序的训练样本与测试样本间相关系数大，分级器分级误差低；相反，不同顺序的训练样本与测试样本间相关系数小，分级器分级误差高。这一结论同样适用水果其它品质的检测与分级。

提高训练样本与测试样本间的相关程度，从而控制补偿模糊神经网络水果分级器分级误差。对于水果外观品质的综合分级，除按等级收购，并按各等级依次分级，避免训练样本与测试样本间等级交叉外，一方面即时更替训练样本，保证训练样本与待分级水果同品种、同产地、同采摘期；另一方面，训练样本与待分级水果以同一品质特征同序排列，避免训练样本与测试样本间排序的标准不统一和排序的顺序不统一。如以大小为特征排序，配置一台机械式水果大小分级机，先分级出各等级的训练样本，每一等级的训练样本训练一套自动分级系统，然后机械式分级机预分级后的各等级水果通过传送装置分别进入对应的自动分级系统完成分级。

[参 考 文 献]

[1] Majid Rashidi, Keyvan Seyfi. Classification of fruit shape in cantaloupe using the analysis of geometrical attributes[J]. World Journal of Agricultural Sciences, 2007, 3(6): 735—740.

- [2] Whitelock D P, Brusewitz G H, Stone M L. Apple shape and rolling orientation[J]. *Applied Engineering in Agriculture*, 2006, 22(1): 87—94.
- [3] Jiangsheng Gui. Fruit Shape Detection Based on Multi-scale Level Set Framework[R]. Yibin Ying. St. Joseph, Michigan: the American Society of Agricultural and Biological Engineers, 2006: 1—11.
- [4] Jiangsheng Gui, Yibin Ying, Xiuqin Rao. A novel fruit shape classification method based on multi-scale analysis[A]. *Optical Sensors and Sensing Systems for Natural Resources and Food Safety and Quality*[C]. *Proceedings of SPIE*, 2005, 5996, 59961F: 1—9.
- [5] Abdullah M Z, Mohamad-Saleh J, Fathinul-Syahir A S, et al. Discrimination and classification of fresh-cut starfruits (*Averrhoa carambola* L.) using automated machine vision system[J]. *Journal of Food Engineering*, 2006, 76(4): 506—523.
- [6] Ingrid Paulus, Eddic Schrevers. Shape characterization of new apple cultivars by fourier expansion of digitized images [J]. *Journal of Agricultural Engineering Research*, 1999, 72(2): 113—118.
- [7] 林开颜, 吴军辉, 徐立鸿. 基于计算机视觉技术的水果形状分级方法[J]. *农业机械学报*, 2005, 36(6): 71—74.
- [8] 蔡健荣, 许月明. 基于主动形状模型的苹果果形分级研究[J]. *农业工程学报*, 2006, 22(6): 123—126.
- [9] 应义斌, 桂江生, 饶秀勤. 基于 Zemike 矩的水果形状分类[J]. *江苏大学学报(自然科学版)*, 2007, 28(1): 1—3, 67.
- [10] Jiangsheng Gui, Yibin Ying, Xiuqin Rao. A new algorithm for fruit shape classification based on level set[A]. *Intelligent Robots and Computer Vision XXIV: Algorithms, Techniques, and Active Vision*[C]. *Proceedings of SPIE*, 2006, 6384, 63840X: 1—8.
- [11] 谢国俊, 曹其新, 刘建政, 等. 基于多方位视觉的果实形状特征的提取研究[J]. *农业工程学报*, 2007, 23(7): 127—132.
- [12] 江学军, 唐焕文. 前馈神经网络泛化性能力的系统分析[J]. *系统工程理论与实际*, 2000, (8): 36—40.
- [13] Robert Tibshirani. A comparison of some error estimates for neural network models[J]. *Neural Computation*, 1996, 8: 152—163.
- [14] 王守觉, 李兆洲, 王柏南, 等. 用前馈神经网络进行带噪声信号的去噪声建模[J]. *电路与系统学报*, 2000, 5(4): 21—26.
- [15] Heinemann P H, Sommer H J, Morrow C T, et al. Machine vision based station for grading of 'Golden Delicions' apples[R]. Chicago: *Proceedings of the FPAC IV Conference*. 1995: 239—248.
- [16] Leemans V. Apple Shape Inspection with Computer Vision[R]. Chicago: *Proceedings of the FPAC IV Conference*. 1995: 316—327.
- [17] 刘 禾, 汪懋华. 基于数字图像处理的苹果表面缺陷分类方法[J]. *农业工程学报*, 2004, 20(6): 138—140.
- [18] 洪添胜, 乔 军, Ning Wang, 等. 基于高光图像技术的雪花梨品质无损检测[J]. *农业工程学报*, 2007, 23(2): 151—155.
- [19] 邹小波, 赵杰文, 潘胤飞, 等. 基于遗传 RBF 网络的电子鼻对苹果质量的评定[J]. *农业机械学报*, 2005, 36(1): 61—64.

## Error of the compensation fuzzy neural network fruit grader

Cao Leping<sup>1</sup>, Wen Zhiyuan<sup>2</sup>

(1. Hunan Biological and Electromechanical Polytechnic, Changsha 410127, China;

2. College of Science, Hunan Agricultural University, Changsha 410128, China)

**Abstract:** This investigation analyses grading errors based on compensating fuzzy neural network to find out methods to improve citrus fruit grading. The first four harmonic components of discrete Fourier transform of citrus fruit images were sorted according to expected fuzzy values, and formed into four training sample sets which were in the order of monotone increasing, monotone decreasing, bell-shaped distribution and sawtooth distribution respectively. The new four sample sets were used to train compensating fuzzy neural network of the same architecture. Test set was sorted in monotone decreasing order. Test results show that model trained with monotone decreasing sample set has the smallest grading error which is 1.875%. Models trained with bell-shaped distribution, monotone increasing and sawtooth distribution have larger grading errors, which are 15%, 63.125%, 75% respectively. Grading error is modeled, and analyzed on correlation of grading error and differential order of sample sets. Results show that grading error is smaller with big correlation coefficient if training and testing sets are in the same order. On the contrary, if training and testing sample sets are in different orders, the correlation coefficient is smaller and the grading error is larger. This discloses testing and training sets shall be sorted in the same order according to qualitative feature as far as possible to improve the correlation coefficients. Performance of auto-grading system based on neural network model can be improved greatly by sorting samples in the same order.

**Key words:** the compensation fuzzy neural networks, discrete fourier transform, harmonic component, fruit shape, grading error