TRANSACTIONS OF THE CHINESE SOCIETY FOR AGRICULTURAL MACHINERY
Vol. 47, No. 11, 2016

doi:10.6041/j.issn.1000-1298.2016.11.001

Fruit Cluster Recognition and Picking Sequence Planning Based on Selective Attention

Wang Bingxin Wang Sun'an Yu Dehong
(*School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an* 710049, *China*)

Abstract: To improve algorithm versatility and picking efficiency, recognizing more kinds of fruits or vegetables and planning picking sequence for fruit clusters in visual field have become hotspots and trends of harvesting robot research. Primates can find objects and shift attention through visual selective attention mechanism, which has similarities with harvesting robots to recognize targets and to plan picking sequence. Hence, by adopting the idea and technology of bionics, a new visual selective attention-based method was proposed for fruit cluster recognition and picking sequence planning. According to Itti visual attention computational model, this algorithm changed computing process for color feature maps, and improved feature integration method for building color conspicuity map and saliency map by introducing priori knowledge about fruits and vegetables. Using the reference of the biological neural network competition mechanism called Winner-Take-All and artificial expertise from professional picking operators, distance, area and saliency were adopted to design the weighted preferential picking sequence strategy for fruit clusters. The experimental results showed that the proposed method in this paper achieved recognition of more than 23 kinds of familiar fruits and vegetables. Recognition correctness was higher than 93.36%. In addition, the results of picking sequence planning showed that its planning way was consistent with artificial expertise from manual picking operation.

Key words: harvesting robot; selective attention; target recognition; sequence planning

0 Introduction

Harvesting robots are a class of perceptive machines that can be programmed to perform a variety of agricultural tasks such as transplanting, cultivating, spraying, trimming and selective harvesting[1].

In recent years, most related researches are about picking a specific type fruit or vegetable, such as apple or citrus[2-9]. And they always use visual sensors to analyze the difference between fruits and their background in color, shape and texture. Bayesian decision, artificial neural network and support vector machine are also adopted to recognize a specific kind of fruit or vegetable[10-13]. Hence, these methods' extension ability is limited. The current researches also rarely discuss about planning fruit cluster (which means a cluster of fruits adhering or overlapping together, a single fruit is a special case of fruit cluster) picking sequence, which makes harvesting robots difficult to be applied to the actual.

By integrating several features of targets and their scene, animals can use visual selective attention to choose their interested targets quickly in complex background and to change the focus of attention though targets' saliency[14]. This mechanism not only has adaptability and interference rejection for complex environment, but also has similarities with harvesting robots to recognize targets and to plan picking sequence.

A bionic method was designed to recognize fruit clusters and to plan their picking sequence by simulating visual selective attention mechanism and improving Itti's visual attention computational model[15-16]. On one hand, it can increase fruit cluster algorithm adaptive capacity for complex agriculture environment and versatility for the algorithm to recognize more kinds of fruits and vegetables; on the other hand, by planning fruit clusters picking sequence

Received date: 2016-04-17 Accepted date: 2016-05-17
Supported by National Natural Science Foundation of China (Grant No. 51375368), Agricultural Science and Technology Innovation and Key Project of Shaanxi Province, China (Grant No. 2015NY008)
Corresponding author: Wang Sun'an, Professor. E-mail: sawang@mail.xjtu.edu.cn. Tel: +86-029-82669321.

ahead, it will decrease repetitive motions and improve work efficiency.

1 Visual selective attention mechanism and its computational model

Primates can select the interested or the most important target quickly from huge visual data by using visual selective attention mechanism. In other words, through attentional process, the focus of attention can be concentrated on a specific location, object or information[17].

Psychological researches show that this mechanism contains data-driven and target-driven processes[17-18]. In data-driven process, feature detector accepts different environment stimuli and combines several visual features (color, motion and so on) together to form the characterizations of objects and their scene. In target-driven process, visual cortex can amend the characterizations in terms of priori knowledge and task expectation.

According to this visual cognitive process, ITTI[15] designed a data-driven visual attention computational model, which combined psychological model using feature integration theory[19] and conceptual model proposed by KOCH and ULLMAN[14] together. In this model, multiple scale features (color, intensity and orientation) and center - surroud difference were adopted to generate conspicuity maps. Then conspicuity maps could be combined into the saliency map. Depending on saliency, this model also used the biological neural network competition mechanism called Winner-Take-All to shift the focus of attention for concerning several targets in visual field successively[20-21]. However, due to the lack of target-driven process[20], Itti computational model only can find the locations which have significant difference to surrounding environment.

In agricultural harvesting process, an operation plant should be chosen firstly from a large amount of plants. And then several fruit clusters in the plant can be recognized, and the harvesting sequence will be planned. Finally, a single fruit can be positioned and isolated from the cluster, which will prepare for picking manipulator to achieve harvesting task.

This paper focuses on fruit cluster recognition and its picking sequence planning in the second process of the above. By introducing familiar fruits and vegetables visual features and harvesting requirement, priori knowledge was supplied to the computational process. According to artificial picking expertise, conspicuity maps and saliency map generated methods were improved to adapt to the actual need in picking task (Fig. 1).

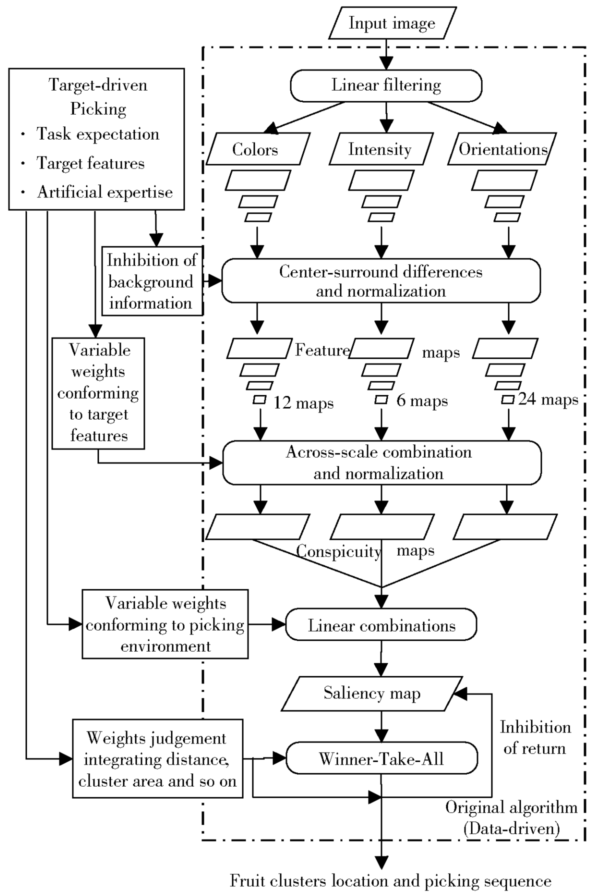


Fig. 1 Improved algorithm sketch of visual selective attention for picking task

2 Variable weight fruit cluster recognition algorithm

2.1 Color feature maps generated method conforming to operation object characteristics

Itti computational model used intensity I, colors (broadly-tuned red channel R, green channel G, blue channel B and yellow channel Y) and orientation O_n ($n = 0°$, $45°$, $90°$ and $135°$) as early visual features. And the four color channels were defined as[15]

$$\begin{cases} R = r - (g + b)/2 \\ G = g - (r + b)/2 \\ B = b - (r + g)/2 \\ Y = (r + g)/2 - |r - g|/2 - b \end{cases} \quad (1)$$

where r, g and b are the red, green and blue channels information of input image.

When extracting color features, obvious contrast area in R/G or B/Y could be found though color antagonism theory (these basic visual perceptual units come in pairs. When one is inhibited, the other one will be activated).

