

doi:10.6041/j.issn.1000-1298.2015.09.045

# 基于语音识别的蔬菜病害视频语义标注与分割方法\*

李鑫星<sup>1</sup> 刘春迪<sup>1,2</sup> 温皓杰<sup>1,2</sup> 苏叶<sup>1,2</sup> 傅泽田<sup>1,3</sup> 张领先<sup>1,2</sup>(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 农业部农业信息获取技术重点实验室, 北京 100083;  
3. 食品质量与安全北京实验室, 北京 100083)

**摘要:** 为了向农民提供蔬菜病害知识, 基于语音识别技术设计了一种蔬菜病害视频标注与分割方法, 可将科研机构录制的蔬菜病害视频分割成适合手机播放的小视频段落。在前期设计的视频镜头切分方法基础上, 进一步设计出基于语音识别技术的视频语义标注及视频镜头聚类方法, 即首先采用成熟的语音识别技术, 将视频镜头的语音讲解识别为文本形式; 进而基于本体对识别文本进行相应的语义处理, 从中提取出能起到指示作用的关键语义实体, 并将其恰当的组织形式作为视频镜头的语义标注; 最终根据用户提供的关键词并结合视频镜头的语义标注, 对视频镜头进行聚类和重组, 从而实现对于蔬菜病害视频的最终分割。所设计的视频镜头语义标注方法对 2 个测试视频的查全率分别达到 96.08%、94.93%, 查准率分别达到 94.31%、95.98%,  $F-1$  测度也分别达到 0.93 和 0.92; 视频镜头聚类方法使得 2 个视频的分割查全率分别达到 94.9%、98.7%, 查准率分别达到 92.1%、90.2%, 查全率平均大于 95%, 查准率大于 90%。证明所设计的蔬菜病害视频标注与分割方法具有理论和实用价值。

**关键词:** 蔬菜病害 视频分割 语音识别 语义标注

**中图分类号:** S2      **文献标识码:** A      **文章编号:** 1000-1298(2015)09-0308-06

## Video Semantic Annotation and Segmentation Method of Vegetable Disease Knowledge Based on Voice Recognition

Li Xinxing<sup>1</sup> Liu Chundi<sup>1,2</sup> Wen Haojie<sup>1,2</sup> Su Ye<sup>1,2</sup> Fu Zetian<sup>1,3</sup> Zhang Lingxian<sup>1,2</sup>

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture, Beijing 100083, China

3. Beijing Laboratory of Food Quality and Safety, Beijing 100083, China)

**Abstract:** To provide farmers with vegetable diseases knowledge, this paper proposes a method based on voice recognition technology to label and split vegetable diseases videos. Through this method, videos about vegetable diseases can be split into several smaller segments which are more suitable for cell phone. The methods of semantic annotation and video shot clustering were based on video segmentation and voice recognition. In this method, the audio signals of videos were transformed into text strings firstly by voice recognition. Then key semantic entities for labelling video shots semantically were split from the text strings. Finally different video shots were clustered and recombined based on keywords provided by user and the semantic labels of video shots. When applying the method of semantic annotation to two videos, the recall ratios were up to 96.08% and 94.93%, the precision ratios were up to 94.31% and 95.98%, and the  $F-1$  measures were up to 0.93 and 0.92. As for method of video shot clustering, the recall ratios were up to 94.9% and 98.7%, and the precision ratios were up to 92.1% and 90.2%. Results of comparative experiments show that the proposed method is valuable both in theory and practice.

