

引用格式:

李昕, 陈泽君, 李立君, 谭季秋, 吴发展. 基于多特征降维的油茶果壳籽粒的分选识别[J]. 湖南农业大学学报(自然科学版), 2022, 48(4): 501-506.

LI X, CHEN Z J, LI L J, TAN J Q, WU F Z. Recognition and sorting of the shell and the seed for *Camellia* based on the reduction of multi-features dimension[J]. Journal of Hunan Agricultural University(Natural Sciences), 2022, 48(4): 501-506.

投稿网址: <http://xb.hunau.edu.cn>



基于多特征降维的油茶果壳籽粒的分选识别

李昕¹, 陈泽君^{2*}, 李立君³, 谭季秋¹, 吴发展⁴

(1.湖南工程学院机械工程学院, 湖南 湘潭 411104; 2.湖南省林业科学院, 湖南 长沙 410004; 3.中南林业科技大学机电工程学院, 湖南 长沙 410004; 4.株洲丰科林业装备科技有限责任公司, 湖南 株洲 412000)

摘要: 针对目前株洲丰科林业装备有限公司智能-1 型油茶脱壳机中分选识别系统存在的识别方法单一、受分选目标颜色影响大、自适应功能较差等问题, 建立了基于多特征降维的油茶果壳籽粒的分选识别方法: 提取油茶果壳籽粒的 6 维形态和颜色特征; 根据这些特征的特性采用降维的方法, 以保证算法的识别效率; 将降维方法融入人工免疫网络算法中进行算法模型的辨识。选用颜色特征区分较明显的分别已晾晒 3 d 和 12 d 的油茶果进行采集分选, 通过降维优化得到 2 分量、4 分量与 6 分量的 3 d 晾晒样本识别率均值达 70%、80%、90%; 晾晒 12 d 的识别率均值达 50%、60%、75%; 晾晒 3 d 识别时间均值为 60 ms、350 ms、450 ms; 晾晒 12 d 的识别时间均值为 80 ms、420 ms、480 ms。

关键词: 油茶果; 分选; 颜色特征; 形态特征; 多特征降维

中图分类号: TP391

文献标志码: A

文章编号: 1007-1032(2022)04-0501-06

Recognition and sorting of the shell and the seed for *Camellia* based on the reduction of multi-features dimension

LI Xin¹, CHEN Zejun^{2*}, LI Lijun³, TAN Jiqiu¹, WU Fazhan⁴

(1.School of Mechanical Engineer Hunan Institute of Engineering, Xiangtan, Hunan 411104, China; 2.Hunan Academy of Forestry, Changsha, Hunan 410004, China; 3.School of Mechanical and Electrical Engineer, Center South University of Forestry and Technology, Changsha, Hunan 410004, China; 4.Fengke Forestry Equipment Technology Co. Ltd, Zhuzhou, Hunan 412000, China)

Abstract: Aiming at the problem of single recognition method, large influence by target color and poor adaptive function for the recognition and sorting system of intelligent-1 type *Camellia* sheller manufactured by Fengke Forestry Equipment Technology Co. Ltd, a new recognition and sorting system method of the shell and the seed was established based on the reduction of multi-features dimension for *Camellia*. Six dimensional of morphological and color features was extracted in *Camellia* shell/seed. According to the characteristics of these features, a dimension reduction method is proposed to ensure the recognition efficiency of the algorithm. The dimension reduction method is integrated into the artificial immune network algorithm for multi-features dimension reduction identification. *Camellia* fruits of drying 3 d and 12 d with obvious color characteristics were selected for collection and sorting, respectively. Through dimension reduction optimization, the average recognition rate of 2-component, 4-component and 6-component for drying 3 d samples reached 70%, 80% and 90%, and 50%, 60% and 75% for drying 12 d. The average recognition time was 60 ms, 350 ms and 450 ms in 3 d, and 80 ms, 420 ms and 480 ms in 12 d.

收稿日期: 2020-10-28

修回日期: 2021-12-08

基金项目: 国家重点研发计划项目子课题(2016YFD0702100)

作者简介: 李昕(1984—), 男, 湖南长沙人, 博士, 讲师, 主要从事机器视觉与人工智能研究, lixin7059@qq.com; *通信作者, 陈泽君, 研究员, 主要从事现代林业技术与装备研究, zejun1166@163.com

Keywords: *Camellia oleifera*; sorting; color characteristics; morphological characteristics; multi-features dimension reduction

油茶产业链的机械化和自动化进程中,油茶采摘后加工榨油环节已经比较成熟,而油茶采摘、脱壳环节的自动化还存在着较大的发展空间。李立君等对油茶果粗选机曲柄摇杆机构进行了运动学分析和试验^[1],后续在油菜果粗选机上加装6自由度林果采摘混联机械臂,运用旋量理论对其进行运动学逆向分解^[2]。高自成等^[3-4]进行了末端执行机构的研发,研究悬挂振动式油茶果采摘机齿梳式采摘机构并进行了现场试验。康地等^[5]设计了油茶果分级滚筒,利用滚筒把油茶果进行分级,再由李阳等^[6]设计出揉搓型油茶果分类脱壳清选系统,对油茶进行晾晒后脱壳与分选,实现了智能化的要求。