In harvesting environment, green and blue always mean background, such as leaves and sky. And the most kinds of mature fruits (less vegetables) are red or yellow. Based on this color feature priori knowledge, the method inhibited the nerve which could perceive green or blue and only activated the nerve which could perceive red or yellow. Thus, color feature maps generated method would be

$$RG(c,s) = \begin{cases} RG'(c,s) & (RG'(c,s) \geqslant 0) \\ 0 & (RG'(c,s) < 0) \end{cases} \tag{2}$$

$$YB(c,s) = \begin{cases} YB'(c,s) & (YB'(c,s) \geqslant 0) \\ 0 & (YB'(c,s) < 0) \end{cases} \tag{3}$$

Among them

$$RG'(c,s) = (R(c) - G(c)) \circ (G(s) - R(s)) \tag{4}$$

$$YB'(c,s) = (Y(c) - B(c)) \circ (B(s) - Y(s)) \tag{5}$$

where $RG(c,s)$ is red/green feature image. $YB(c,s)$ is yellow/blue feature image. $RG'(c,s)$ is the comparative results between red and green features. $YB'(c,s)$ is the comparative results between yellow and blue features. c, s are the scale c, s in Gaussian pyramid, $c \in \{2, 3, 4\}$, $\delta \in \{3, 4\}$, $s = c + \delta$. $\circ$ is a subtraction operator, which uses linear interpolation to make the image at scale c have the same resolution as the image at scale s, then the two images can have a subtraction. $R(c)$ and $R(s)$ are red information at scale c and s. $G(c)$ and $G(s)$ are green information at scale c and s. $Y(c)$ and $Y(s)$ are yellow information at scale c and s. $B(c)$ and $B(s)$ are blue information at scale c and s.

2.2 Conspicuity and saliency maps generated method

Using multiple feature integration to recognize target like people, Itti basic model has improved the algorithm adaptability to complex environment. However, this model only imitates data-driven process. It cannot recognize a specific target or amend recognition results with task expectation. Sometimes, noise may be even introduced to the maps.

As shown in Fig. 2a, this kind of apple has lighter red, which makes $RG(c,s)$ low. However, pixels in the cycle has more yellow information, which makes $YB(c,s)$ large. So if $RG(c,s)$ and $YB(c,s)$ are added in equal proportions using Itti basic model, noise will be introduced to color conspicuity map $\overline{C}$ as shown in Fig. 2b.

To adapt to different kinds of fruits and vegetables, variable weights were introduced into generating color conspicuity map. The values could be computed by supervised learning method, and would be adjusted dynamically according to the current picking object color feature. Hence

$$\overline{C} = \bigoplus_{c=2}^{4} \bigoplus_{s=c+3}^{c+4} [\omega_{RG} N(RG(c,s)) + \omega_{YB} N(YB(c,s))] \tag{6}$$

where $\oplus$ is an operation symbol, which makes the two images all at scale 4 add point-to-point. ω_{RG} is red/green feature weight. $N(\ )$ is a standardized operator[15]. ω_{YB} is yellow/blue feature weight.

As shown in Fig. 2c, based on this kind of apple's color feature, the method not only keeps the effect of $YB(c,s)$, but also reduces noise when $\omega_{RG} : \omega_{YB} = 0.7 : 0.3$.

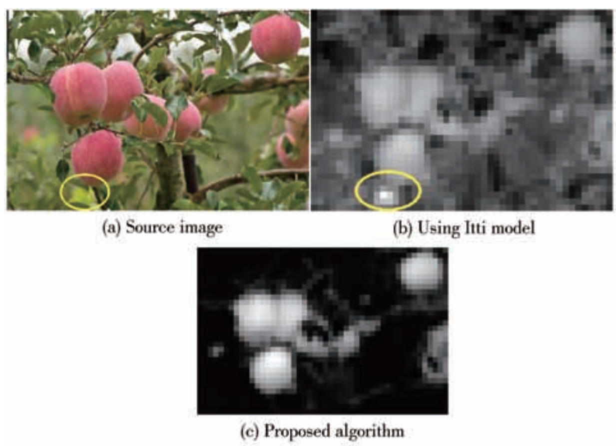


Fig. 2 Comparative experiment results of color conspicuity map

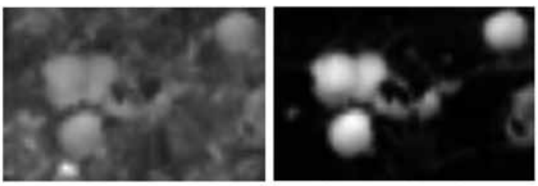
Similarly, when Itti basic model is used to add intensity, color and oritantion conspicuity maps ($\overline{I}$, $\overline{C}$ and $\overline{O}$) in equal proportions, high bright sky or other background will be introduced into saliency map S from $\overline{I}$. Hence, when computing S, weights were also used in the algorithm, and the values could be adjusted timely depending on the current light condition. Therefore

$$S = \omega_I N(\overline{I}) + \omega_C N(\overline{C}) + \omega_O N(\overline{O}) \tag{7}$$

where ω_I is intensity conspicuity map weight. $\overline{I}$ is intensity conspicuity map. ω_C is color conspicuity map *weight*. $\overline{C}$ is color conspicuity map. ω_O is orientation conspicuity map weight. $\overline{O}$ is intensity conspicuity map.

As shown in Fig. 3, compared with Itti basic model, the algorithm using variable weights makes fruit clusters more prominent in saliency map (In Fig. 3b, ω_I: ω_C: ω_O is 0.25:0.5:0.25).

Because the boundary between foreground and background in saliency map is blurry, it will be hard to use fixed threshold segmentation method. Otsu method[5], which used dynamic threshold, was adopted to extract fruit cluster information here. The results of segmenting fruit clusters out and eliminating too small area or too far clusters are shown in Fig. 4.



(a) Using Itti model (b) Proposed algorithm

Fig. 3 Comparison of saliency map about Fig. 2a

(a) Using Otsu (b) Eliminating too small area

Fig. 4 Results of fruit clusters recognition in Fig. 2a

3 Picking sequence planningbased on Winner-Take-All

Winner-Take-All is a biological neural network competition mechanism[14-15]. Neurons in this mechanism and pixels in saliency map S have one-to-one correspondence, and the neurons are independent of each other. These neurons can generate different strength electrical signals because of different saliency signal stimuli, and they will begin to accumulate. Only the neuron who is the first to reach threshold voltage can be the winner and the current focus of attention. It will not join to compete again. The other neurons are inhibited and reset until the next competition starts, they will be reactivated. Hence, this process can achieve shifting focus of attention.

Higher salicency means stronger electrical signal, so attention sequence is always from the highest to the lowest depended on saliency. However, when people pick fruits, they will not simply consider saliency.

Our surveys in Weinan city and Xianyang city, Shaanxi province show that, picking works usually start from the nearest cluster and then gradually move further and further away. For reducing repeating action, they will pick the nearest cluster to current picking cluster. Meanwhile, in order to prevent fruits from falling by shake, the cluster with more fruits will be harvested first. Hence, picking task is a process which takes clusters distance and fruit quantity into account, and also a process which is from near to far and from more to less.

Based on above expert experience, fruit cluster picking sequence can be regarded as a competition among clusters. In monocular vision, when the cluster has more fruits and much closer to observation point, its area S'_i will be larger. Thus, the proposed algorithm stipulated that planning work started from the most salient cluster. If j stood for all clusters which needed to pick, the picking sequence would be

$$F_{i+1} = \max_j (\omega_s S'_j + \omega_{SM} V_{SMj} - \omega_d d_{ij}) \tag{8}$$

where F_{i+1} is the next picking cluster. ω_s is cluster area weight. S'_j is the normalized area of each clusters, which wait for picking. ω_{SM} is saliency weight. V_{SMj} is the most salient point in cluster j. ω_d is distance weight. d_{ij} is the normalized Euclidean distance between the most salient point in last determined cluster i and the most salient point in undetermined clusters j.

Weights in Eq. 8 can be decided by different picking expert experiments for different kinds of fruits and vegetables.

As shown in Fig. 5, compared with Winner-Take-All mechanism which only uses saliency, the weighted preferential picking sequence strategy using distance, area and saliency is more aligned with artificial picking experience and reduces repetitive movements better.