**Key words:** Vegetable diseases Video segmentation Voice recognition Semantic annotation

收稿日期: 2015-01-22 修回日期: 2015-03-06

\* 国家自然科学基金资助项目(31271618)、现代农业产业技术体系北京市叶类蔬菜创新团队建设专项科研资金资助项目(blvt-20)、中央高校基本科研业务费专项资金资助项目(2013XJ021)和北京市大学生科学研究与创业行动计划资助项目(014bj091)

作者简介: 李鑫星, 讲师, 主要从事农业信息化技术研究, E-mail: lxxcau@cau.edu.cn

通讯作者: 张领先, 副教授, 主要从事农业信息化技术研究, E-mail: zlx131@163.com

## 引言

蔬菜病害对蔬菜质量和菜农经济收益造成巨大影响,然而菜农受教育水平相对偏低,其生产仍需要先进病害防治知识的指导。目前,科研机构已经录制了大量蔬菜病害知识视频,这些视频都由专业人员录制,知识性和专业性很强,具有很高的推广价值。若能设计出适合蔬菜病害知识特点的视频分割方法,将大段视频分割成适合手机播放的小片段并推送给用户,对菜农具有现实意义。

视频的结构包含幕、场景、镜头、帧等,视频分割实际上要将视频分割成一个个场景。现有的场景分割方法通常分为2步:依据视频低层特征(颜色、纹理等)检测出镜头边缘;通过聚类算法将视频特征相似的多个镜头合并为一个场景<sup>[1]</sup>。本文基于自适应双阈值的视频镜头切分方法<sup>[2]</sup>,进一步研究视频镜头的聚类问题。

国内外学者对于视频镜头聚类也做了一些研究,如 Mohanta 等<sup>[3]</sup>利用  $k$  均值硬聚类算法将视频镜头分类并根据 Davies - Bouldin 加以标签最终实现了视频镜头到场景的聚合;Tian 等<sup>[4]</sup>提出了一种基于物体形状聚类识别的视频物体分割技术,识别物体后根据特征实现关键帧的精确划分。李健等<sup>[5]</sup>针对基于内容的视频检索中的镜头聚类问题,采用了一种基于模拟退化思想改进的  $k$  均值聚类算法;张婵等<sup>[6]</sup>提出一种以可能性 C 模式聚类为基础的視頻关键帧提取方法;张建宁等<sup>[7]</sup>提出一种基于最优化分类的视频镜头谱聚类算法。但这些方法都是将低层特征相似的镜头合并,即通过提取每个镜头的关键帧,比较其颜色直方图等低层特征,如果两帧具有相似性(差值在一定范围内),就将这些镜头聚合在一起。这使得低层特征差别较大但在内容上密切相关、本应被并入同一场景的镜头,有可能被错误的分割成2个场景。

为解决此问题,本文设计一种适合蔬菜病害知识特点的视频镜头聚类方法。该方法是一种“动态”镜头聚类重组方法,即视频镜头是最小的视频存储物理单元,在对镜头聚类时可根据需求(通常是根据用户提供的查询关键词),从镜头库中选择所需的镜头进行合并,这种合并是多个镜头的顺序播放,而非将多个镜头的物理组织形式重新合并在一起。

## 1 视频镜头聚类总体方案

若镜头聚类时只考虑视频的低层特征,而忽略了视频的语义特征,将造成本应聚类在一起的视频

镜头分割成2个场景的问题。文本则可以很好地帮助人们理解视频语义,与视频内容匹配的文本就成为提取视频语义特征最有力的工具。本文的研究对象是讲座类视频,很多时候具有“音频为主,视频为辅”的特点,通常通过语音介绍来表达视频的语义,如能通过语音识别技术将视频的讲解语音识别为文本,则可以有效获得视频的语义信息。因此本文设计出基于语音识别技术的视频镜头聚类方法,流程如图1所示。