(a) Using Winner-Take-All (b) ω_s:ω_{SM}:ω_d is 0.2:0.3:0.5

Fig. 5 Comparative results of picking sequence planning

4 Experiment results and analysis

For verifying the method application effects in practical agricultural environment, 10 images in each 23 kinds of fruits and vegetables (only Huanghemi has 6 images) from the web were tested. These images could imitate single visual sensor to acquire image, so they were used to analyze monocular vison effect to recognize fruit clusters and plan picking sequence.

4.1 Cluster recognition results and analysis

As shown in Fig. 6, introducing fruits and vegetables color feature as priori knowledge guarantees that the more salient areas will all be clusters, when the saliency map is generated with multiple visual features. What's more, this method improves the signal-to-noise ratio obviously.

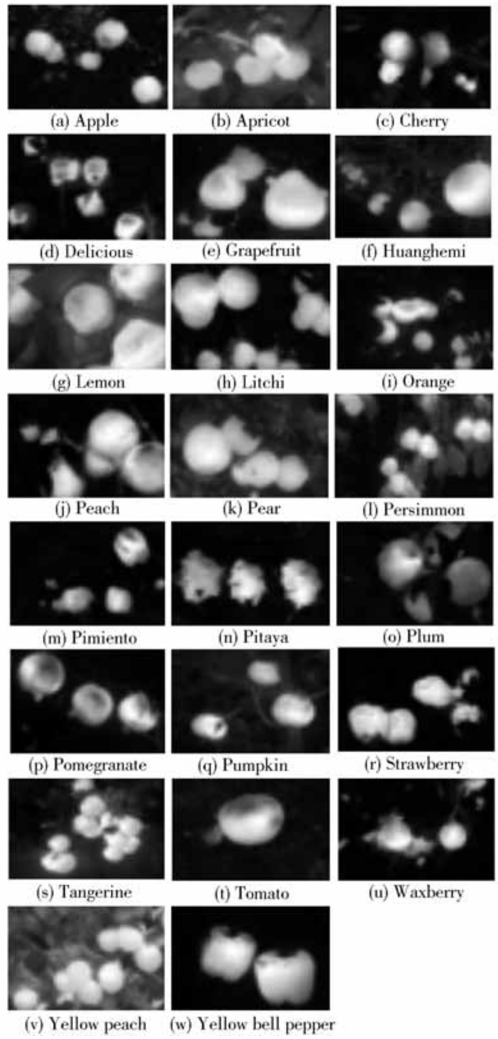
Fig. 7 is on the basis of Fig. 6. It is the fruit clusters images after using Otsu to make dynamic threshold segmentation. In these images, very few branches and leaves are mistakenly classified, so recognition results are much better than basic Itti model[15] (Fig. 8). Statistical results in Tab. 1 show that, the proposed algorithm recognizes clusters accurately except in few images where branches and leaves color or shape are too close to fruits. Average recognition correctness is higher than 93.36%.

Compared with the method only using single feature like chromatic aberration[10,22], the algorithm on one hand improves fruit cluster recognition accuracy, which will provide noise-free fruit cluster location information for planning picking sequence; on the other hand, it has recognized more than 23 kinds of fruits and vegetables so that it has better generality and adaptability to satisfy practical picking work needs.

4.2 Picking sequence planning results and analysis

As shown in Fig. 9, the weighted preferential picking sequence strategy using distance, area and saliency makes picking work start from the most saliency cluster. The sequence after that has balanced cluster saliency, fruit quantity and the distance to the last operation position.

Compared with Winner-Take-All which only uses saliency or the picking sequence planning method and only considers distance between fruits[23], the proposed method is more aligned with artificial picking experience and closer to real picking process. Using saliency to find the first operation object can be relevant for people to choose the focus of attention.



(a) Apple (b) Apricot (c) Cherry (d) Delicious (e) Grapefruit (f) Huanghemi (g) Lemon (h) Litchi (i) Orange (j) Peach (k) Pear (l) Persimmon (m) Pimiento (n) Pitaya (o) Plum (p) Pomegranate (q) Pumpkin (r) Strawberry (s) Tangerine (t) Tomato (u) Waxberry (v) Yellow peach (w) Yellow bell pepper

Fig. 6 Saliency maps of 23 kinds of fruits and vegetables using the proposed algorithm

Introducing distance makes sure that when harvesting robot picked fruits, it will have a short motion distance and high efficiency. Adopting area makes picking work start from the clusters with more fruits to the clusters with less fruits, which reduces fruit damage rate.

4.3 Algorithm execution time

The algorithm was programmed on Matlab R2012a. Computer configuration: processor was Intel (R), Core(TM) i3-2120 CPU @ 3.30 Hz; memory was 4.00 GB.

As shown in Tab. 1, average execution time for 226 images using the proposed algorithm is 967.7 ms. The shortest time is 872.2 ms, and the longest is 1 070 ms. It has a similar executing time with the fast tracing recognition method for apple[5].

Because the calculation using feature integration to establish saliency map is invariable, the main reason of time differences is the fruit quantities and cluster areas difference. In other words, execution time will be

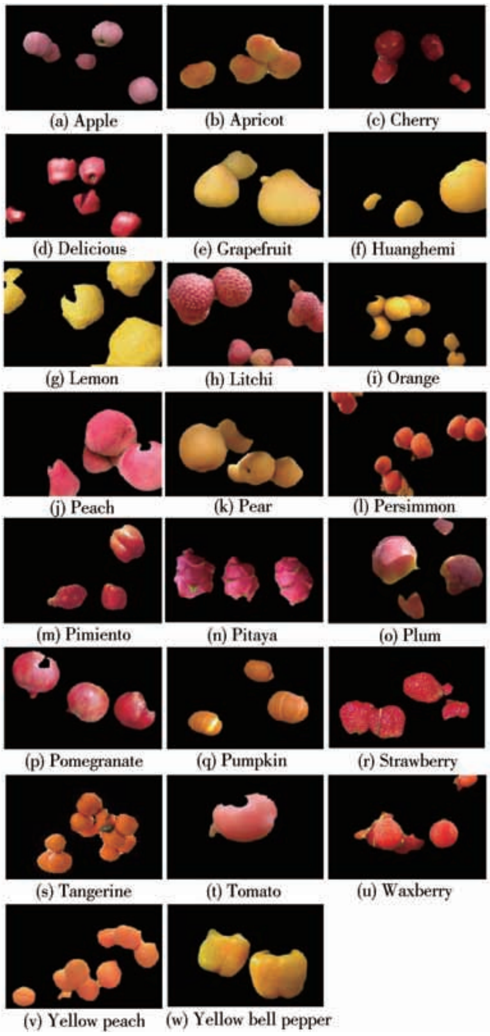


Fig. 7 Segmentation results by Otsu

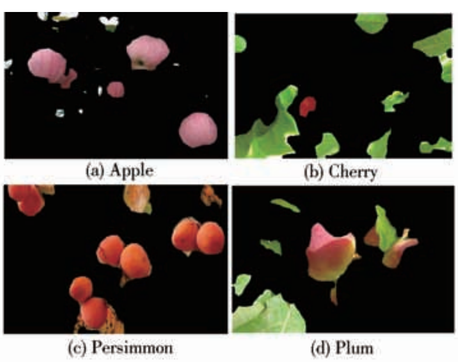


Fig. 8 Recognition results using Itti visual attention computational model

longer, if there is more clusters in image, which need more time to calculate connected domain area, eliminate too small area and get the edge information about connected domain and so on.

5 Conclusions

(1) A new method for fruit cluster recognition and picking sequence planning was proposed, by imitating

Tab. 1 Recognition correctness and whole execution time of 23 kinds of common fruits and vegetables

Name	Recognition correctness/%	Execution time/ms
Apple	100	973. 9
Apricot	90	971. 5
Cherry	90	970. 7
Delicious	90	987. 1
Grapefruit	80	975. 6
Huanghemi	83. 3	985. 0
Lemon	90	970. 8
Litchi	100	975. 1
Orange	100	965. 0
Peach	100	968. 5
Pear	90	958. 4
Persimmon	90	966. 0
Pimiento	100	963. 4
Pitaya	80	967. 1
Plum	90	962. 0
Pomegranate	90	967. 3
Pumpkin	100	953. 4
Strawberry	100	951. 9
Tangerine	100	976. 3
Tomato	100	962. 5
Waxberry	80	961. 9
Yellow peach	100	959. 5
Yellow bell pepper	100	970. 0

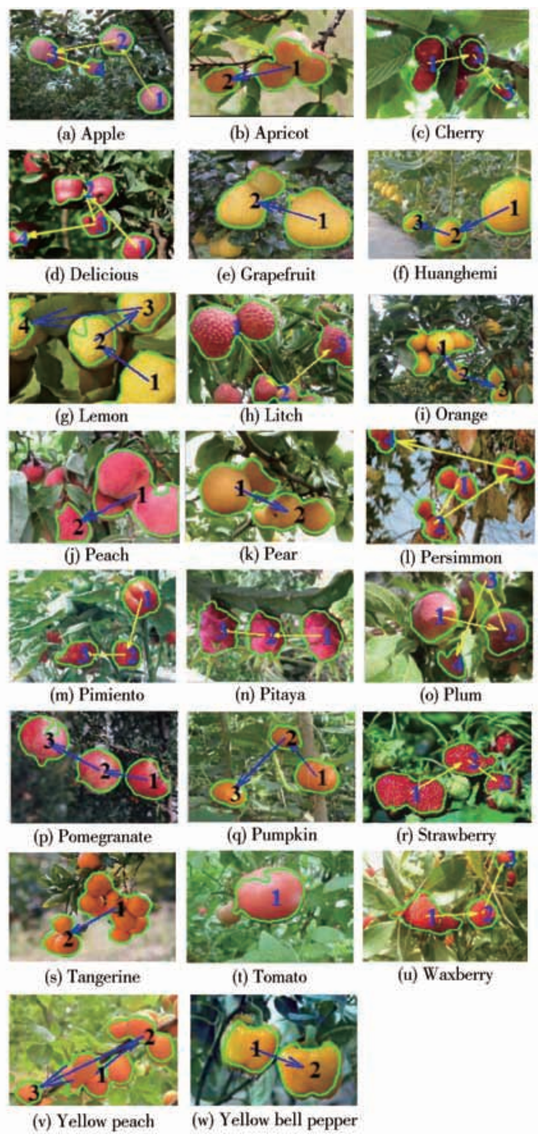


Fig. 9 Results of picking sequence planning

primates' visual selective attention mechanism, using different picking objects common features as priori knowledge and introducing picking experts operational experience. The results show that the method has achieved recognition of more than 23 kinds of familiar fruits and vegetables. Picking sequence planning result is consistent with artificial expertise from manual picking operation.

(2) Compared with visual attention model using only data-driven or fruit recognition method using only single feature, the proposed method not only adapts to the complex agricultural environment better, but also improves algorithm versatility and the ability to recognize more kinds of picking objects.

(3) Based on the above research, further study will be needed to segregate a single fruit from the cluster, plan picking sequence for individual fruits and locate them. This can provide more accurate information for harvesting robot to achieve concrete operations.

References

[1] EDAN Y, MILES G E. Systems engineering of agricultural robot design [J]. IEEE Transactions on Systems Man & Cybernetics, 1994, 24(8): 1259 – 1265.

[2] KAPACH K, BARNEA E, MAIRON R, et al. Computer vision for fruit harvesting robots state of the art and challenges ahead [J]. International Journal of Computational Vision and Robotics, 2012, 3(1/2): 4 – 34.

[3] SONG Jian, ZHANG Tiezhong, XU Liming, et al. Research actuality and prospect of picking robot for fruits and vegetables[J]. Transactions of the Chinese Society for Agricultural Machinery, 2006, 37(5): 158 – 162. (in Chinese)

[4] WANG Rujing, SUN Bingyu. Development status and expectation of agricultural robot[J]. Bulletin of Chinese Academy of Sciences, 2015, 30(6): 803 – 809. (in Chinese)

[5] LV Jidong, ZHAO Dean, JI Wei. Fast tracing recognition method of target fruit for apple harvesting robot [J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(1): 65 – 72. (in Chinese)

[6] NGUYEN T T, VANDEVOORDE K, WOUTERS N, et al. Detection of red and bicoloured apples on tree with an RGB – D camera[J]. Biosystems Engineering, 2016, 146: 33 – 44.

[7] JI W, ZHAO D, CHENG F, et al. Automatic recognition vision system guided for apple harvesting robot [J]. Computers & Electrical Engineering, 2012, 38(5): 1186 – 1195.

[8] TANIGAKI K, FUJIURA T, AKASE A, et al. Cherry-harvesting robot [J]. Computers and Electronics in Agriculture, 2008, 63(1): 65 – 72.

[9] LI P, LEE S H, HSU H Y. Study on citrus fruit image data separability by segmentation methods[J]. Procedia Engineering, 2011, 23(5): 408 – 416.

[10] QIAN Jianping, YANG Xinting, WU Xiaoming, et al. Mature apple recognition based on hybrid color space in natural scene[J]. Transactions of the CSAE, 2012, 28(17): 137 – 142. (in Chinese)

[11] YANG Guobin, ZHAO Jiewen, XIANG Zhongping. Discriminating tomatoes based on computer vision technology in natural outdoor scenes [J]. Journal of Agricultural Mechanization Research, 2003(1): 60 – 62. (in Chinese)

[12] JIMÉNEZ A R, CERES R, PONS J L. A vision system based on a laser range-finder applied to robotic fruit harvesting [J]. Machine Vision and Applications, 2000, 11(6): 321 – 329.

[13] MA Xiaodan, LIU Gang, ZHOU Wei, et al. Apple recognition based on fuzzy neural network and quantum genetic algorithm [J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(12): 227 – 232, 251. (in Chinese)

[14] KOCH C, ULLMAN S. Shifts in selective visual attention: towards the underlying neural circuitry[J]. Human Neurobiology, 1985, 4(4): 219 – 246.

[15] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254 – 1259.

[16] ITTI L, KOCH C. Computational modelling of visual attention[J]. Nature Reviews Neuroscience, 2001, 2(3): 194 – 203.

[17] GOLDSTEIN E B. Cognitive psychology (Third edition) [M]. Zhang Ming. Beijing: China Light Industry Press, 2015. (in Chinese)

[18] MYERS D. Psychology (Ninth edition) [M]. Huang Xiting. Beijing: Posts & Telecom Press, 2013. (in Chinese)

[19] TREISMAN A M, GELADE G. A feature-integration theory of attention[J]. Cognitive Psychology, 1980, 12(1): 97 – 136.

[20] LI Wanyi, WANG Peng, QIAO Hong. A survey of visual attention based methods for object tracking[J]. Acta Automatica Sinica, 2014, 40(4): 561 – 576. (in Chinese)

[21] BORJI A, ITTI L. State-of-the-art in visual attention modeling[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 185 – 207.

[22] MENG Jiao, WANG Sun'an. The recognition of overlapping apple fruits based on boundary curvature estimation[C]//2015 Sixth International Conference on Intelligent Systems Design & Engineering Applications, 2015: 874 – 877.