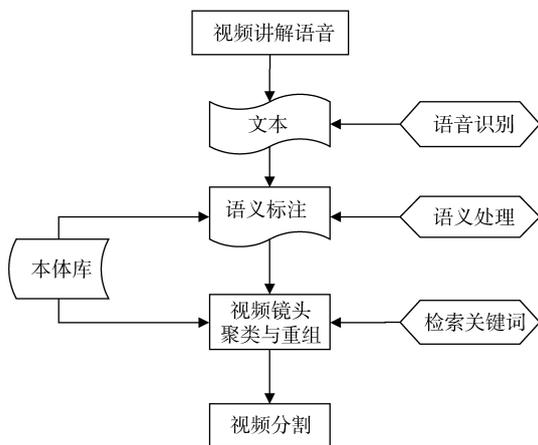


图1 基于语音识别技术的视频镜头聚类流程图

Fig. 1 Video shot clustering process based on speech recognition technology

(1)采用语音识别技术,将视频讲解语音识别为文本。现有对于语音识别技术的研究已经相对成熟,本文参照文献[8-10]的方法,采用基于隐马尔科夫模型的语音识别技术,实现视频讲解语音的识别。虽然识别文本中仍然存在误差(并没能转换为通顺、连贯、断句恰当的语句),但因为词是汉语中表达语义的关键实体,而语音识别基本可以识别出所有词语,所以这些识别文本也完全可以满足语义标注的要求。

(2)基于识别文本对视频镜头进行语义标注。需要指出的是,本文通过语音识别技术将语音识别为文本后,并不是将所有文本都作为视频镜头的标注,因文本中会包含一些不能表达语义的一般性语料,如将这些语料都标注在视频镜头上,会影响后续基于标注的视频镜头聚类与重组的效率和准确性。为此,本文设计了基于本体的视频镜头语义标注方法,在标注过程中将参照文献[11]引入蔬菜知识本体,即从识别文本中提取出能起到指示作用的关键语义实体,并将其以恰当的组织形式作为视频镜头的语义标注。

(3)基于视频的语义标注,对视频镜头进行聚类与重组,最终实现视频分割。当用户提供关键词后,将根据关键词检索视频镜头的语义标注进而锁

定该镜头,再将所有相关镜头按照人类认知习惯进行聚类与重组并反馈给用户,最终仅为用户提供其所需的视频场景(视频镜头组合),即实现视频分割。

## 2 基于本体的视频镜头语义标注

为了从识别文本中提取出能起到指示作用的关键语义实体,并以恰当的组织形式作为视频镜头的语义标注,需要对文本进行语义处理,处理流程图如图2所示。

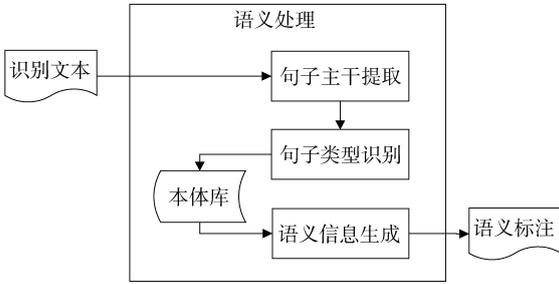


图2 语义处理流程图

Fig.2 Flow chart of semantic processing

处理步骤如下<sup>[12]</sup>：

(1)句子主干提取。在汉语中,句子的语义都是通过句子主干表达的。本文认为本体库中的实例都是语义实体,其实就是句子主干,即在本体模型的指导下,提取出句中出现的、在本体库中包含相同实例的词。

(2)句子类型识别。在构建本体库时,已经确定了每个实例的词性,例如蔬菜病害知识包含3类子集:名词性概念类子集、个体类子集以及谓词性概念类子集,而3类概念集的词性依次为名词、代词和动词,而句子类型则可以通过句子主干的词性确定。如“芹菜患上斑枯病”的句子成分为代词+动词+名词,则句子类型为主谓宾形式的陈述句。确定了句子类型,句中概念间的语义关系随即确定,如被字句中“芹菜被喷药”,即可确定句中概念的语义关系