[23] TANG Xiuying, ZHANG Tiezhong. Obstacle avoidance trajectory planning of cucumber harvesting robot [J]. Journal of Agricultural Mechanization Research, 2010, 32(5): 98 – 102. (in Chinese)

doi:10.6041/j.issn.1000-1298.2016.11.001

基于选择性注意机制的果实簇识别与采摘顺序规划

王冰心　王孙安　于德弘
（西安交通大学机械工程学院，西安 710049）

摘要：基于仿生学思想设计了一种果实簇识别与采摘顺序规划方法。该方法以 Itti 视觉注意基本模型为基础，通过改进视觉特征整合方法，构建依赖先验知识的视觉显著图，改善果实簇的识别效果；借鉴“赢者取全”的生物神经竞争机制和采摘专业人员的操作经验，设计距离、面积和显著度加权择优的采摘顺序规划策略，提高采摘工作效率。试验结果表明，设计的算法能够识别 20 余种常见果蔬，识别正确率达到 93.36%；果实簇的采摘顺序规划结果符合人工采摘的专家经验。

关键词：采摘机器人；选择性注意；目标识别；顺序规划
中图分类号：TP242.6　　**文献标识码**：A　　**文章编号**：1000-1298(2016)11-0001-07

Fruit Cluster Recognition and Picking Sequence Planning Based on Selective Attention

Wang Bingxin　Wang Sun'an　Yu Dehong
(*School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an* 710049, *China*)

Abstract: To improve algorithm versatility and picking efficiency, recognizing more kinds of fruits or vegetables and planning picking sequence for fruit clusters in visual field have become hotspots and trends of harvesting robot research. Primates can find objects and shift attention through visual selective attention mechanism, which has similarities with harvesting robots to recognize targets and to plan picking sequence. Hence, by adopting the idea and technology of bionics, a new visual selective attention-based method was proposed for fruit cluster recognition and picking sequence planning. According to Itti visual attention computational model, this algorithm changed computing process for color feature maps, and improved feature integration method for building color conspicuity map and saliency map by introducing priori knowledge about fruits and vegetables. Using the reference of the biological neural network competition mechanism called Winner-Take-All and artificial expertise from professional picking operators, distance, area and saliency were adopted to design the weighted preferential picking sequence strategy for fruit clusters. The experimental results showed that the proposed method in this paper achieved recognition of more than 23 kinds of familiar fruits and vegetables. Recognition correctness was higher than 93.36%. In addition, the results of picking sequence planning showed that its planning way was consistent with artificial expertise from manual picking operation.

Key words: harvesting robot; selective attention; target recognition; sequence planning

引言

采摘机器人是一类针对水果和蔬菜，可以通过编程来完成采摘等相关作业任务的具有感知能力的自动化机械收获装置[1]。

近年来，相关领域的研究多是关于某一特定品

收稿日期：2016-04-17　修回日期：2016-05-17
基金项目：国家自然科学基金项目(51375368)和陕西省农业科技创新与攻关项目(2015NY008)
作者简介：王冰心(1990—)，女，博士生，主要从事机器视觉与三维全场测试技术研究，E-mail：icywangbx@126.com
通信作者：王孙安(1957—)，男，教授，博士生导师，主要从事智能与仿生机器人及其应用研究，E-mail：sawang@mail.xjtu.edu.cn

种的果实采摘问题(如苹果、柑橘)[2-9]。感知阶段多利用视觉传感器,通过分析果实与背景环境的颜色、形状或纹理差异,结合贝叶斯决策、人工神经网络、支持向量机等方法实现对单一品种果蔬的识别[10-13],拓展能力有限。研究中也鲜有规划成簇果实(多个粘连或重叠在一起的果实称为一个果实簇,单个果实是果实簇的特例)采摘顺序的成果,使得采摘机器人难以应用于实际。

生物视觉的选择性注意机制通过结合目标及目标所处场景的多种视觉特征,能够从复杂背景中快速选取出感兴趣的目标,并根据目标的显著性转移关注点位置[14]。这一机制不仅对复杂环境具有一定的适应性和抗扰能力,而且其目标筛选、关注点转换流程与采摘作业中的果实识别、采摘顺序规划过程具有一致性。

据此,本文通过模拟视觉选择性注意机制、改进Itti视觉注意基本模型[15-16],设计一种果实簇识别与采摘顺序规划的仿生学方法。一方面,用于提高果实簇识别算法对复杂农业环境的适应能力,提高算法对不同种类果实识别的通用性;另一方面,通过提前规划视域内多个果实簇的采摘顺序,减少采摘过程中的重复动作,提高作业效率。

1 视觉选择性注意机制及其计算模型

灵长类依赖视觉选择性注意机制,能够从获取的大量视觉信息中快速地筛选出最重要、最感兴趣的信息,即通过注意过程将关注点集中在一个特定的位置、客体或者信息上[17]。

心理学研究表明,该过程包括自下而上的数据驱动和自上而下的目标驱动[17-18]:数据驱动过程中,特征觉察器接收来自环境的不同刺激,通过结合多种视觉特征(颜色、运动状态等),形成对物体、场景的表征;目标驱动过程中,大脑更高级的视觉区根据已有的先验知识或任务期望,修正物体、场景的表征结果。

根据灵长类的这一视觉认知过程,并结合TREISMAN等提出的“特征整合理论”心理学模型[19]和KOCH、ULLMAN提出的概念模型[14],ITTI等在1998年提出了一种自下而上的视觉注意计算模型[15]。该模型不仅能够利用多尺度的颜色、亮度和方向特征,通过中央周边操作产生体现显著性度量的特征图,并可将这些单特征维的显著图合并为最终显著图;还可以利用“赢者取全”的生物神经竞争机制,根据显著性强弱,依次注视视域内的多个目标,切换关注点位置[20-21]。但是,由于Itti基本模型缺少自上而下的目标信息指导[20],因此该模型仅能发现与周边区域有明显特征差异的地方,无法根据目标特征确定关注点位置。

农业采摘作业过程中,首先需要从多个待作业植株中选择一个作业对象。然后对植株上的多个果实簇进行识别与采摘顺序规划。最后从当前采摘果实簇中,分离出单一果实并对其定位,指导采摘机械臂完成最终的采摘任务。

本文针对上述过程中对果实簇的识别与采摘顺序规划,通过引入常见果蔬的视觉特征及采摘作业的任务需求,为该模型的计算过程提供先验知识。根据人工采摘的专家经验,改进Itti注意计算模型的显著图生成方法及关注点切换策略,以适应采摘作业的实际需求(图1)。

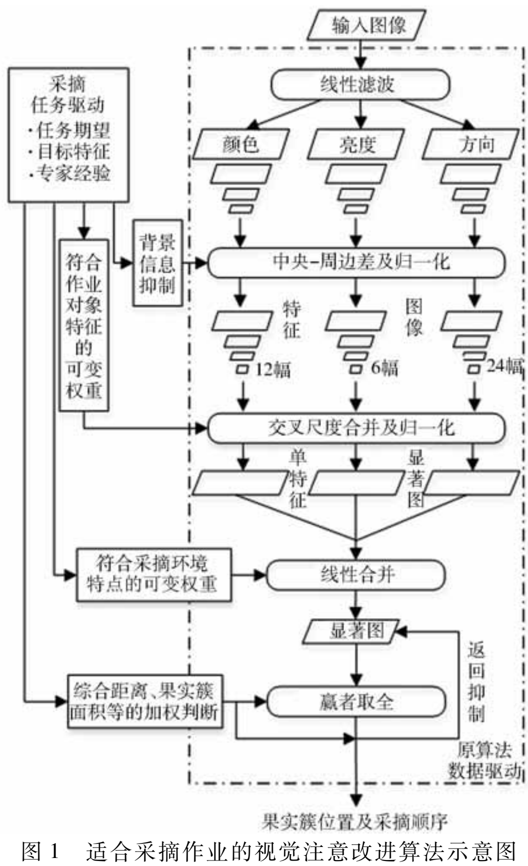


图1 适合采摘作业的视觉注意改进算法示意图

Fig. 1 Improved algorithm sketch of visual selective attention for picking task

2 变权重的果实簇识别算法

2.1 适应作业对象特征的颜色特征图构建方法

Itti基本模型以亮度信息 I、颜色信息(红色信息 R、绿色信息 G、蓝色信息 B、黄色信息 Y)和方向信息 O_n($n=0°$、$45°$、$90°$、$135°$)作为初级图像特征,并定义4个颜色信息为[15]

$$\begin{cases} R = r - (g+b)/2 \\ G = g - (r+b)/2 \\ B = b - (r+g)/2 \\ Y = (r+g)/2 - |r-g|/2 - b \end{cases} \tag{1}$$

式中　r、g、b——输入图像的红、绿、蓝 3 个通道值

提取颜色特征时，通过颜色拮抗理论（成对出现的视觉感应基本单元，其中一个被抑制时，另一个被激活），提取 R/G、B/Y 对比明显的区域，也就是与周围颜色差异大的区域。

由于采摘作业环境中的绿色和蓝色多为枝叶、天空等背景信息，且多数成熟果实（少数蔬菜）属红、黄色系，因此依据人对果蔬颜色特征的先验知识，采用抑制大脑中感应绿色、蓝色神经元的方法，使其不被激活，只提取红色、黄色的果蔬信息，将颜色特征图的构建方法修改为

$$RG(c,s)=\begin{cases}RG'(c,s) & (RG'(c,s)\geq 0)\\ 0 & (RG'(c,s)<0)\end{cases} \tag{2}$$

$$YB(c,s)=\begin{cases}YB'(c,s) & (YB'(c,s)\geq 0)\\ 0 & (YB'(c,s)<0)\end{cases} \tag{3}$$

其中

$$RG'(c,s)=(R(c)-G(c))\circ(G(s)-R(s)) \tag{4}$$

$$YB'(c,s)=(Y(c)-B(c))\circ(B(s)-Y(s)) \tag{5}$$

式中　$RG(c,s)$——红/绿特征图像信息

$YB(c,s)$——黄/蓝特征图像信息

$RG'(c,s)$——红/绿特征对比计算结果

$YB'(c,s)$——黄/蓝特征对比计算结果

c、s——高斯金字塔第 c、s 层的图像，$c\in\{2,3,4\}$，$\delta\in\{3,4\}$，$s=c+\delta$

$\circ$——通过线性插值，使第 s 层图像具有第 c 层的分辨率，然后对 2 幅图像进行像素点间的减法操作算子

$R(c)$、$R(s)$——第 c、s 层的红色信息

$G(c)$、$G(s)$——第 c、s 层的绿色信息

$Y(c)$、$Y(s)$——第 c、s 层的黄色信息

$B(c)$、$B(s)$——第 c、s 层的蓝色信息

2.2　变权重的显著图生成

Itti 基本模型通过模拟人利用视觉识别目标时的多特征整合方法，提高了识别算法对复杂环境的适应性。但是，该模型仅模拟了自下而上的数据驱动，无法在识别特定目标时依据任务期望修正识别结果，甚至会引入噪声。

如图 2a 所示，该品种苹果的红色较弱，使得 $RG(c,s)$ 整体偏小，但属于背景的圈内部分含有较高的黄色信息，使得 $YB(c,s)$ 在该区域的值过大。这时若采用基本模型，等比例合并 $RG(c,s)$ 和 $YB(c,s)$，则颜色特征显著图 $\overline{C}$ 中会引入如图 2b 所示的噪声。

为适应不同果蔬的特点，在构建颜色特征显著图时引入了权重。权值可通过有监督的学习方法计算得到，并根据当前作业对象颜色特征进行动态调整。则

$$\overline{C}=\bigoplus_{c=2}^{4}\bigoplus_{s=c+3}^{c+4}\left[\omega_{RG}N(RG(c,s))+\omega_{YB}N(YB(c,s))\right] \tag{6}$$

式中　$\oplus$——将图像均缩放至高斯金字塔的第 4 层，然后进行点对点相加的运算符号

ω_{RG}——红/绿特征的权重

$N(\,)$——标准化操作算子[15]

ω_{YB}——蓝/黄特征的权重

如图 2c 所示，根据该品种苹果的颜色特征，$\omega_{RG}:\omega_{YB}$ 为 0.7∶0.3 时，不仅可以保留 $YB(c,s)$ 的作用，也削弱了噪声的影响。

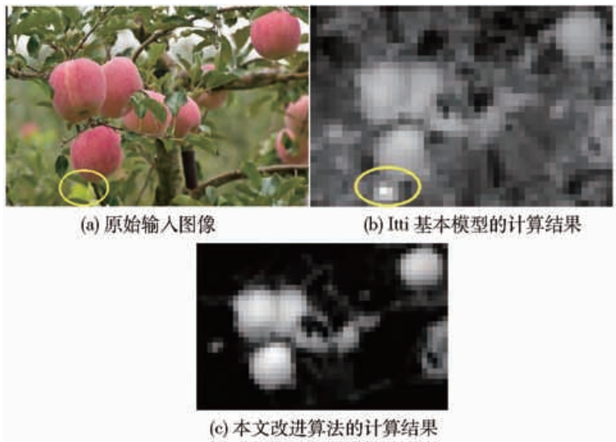

(a) 原始输入图像　(b) Itti 基本模型的计算结果　(c) 本文改进算法的计算结果

图 2　颜色特征显著图的效果对比

Fig. 2　Comparative experiment results of color conspicuity map

同理，用基本模型等比例合并亮度、颜色和方向显著图 $\overline{I}$、$\overline{C}$ 和 $\overline{O}$ 时，过亮的天空等背景噪声也会从 $\overline{I}$ 进入最终显著图 S。因此，在构建 S 的过程中也引入了权重。权值大小可根据采摘作业环境的当前光照条件实时调整，则

$$S=\omega_I N(\overline{I})+\omega_C N(\overline{C})+\omega_O N(\overline{O}) \tag{7}$$

式中　ω_I——亮度特征显著图的权重

$\overline{I}$——亮度特征显著图

ω_C——颜色特征显著图的权重

$\overline{C}$——颜色特征显著图

ω_O——方向特征显著图的权重

$\overline{O}$——方向特征显著图

如图 3 所示，本文的变权重改进算法较 Itti 基本模型构建的最终显著图更能够突出果实簇区域（图 3b 中 $\omega_I:\omega_C:\omega_O$ 为 0.25∶0.5∶0.25）。

由于最终显著图中前、后景的边界模糊，不宜采用固定阈值分割图像，因此采用 Otsu 法[5]进行动态阈值分割，提取果实簇信息。分割后的果实簇图像，以及剔除面积过小、距离过远果实簇后的识别结果如图 4 所示。

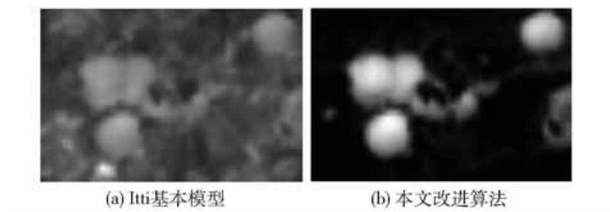


图 3 针对图 2a 的最终显著图效果对比

Fig. 3 Comparison of saliency map about Fig. 2a

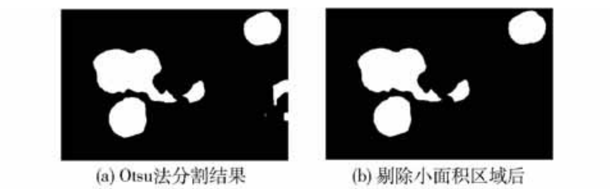


图 4 针对图 2a 的果实簇识别结果

Fig. 4 Results of fruit clusters recognition in Fig. 2a

3 基于“赢者取全”的采摘顺序规划

“赢者取全”是一种生物学上的神经竞争机制[14-15]。该机制的神经元与最终显著图 S 上的点一一对应，且相互独立。这些神经元在不同显著强度的信号激励下形成不同强度的电信号，并开始累积。只有最先达到阈值电压的神经元（赢者）才能成为当前的关注点，而且之后不再参与竞争。其他神经元则被抑制（复位），直到下一轮竞争开始时被重新激活，再次竞争赢者，从而实现关注点的切换。

由于显著度越高，生成的神经元电信号越强，因此“赢者取全”的实际关注顺序往往是根据显著度由高至低排列的。但在人工采摘作业时，依靠的并不仅是显著度。

在陕西省渭南市、咸阳市等我国苹果主产区调查发现，采摘作业一般从最近的果实簇开始，逐渐向远处进行。为减少重复动作，会先采摘距当前果实簇较近的一簇果实。同时，为了防止在采摘过程中其他果实因枝叶晃动掉落，会先摘果实较多的簇。所以，采摘作业实际上是一个综合考虑果实簇间距、簇内果实数量的由近到远、从多到少的过程。