为施事(药、喷),受事(喷、芹菜)。

(3)语义信息生成。根据前面对于句子的处理结果,语义信息则可提取出,并以某种形式表示。本文采用语义片段作为语义信息的表现形式,例如:“\_施事\_A\_B,\_时间\_C\_D,\_主谓宾\_A\_B\_C”等,语义片段可根据长度分成1级、2级、多级等不同级别。生成的语义信息中只包含能表达实际语义的词句(本体库中包含的实体),且为了满足后续镜头聚类与重组需要,对其格式也有明确规定。

经上述流程,可将语音识别文本转换为包含语义片段的形式,例如一段描述芹菜得了斑枯病为其施药的镜头,经过上述处理,可得:

视频:芹菜得斑枯病后为其施药。

识别文本:芹菜如果得了斑枯病,应该用75%百菌清600倍液或64%毒杀矾500倍液,每隔7~10d喷药2~3次。

结果:\_施事\_芹菜\_患上。

\_受事\_患上\_斑枯病。

\_施事\_75%百菌清600倍液或64%毒杀矾500倍液\_喷。

\_时间\_喷\_每隔7~10d喷药2~3次。

语句中“得了”一词,并非本体库中谓词性概念子集的实体,但其语义与实体“患上”相同,因此在标注时,将“得了”替换为“患上”,后续的视频镜头聚类根据语义标注实现。

## 3 基于标注的视频镜头动态聚类与重组

上文已经完成对视频镜头的语义标注,现将对视频镜头进行聚类与重组。本文通过对蔬菜病害知识体系的研究构建了蔬菜知识分类索引(图3)并借鉴知识网络中知识节点的结构模型及蔬菜视频语义镜头标注文本,以视频场景(聚类重组后的视频镜头集合)为最小单位,建立了基于时空顺序的蔬菜视频镜头聚类与重组实施路线模型(图4),对切分

A	B	C	D	E	F	G	H	I	J	K	L
生长阶段	生长部位	空间部位	病害名称	品种	地理位置	气候	症状表现	颜色	形状	药品名称	
种期	苗	上部	霜霉病	津春号	北方	潮湿	斑	褪绿	不规则	尿素	
苗期	根	下部	白粉病	津研4号	南方	干燥	霉	枯黄	椭圆	磷酸二氢钾	
定植期	茎	正面	角斑病	津杂5号		低温	菌丝	灰	圆	百菌清	
株期	卷须	背面	炭疽病	棚优		高温	菌核	白	水渍状	乙磷锰锌	
花期	叶	中心	灰霉病	春宝		高湿	胶	黄	鸡爪	多菌灵	
果期	子叶		黑星病	爱丰			黏物	褐	线形	农抗120	
采收期	叶脉		枯萎病	津研7号			突瘤	灰白	鼠粪状	高脂膜	
	花果		蔓枯病	夏丰1号			线虫瘿瘤	灰褐	疮瘤状	粉锈宁	
			病毒病	早丰1号			蚜虫	黑	晕圈	特富灵	
			疫病	宁丰3号				红褐		琥胶肥酸硼	
			白绢病	露地2号				茶褐		农用链霉素	
			菌核病	露地3号				粉红		甲基托布津	
			根结虫	津绿5号				绿		速克灵	
			斑点病	中农20号						扑海因	
			黑斑病	中农106号						防霉宝	
			靶斑病	方优2号						双效灵	
			叶斑病	哈研1号							
			根腐病								

图3 蔬菜病害知识分类索引表的实例

Fig.3 An instance of vegetable diseases knowledge classification index

的视频镜头按照人类认知习惯进行排序。当用户提供关键词后,将根据关键词检索视频镜头的语义标注进而锁定该镜头,再将所有相关镜头按照人类认知习惯进行聚类与重组并反馈给用户,最终实现仅为用户提供其所需的视频场景,即实现视频分割的最终目的。

图4中“提取”、“查重”、“排序”为3个步骤,节点1~4为视频镜头聚类与重组时涉及的4个知识节点,图5~8详细介绍了各节点的知识内容。

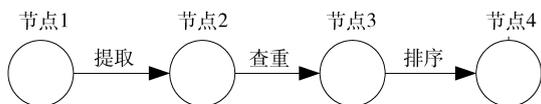


图4 视频镜头聚类与重组

Fig. 4 Video shot clustering and reorganization

图5中,  $A$  为类别;  $a_i$  为  $A$  类中的第  $i$  个关键词,  $a_i \in A$ , 其余类推。



图5 知识节点1概念模型

Fig. 5 Conceptual model of knowledge node 1

图6中,  $E-n$  为视频镜头;  $E$  为初始视频;  $n$  为视频  $E$  的第  $n$  个镜头, 其余类推。

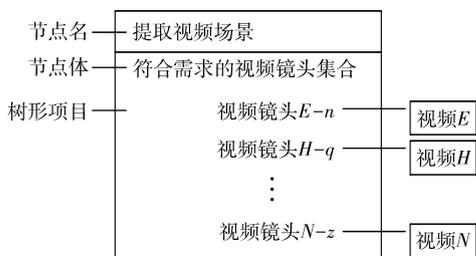


图6 知识节点2概念模型

Fig. 6 Conceptual model of knowledge node 2

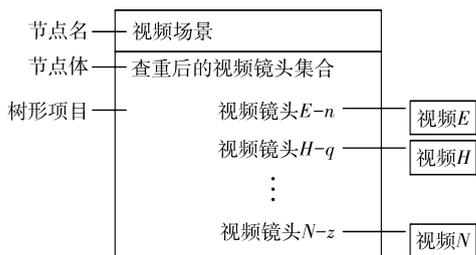


图7 知识节点3概念模型

Fig. 7 Conceptual model of knowledge node 3

基于语义标注的蔬菜视频镜头聚类与重组方法的具体实施步骤如下：

(1) 根据用户输入检索内容生成关键词集合

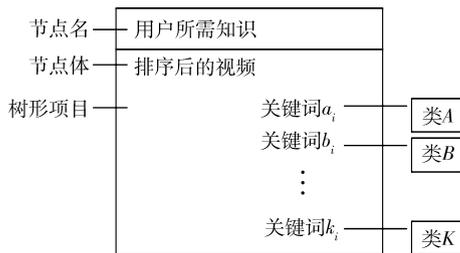


图8 知识节点4概念模型

Fig. 8 Conceptual model of knowledge node 4

$S_x = \{s_x\}$ 。

(2) 在表1中检索关键词包含集合  $S_x$  的视频镜头, 以及与关键词  $s_x$  语义相似度最高的视频镜头, 检索得到视频镜头编号集合  $N_z = \{N-z\}$ 。

(3) 对比视频镜头编号集合  $N_z$  中视频镜头集合  $N-z$  的关键词集合  $K_{N-z} = \{k_j\}$ , 剔除关键词集合完全一致的视频镜头, 获得视频镜头编号集合  $N'_z = \{N'-z\}$ 。

(4) 当视频镜头  $N'-z$  的关键词集合  $K'_{N'-z}$  中元素  $k'_j \in A \cup B \cup C$  时(其中  $A, B, C$  为顺序关键词类), 按照元素  $k'_j$  在顺序关键词类中的位置基于  $A > B > C$  的原则对视频镜头编号集合  $N'_z$  排序, 否则跳转至步骤(5)。

(5) 对视频镜头  $N'-z$  按照所属源视频  $N'$  以  $1 > 2 > \dots$  的顺序排列集合  $N'_z$ 。

(6) 对来自于相同源视频的视频镜头  $N'-z$  按照分割顺序  $z$  以  $1 > 2 > \dots$  的顺序排列集合  $N'_z$ 。

(7) 播放排序后的视频镜头集合(即为视频场景)  $N'_z = \{N'-z\}$ 。

表1 蔬菜视频标注关键词对照

Tab. 1 Annotation keywords comparison of vegetable video

视频编号	关键词
1-1	$e_1, d_1, b_5$
1-2	$e_1, d_2, b_7$
⋮	⋮
2-1	$e_1, d_3, f_1, g_1$
2-2	$e_1, d_4, f_3, g_3$
⋮	⋮