根据上述专家经验，果实簇的作业顺序规划问题也可以看作一个簇与簇的竞争问题。由于在单目视觉中，果实簇距离观测点越近、簇中果实越多，在图像中果实簇的面积 S_i' 越大。因此，在算法中规定，对视域内多个果实簇的采摘从显著度最高的一个簇开始。若 j 表示尚未确定作业顺序的所有果实簇，则待作业果实簇的采摘顺序为

$$F_{i+1}=\max_{j}(\omega_s S_j'+\omega_{SM}V_{SMj}-\omega_d d_{ij}) \tag{8}$$

式中 F_{i+1}——下一作业果实簇

ω_s——果实簇面积所占权重

S_j'——尚未规划采摘顺序的各果实簇的归一化面积

ω_{SM}——显著性所占权重

V_{SMj}——第 j 簇中显著度最高点的值

ω_d——距离所占权重

d_{ij}——已确定的最后一个作业果实簇 i 中显著度最高点与未确定采摘顺序的果实簇显著度最高点的归一化欧氏距离

式（8）中的权值可根据不同品种果实采摘的专家经验获得。

如图 5 所示，采用距离、面积和显著度加权择优策略规划的采摘路径与仅依靠显著度的“赢者取全”机制相比，更符合人工操作习惯，能够更好地减少重复动作。

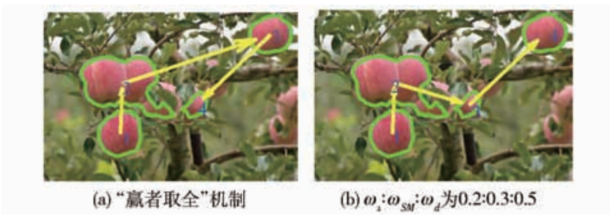


图 5 采摘顺序规划的结果对比

Fig. 5 Comparative results of picking sequence planning

4 试验与结果分析

为验证算法在实际农业作业环境中的效果，本文测试了源自于网络的 23 种果蔬的各 10 幅图像（黄河蜜为 6 幅），以模拟单个视觉传感器获取图像，从而分析单目视觉识别果实簇与规划采摘顺序的效果。

4.1 果实簇识别效果及分析

如图 6 所示，通过引入果蔬颜色特征作为先验知识，不仅保证了在构建多视觉特征的显著图时显著性高的区域均为果实簇，而且明显提高了识别信噪比。

图 7 是在图 6 基础上，通过 Otsu 法进行动态阈值分割后的果实簇区域图像。这些图像中，几乎没有果实簇周边的枝叶被误划入，识别效果明显优于标准 Itti 模型[15]（图 8）。表 1 的统计结果表明，除少数图像中由于枝叶与果实颜色、形状过于相近被误识别外，算法能够准确识别果实簇，且平均正确率高于 93.36%。

与仅采用色差等单一特征识别果实的算法[10,22]相比，本文算法一方面能够提高果实簇识别的正确率，为采摘顺序规划提供无噪声干扰的果实簇位置信息；另一方面本文算法能够识别至少 23 种常见果蔬，较识别单一品种果蔬的算法具有更好的

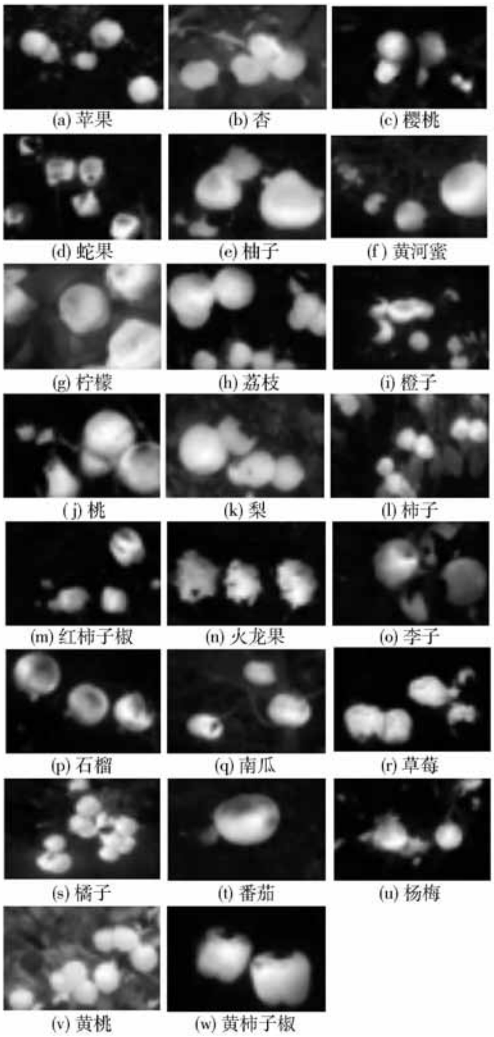


图 6　采用本文算法所得 23 种水果的最终显著图

Fig. 6　Saliency maps of 23 kinds of fruits and vegetables using the proposed algorithm

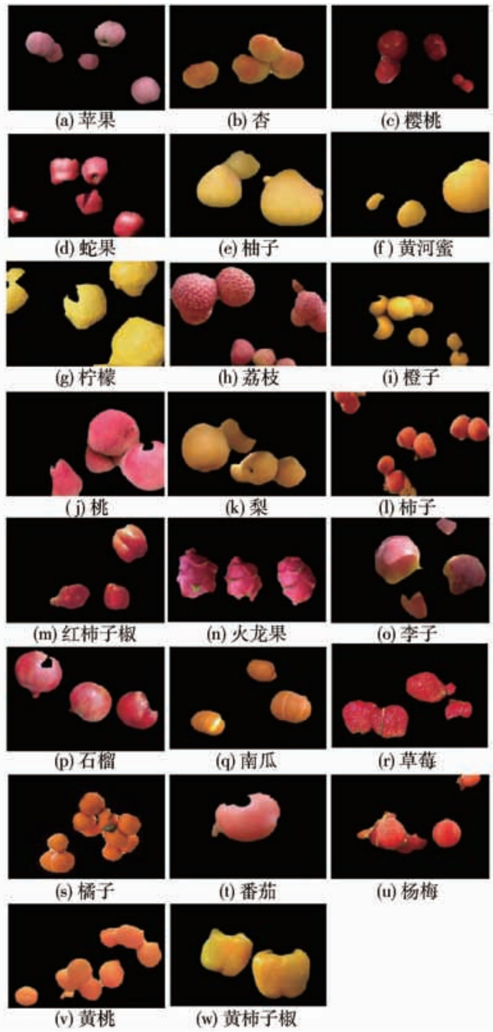


图 7　采用 Otsu 法所得分割结果

Fig. 7　Segmentation results by Otsu

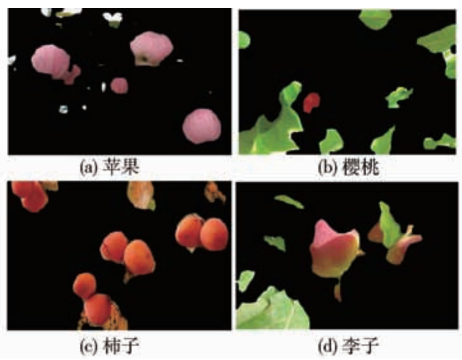


图 8　采用 Itti 标准模型的部分果实识别效果

Fig. 8　Recognition results using Itti visual attention computational model

通用性和适应性，更满足实际采摘的需求。

4.2　采摘顺序规划效果及分析

如图 9 所示，本文设计的距离、面积和显著度加权择优的“赢者取全”改进算法，使得采摘位置顺序从显著度最高的果实簇开始。之后确定作业位置时，兼顾了果实簇的显著度、果实数量，以及与上一作业位置间的距离。

与“赢者取全”机制仅依靠显著性和仅参考果实间距离的采摘顺序规划方法[23]相比，本文算法更符合采摘专家进行人工操作时的经验，贴近真实作业过程。显著性的引入符合人在选择注意点时的习惯，有助于选择第一个作业位置；距离的引入保证机器人采摘过程中运动距离短、效率高；面积的引入使得采摘成簇果实时，能够从果实多的簇开始，从而降低果实受损率。

4.3　算法执行时间

本文算法由 Matlab R2012a 编程开发，试验用计算机配置：处理器为 Intel(R)，Core(TM) i3－2120 CPU@ 3.30 Hz；内存 4.00 GB。

如表 1 所示，本文算法对 226 幅图的平均执行时间约 967.7 ms，最短执行时间约 872.2 ms，最长执行时间约 1 070 ms，与苹果果实的快速识别算法[5]执行时间相似。

由于通过特征整合获得最终显著图的计算量相同，因此产生时间差异的主要原因是果实簇的数量和面积的差异。也就是说，果实簇越多，通过显著图

表 1 23 种常见果蔬的识别正确率与算法执行时间

Tab. 1 Recognition correctness and whole execution time of 23 kinds of common fruits and vegetables

名称	识别正确率/%	执行时间/ms	名称	识别正确率/%	执行时间/ms
苹果	100	973.9	红柿子椒	100	963.4
杏	90	971.5	火龙果	80	967.1
樱桃	90	970.7	李子	90	962.0
蛇果	90	987.1	石榴	90	967.3
柚子	80	975.6	南瓜	100	953.4
黄河蜜	83.3	985.0	草莓	100	951.9
柠檬	90	970.8	橘子	100	976.3
荔枝	100	975.1	番茄	100	962.5
橙子	100	965.0	杨梅	80	961.9
桃	100	968.5	黄桃	100	959.5
梨	90	958.4	黄柿子椒	100	970.0
柿子	90	966.0			

计算连通域面积、剔除过小面积区域、获取连通域边缘信息等计算量就越大,因此算法的执行时间就越长。

5 结论

(1)借鉴生物视觉选择性注意机制,结合采摘对象共性特征作为先验知识,引入采摘专家的操作经验,设计了一种果实簇识别与采摘顺序规划的新方法。结果表明,算法实现了对至少 23 种常见果蔬的识别,并规划了符合人工操作习惯的多个果实簇的作业顺序。

(2)本文算法与采用常规自下而上的视觉注意机制模型、采用单一特征识别果实的算法相比,不仅更好地适应了复杂的农业作业环境,而且提高了算法识别多种不同类型采摘作业对象的能力和算法的通用性。

(3)在上述研究基础上,将针对果实簇中粘连果实分割及簇内单个果实的采摘顺序规划及定位问题做进一步研究,为采摘机器人的具体操作过程提供更为精准的目标信息。

图 9 采摘顺序规划结果

Fig. 9 Results of picking sequence planning

参 考 文 献

1 EDAN Y, MILES G E. Systems engineering of agricultural robot design[J]. IEEE Transactions on Systems Man & Cybernetics, 1994, 24(8): 1259 - 1265.

2 KAPACH K, BARNEA E, MAIRON R, et al. Computer vision for fruit harvesting robots state of the art and challenges ahead[J]. International Journal of Computational Vision and Robotics, 2012, 3(1/2): 4 - 34.

3 宋健,张铁中,徐丽明,等. 果蔬采摘机器人研究进展与展望[J]. 农业机械学报,2006,37(5):158 - 162.
SONG Jian, ZHANG Tiezhong, XU Liming, et al. Research actuality and prospect of picking robot for fruits and vegetables[J]. Transactions of the Chinese Society for Agricultural Machinery, 2006, 37(5): 158 - 162. (in Chinese)

4 王儒敬,孙丙宇. 农业机器人的发展现状及展望[J]. 中国科学院院刊,2015,30(6):803 - 809.
WANG Rujing, SUN Bingyu. Development status and expectation of agricultural robot[J]. Bulletin of Chinese Academy of Sciences, 2015, 30(6): 803 - 809. (in Chinese)

5 吕继东,赵德安,姬伟. 苹果采摘机器人目标果实快速跟踪识别方法[J]. 农业机械学报,2014,45(1):65 - 72.
LV Jidong, ZHAO Dean, JI Wei. Fast tracing recognition method of target fruit for apple harvesting robot[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(1): 65 - 72. (in Chinese)

6 NGUYEN T T, VANDEVOORDE K, WOUTERS N, et al. Detection of red and bicoloured apples on tree with an RGB - D camera

[J]. Biosystems Engineering, 2016, 146: 33 - 44.
7 JI W, ZHAO D, CHENG F, et al. Automatic recognition vision system guided for apple harvesting robot[J]. Computers & Electrical Engineering, 2012, 38(5): 1186 - 1195.
8 TANIGAKI K, FUJIURA T, AKASE A, et al. Cherry-harvesting robot[J]. Computers and Electronics in Agriculture, 2008, 63(1): 65 - 72.
9 LI P, LEE S H, HSU H Y. Study on citrus fruit image data separability by segmentation methods[J]. Procedia Engineering, 2011, 23(5): 408 - 416.
10 钱建平,杨信廷,吴晓明,等. 自然场景下基于混合颜色空间的成熟期苹果识别方法[J]. 农业工程学报,2012,28(17):137 - 142.
QIAN Jianping, YANG Xinting, WU Xiaoming, et al. Mature apple recognition based on hybrid color space in natural scene[J]. Transactions of the CSAE, 2012, 28(17): 137 - 142. (in Chinese)
11 杨国彬,赵杰文,向忠平. 利用计算机视觉对自然背景下西红柿进行判别[J]. 农机化研究,2003(1):60 - 62.
YANG Guobin, ZHAO Jiewen, XIANG Zhongping. Discriminating tomatoes based on computer vision technology in natural outdoor scenes[J]. Journal of Agricultural Mechanization Research, 2003(1): 60 - 62. (in Chinese)
12 JIMÉNEZ A R, CERES R, PONS J L. A vision system based on a laser range-finder applied to robotic fruit harvesting[J]. Machine Vision and Applications, 2000, 11(6): 321 - 329.
13 马晓丹,刘刚,周薇,等. 基于量子遗传模糊神经网络的苹果果实识别[J]. 农业机械学报,2013,44(12):227 - 232,251.
MA Xiaodan, LIU Gang, ZHOU Wei, et al. Apple recognition based on fuzzy neural network and quantum genetic algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(12): 227 - 232, 251. (in Chinese)
14 KOCH C, ULLMAN S. Shifts in selective visual attention: towards the underlying neural circuitry[J]. Human Neurobiology, 1985, 4(4): 219 - 246.
15 ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254 - 1259.
16 ITTI L, KOCH C. Computational modelling of visual attention[J]. Nature Reviews Neuroscience, 2001, 2(3): 194 - 203.
17 GOLDSTEIN E B. 认知心理学——心智、研究与你的生活[M]. 张明,译. 北京:中国轻工业出版社,2015.
18 MYERS D. 心理学[M]. 9 版. 黄希庭,译. 北京:人民邮电出版社,2013.
19 TREISMAN A M, GELADE G. A feature-integration theory of attention[J]. Cognitive Psychology, 1980, 12(1): 97 - 136.
20 黎万义,王鹏,乔红. 引入视觉注意机制的目标跟踪方法综述[J]. 自动化学报,2014,40(4):561 - 576.
LI Wanyi, WANG Peng, QIAO Hong. A survey of visual attention based methods for object tracking[J]. Acta Automatica Sinica, 2014, 40(4): 561 - 576. (in Chinese)
21 BORJI A, ITTI L. State-of-the-art in visual attention modeling[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 185 - 207.
22 MENG Jiao, WANG Sun'an. The recognition of overlapping apple fruits based on boundary curvature estimation[C]//2015 Sixth International Conference on Intelligent Systems Design & Engineering Applications, 2015: 874 - 877.
23 汤修映,张铁中. 黄瓜收获机器人避碰轨迹规划[J]. 农机化研究,2010,32(5):98 - 102.
TANG Xiuying, ZHANG Tiezhong. Obstacle avoidance trajectory planning of cucumber harvesting robot[J]. Journal of Agricultural Mechanization Research, 2010, 32(5): 98 - 102. (in Chinese)