## 4 实验数据与结果分析

本文从中国农技网选取2个视频作为实验材料: 视频1是芹菜病害视频, 视频长度为2.92 min, 视频的内容包括芹菜斑枯病、早疫病、黑斑病等芹菜常见病害的特征及防治方法; 视频2为菠菜病害视频, 长度为1.65 min, 内容包括菠菜霜霉病、炭疽病、斑点病、枯萎病、叶斑病、灰霉病等。

为对本文的视频语义标注方法进行评价, 将本

文方法与最大匹配法、三元语法 2 种传统方法做对比实验,评价指标选取查准率、查全率和  $F$  测度 3 个评价指标,各评价指标的定义如下:

查全率( $R$ )<sup>[13]</sup>为净分词结果中切分正确的总词数占标准文本中总词数的百分数。

查准率( $P$ )<sup>[13]</sup>为分词结果中切分正确的总词数占有输出结果总词数的百分数。

$F$  测度<sup>[14]</sup>计算公式为

$$F = (\beta^2 + 1)PR / (\beta^2P + R) \quad (1)$$

在通常情况下,取  $\beta = 1$ ,则得到  $F - 1$  测度的计算公式为

$$F = 2PR / (P + R) \quad (2)$$

在评价本文视频聚类方法时,用本文方法和直方图法、双阈值法 2 种传统方法做比较,评价指标为查全率和查准率,其定义为<sup>[15]</sup>

$$V_r = \frac{N_c}{N_c + N_m} \times 100\% \quad (3)$$

$$V_p = \frac{N_c}{N_c + N_f} \times 100\% \quad (4)$$

式中  $N_c$ ——正确检测到的镜头个数

$N_m$ ——漏检的镜头个数

$N_f$ ——误检的镜头个数

实验结果如图 9、10 和表 2、3 所示。



图 9 镜头分割结果

Fig. 9 Video shot segmentation result



图 10 视频镜头集合

Fig. 10 Video shot set

表 2、3 的数据是各方法运行在相同的系统环境下得到的,从表中可看出,本文设计的视频镜头语义标注方法对 2 个视频的查全率分别达到 96.08%、

表 2 视频镜头语义标注实验数据及结果

Tab. 2 Experiment data and result of video shot

视频	标注方法	查全率/%	查准率/%	$F - 1$
1	最大匹配	95.09	88.31	0.93
	三元语法	93.37	93.06	0.92
	本文方法	96.08	94.31	0.93
2	最大匹配	94.02	89.07	0.90
	三元语法	93.31	92.35	0.91
	本文方法	94.93	95.98	0.92

表 3 视频镜头聚类实验数据及结果

Tab. 3 Experiment data and result of video shot clustering

视频	分割方法	查全率/%	查准率/%	分割时间/s
1	直方图法	81.8	73.0	93
	双阈值法	93.3	86.4	104
	本文方法	94.9	92.1	121
2	直方图法	77.4	73.8	87
	双阈值法	97.2	89.6	98
	本文方法	98.7	90.2	109

94.93%, 查准率分别达到 94.31%、95.98%,  $F - 1$  测度也分别达到 0.93、0.92; 视频镜头聚类方法中 2 个视频的查全率分别达到 94.9%、98.7%, 查准率分别达到 92.1%、90.2%, 查全率平均大于 95%, 查准率大于 90%。另外, 实验也特别对视频分割时间进行了对比, 由表中得出各个算法所需的时间相差不大, 所以本文设计的方法在保证准确性的前提下并没有降低分割效率。整体来看, 本文提出的视频镜头语义标注和视频镜头聚类方法, 能有效地通过对 2 个蔬菜病害知识视频进行分割, 同时较好地兼顾了分割的准确性、可靠性和实时性。

## 5 结束语

针对蔬菜病害知识视频特点, 设计出基于语音识别技术的蔬菜病害知识视频分割方法。该方法基于前期研究提出的自适应双阈值的视频镜头切分方法, 结合农业知识讲座类视频“音频为主, 视频为辅”的特点, 通过语音识别技术对视频镜头进行语义标注, 进而基于语义标注对视频进行聚类 and 重组, 最终实现视频的分割。本文设计的视频镜头语义标注方法对 2 个视频的查全率分别达到 96.08%、94.93%, 查准率分别达到 94.31%、95.98%,  $F - 1$  测度也分别到了 0.93、0.92; 视频镜头聚类方法使得 2 个视频的查全率分别达到 94.9%、98.7%, 查准率分别达到 92.1%、90.2%, 查全率平均大于 95%, 查准率大于 90%。与传统方法比较, 证明本文设计的视频分割方法具有理论和实用价值。

## 参 考 文 献

- 1 屈有政. 基于内容的暴力视频检索[D]. 北京:北京交通大学,2011.
- 2 傅泽田,苏叶,张领先,等. 基于自适应双阈值的蔬菜病害知识视频分割方法[J]. 农业工程学报,2013,29(9):148-155.  
Fu Zetian, Su Ye, Zhang Lingxian, et al. Method of video segmentation for vegetable disease based on self-adaptive dual thresholds[J]. Transactions of the CSAE, 2013, 29(9): 148-155. (in Chinese)
- 3 Mohanta P P, Saha S K, Chanda B. A heuristic algorithm for video scene detection using shot cluster sequence analysis[C]//7th Indian Conference on Computer Vision, Graphics and Image Processing, 2010: 464-471.
- 4 Tian Zhiqiang, Xue Jianru, Lan Xuguang, et al. Object segmentation and key-pose based summarization for motion video[J]. Multimedia Tools and Applications, 2014,72(2): 1773-1802.
- 5 李健,宋立新. 模拟退火改进K均值算法在镜头聚类中的应用[J]. 哈尔滨理工大学学报,2010,15(6):13-16.  
Li Jian, Song Lixin. Simulated annealing K-means clustering algorithm for video shots[J]. Journal of Harbin University of Science and Technology, 2010,15(6): 13-16. (in Chinese)
- 6 张婵,高新波,姬红兵,等. 视频关键帧提取的可能性C-模式聚类算法[J]. 计算机辅助设计与图形学学报,2005,17(9):2040-2045.  
Zhang Chan, Gao Xinbo, Ji Hongbing et al. Video key frame extraction algorithm based on possibilistic C-pattern clustering[J]. Journal of Computer-aided Design & Computer Graphics, 2005,17(9):2040-2045. (in Chinese)
- 7 张建宇,孙立峰,钟玉琢,等. 基于最优化分类的视频镜头谱聚类算法[J]. 清华大学学报:自然科学版,2007,47(10):1700-1703.  
Zhang Jianning, Sun Lifeng, Zhong Yuzhuo, et al. Video shot spectral clustering algorithm by optimized automatic cluster model selection[J]. Journal of Tsinghua University: Science and Technology, 2007, 47(10): 1700-1703. (in Chinese)
- 8 Saeed V Vaseghi. 现代数字信号处理与噪声降低[M]. 邱天爽,刘文红,郭莹,译. 北京:电子工业出版社,2007:99-107.
- 9 杨笔锋. 基于改进训练算法的HMM语音识别技术研究[D]. 长沙:湖南大学,2010.
- 10 黄文龙. 语音识别关键技术研究及系统实现[D]. 重庆:重庆大学,2010.
- 11 岳峻. 基于本体的蔬菜供应链知识获取系统研究[D]. 北京:中国农业大学,2007.
- 12 周志伟. 支持语义的视频检索技术研究[D]. 合肥:中国科学技术大学,2011.
- 13 张彩琴,袁健. 改进的正向最大匹配分词算法[J]. 计算机工程与设计,2010,31(11):2595-2633.  
Zhang Caiqin, Yuan Jian. Improved forward algorithm for maximum matching word segmentation [J]. Computer Engineering and Design, 2010,31(11):2595-2633. (in Chinese)
- 14 杨超. 基于最大匹配的书面汉语自动分词研究[D]. 长沙:湖南大学,2004.
- 15 张玉珍,杨明,王建宇,等. 基于运动补偿和自适应双阈值的镜头分割[J]. 计算机科学,2010,37(9):282-285.  
Zhang Yuzhen, Yang Ming, Wang Jianyu, et al. Shot segmentation based on motion compensation and self-adaptive dual thresholds[J]. Computer Science, 2010, 37(9): 282-285. (in Chinese)

## (上接第250页)

- 12 李晶,张东兴,刘宝. 苹果分级机输送与翻转机构设计[J]. 农业机械学报,2009,40(5):158-161.  
Li Jing, Zhang Dongxing, Liu Bao. Design of conveyor and turnover mechanism of apple grader [J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(5):158-161. (in Chinese)
- 13 魏新华,胡学同,周杏鹏,等. 水果机器视觉自动分选机同步控制系统设计[J]. 农业机械学报,2008,39(11):99-103.  
Wei Xinhua, Hu Xuetong, Zhou Xingpeng, et al. Design of a synchronization control system for fruit machine vision grader [J]. Transactions of the Chinese Society for Agricultural Machinery, 2008, 39(11):99-103. (in Chinese)
- 14 王松磊,刘民法,何建国,等. 红枣自动快速无损检测分级机研究[J]. 农业机械学报,2014,45(9):79-83.  
Wang Songlei, Liu Minfa, He Jianguo, et al. Study of rapid and automatic non-destructive testing grader for red jujubes [J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(9):79-83. (in Chinese)
- 15 陈红,夏青,左婷,等. 基于机器视觉的花菇分选技术[J]. 农业机械学报,2014,45(1):281-287.  
Chen Hong, Xia Qing, Zuo Ting, et al. Determination of shiitake mushroom grading based on machine vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(1):281-287. (in Chinese)
- 16 郑冠楠,谭豫之,张俊雄,等. 基于计算机视觉的马铃薯自动检测分级[J]. 农业机械学报,2009,40(4):166-168, 156.  
Zheng Guannan, Tan Yuzhi, Zhang Junxiong, et al. Automatic detecting and grading method of potatoes with computer vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(4):166-168, 156. (in Chinese)
- 17 张俊雄,荀一,李伟. 山竹的计算机视觉分级方法[J]. 农业机械学报,2009,40(11):176-179, 188.  
Zhang Junxiong, Xun Yi, Li Wei. Mangosteen grading based on computer vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(11):176-179, 188. (in Chinese)
- 18 殷勇,陶凯,于慧春. 基于机器视觉的苹果分级中特征参量选择方法[J]. 农业机械学报,2012,43(6):118-121.  
Yin Yong, Tao Kai, Yu Huichun. Feature selection method for apple grading based on machine vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(6):118-121. (in Chinese)
- 19 王润涛,张长利,房俊龙,等. 基于机器视觉的大豆籽粒精选技术[J]. 农业工程学报,2011,27(8):355-359.  
Wang Runtao, Zhang Changli, Fang Junlong, et al. Soybean seeds selection based on computer vision [J]. Transactions of the CSAE, 2011, 27(8):355-359. (in Chinese)
- 20 熊俊涛,邹湘军,刘念,等. 基于机器视觉的荔枝果实采摘时品质检测技术[J]. 农业机械学报,2014,45(7):54-60.  
Xiong Juntao, Zou Xiangjun, Liu Nian, et al. Fruit quality detection based on machine vision technology when picking litchi [J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(7):54-60. (in Chinese)