苗中华等^[7]研究利用K-means聚类分割法对自环境下重叠果实图像识别的算法,发现K-means的聚类特性取得很好的分割结果,但是稳定性还有所欠缺;周志宇等^[8]采用自适应Canny算子检测柑橘的边缘,还原柑橘的几何形态,但受自然光影响过大;WANG等^[9]对樱桃图像的RGB特征值进行分析,设计了一套基于RGB的樱桃色选分级系统,对于颜色分级系统的开发和升级来说有一定意义。PEARSON等^[10]分析了RGB、HSV、Lab 3种颜色模型,提出了一套基于多特征颜色模型的玉米色选系统,但是多特征颜色数据库在一定程度上存在着复杂度较高的问题。王丹丹等^[11]将深度卷积神经网络的多级传递特征应用于机器人视觉系统之中,将苹果的特征值代入深度卷积神经网络之中,以提高识别率,考虑到算法的实时性,距离实用还有一定差距。

周敬东等^[12]采用模糊聚类的色选方法对油茶果籽粒进行分选,考虑到油茶果分选的实时性要求特点,单一的模糊聚类算法可能存在运算冗余的问题,不适用于实时性较高的工程应用。刘浩等^[13]利用CCD相机与色选图像处理实现了油茶果皮与籽粒的清选分离,但是其采用CCD工业阵线相机与图像处理控制系统结合使用,相对来说成本较高。笔者在前期采用单一的色选识别方法进行分选,发现由于识别方法与识别维度单一,现有方法和设备存在着自适应性不强、识别效率较低的问题,很难解决智能化分选识别过程中所需的高效率与实时性要求。

针对现有分选设备存在的识别精准度较低的问题,基于株洲丰科林业装备有限公司智能-1型油

茶脱壳机中分选识别系统的研究基础,采取多特征综合辨识的方法,为保证识别效率,将多特征参数进行降维处理,进而将多特征降维系统应用于智能免疫算法的分选识别系统之中。试验结果表明,提出的多特征降维智能分选方法识别效率提高。现将结果报告如下。

1 数据处理与平台选择

丰科林业装备公司的油茶果成套处理果籽分选控制系统分为揉搓脱壳环节与物料分选识别吹飞环节两大部分。油茶果经过揉搓脱壳后进入物料分选系统。物料分选识别基于双视双带系统识别,即第一层分选机构中,籽壳混合物沿运输带高速作抛物运动,识别系统通过油茶籽与果壳的识别,发送判断信息给高压喷嘴;当识别为壳时,喷嘴瞬时高压气流将壳改变运动轨迹,实现籽壳分离。为提高分选精度,增加了第二层分选机构,剩余的物料混合物沿运输带继续作高速抛物运动,识别系统通过对油茶籽与壳的识别,当识别为籽时,喷嘴瞬时高压气流将壳改变运动轨迹,实现籽粒的进一步分离。

视觉系统是建立在双视双带分选系统基础上的视觉分选优化产品,将其形状和颜色作为参数特征融入多特征人工免疫系统,进行智能化实时识别。运算处理设备为安晶龙分选控制系统,安晶龙分选控制系统视觉识别硬件选用日本Computar-3s工业镜头,整体系统采用C++程序设计,工控机配置为intel至强E5-2670 8核心16线程CPU,内存80G。工控机上安装MATLAB与C++系统,以满足仿真实验及实际控制的要求。

1.1 数据采集

2021年10—11月,在湖南株洲丰科林业装备科技有限责任公司采集分别晾晒3d和12d的油茶果壳和经生产线进行脱壳后的图片。

从图像分析可知,油茶果脱壳后,晾晒3d的油茶果壳、籽粒特征区较为明显,而晾晒12d的油茶果的果壳开始逐渐转变为黑色,逐渐接近籽粒颜色。部分油茶果壳籽粒数据库图像如图1所示。

试验共采集待分选图片350幅进行分析,图像分辨率为1682×1430。



图 1 油茶果实和果壳样本

Fig.1 *Camellia* fruit and shell samples

1.2 图像分割

基于油茶分选处理系统良好的识别采集环境，为保证整体图像辨识系统的运算速度，选取最大类间分割法(Otsu)图像分割算法^[14]进行分割。

考虑视觉辨识系统对整体分选系统的实时契合度要求，对目标区域范围内的油茶果壳、籽粒的 RGB 色差分量进行提取，再进行整体图像的 R-B 色差运算，再将得到的 R-B 色差进行 Otsu 分割，以得到更清晰的分割结果。

经典 Otsu 算法分割后得到的目标可能会包含孔洞以及噪点杂质等非区域目标(图 2-b)。通过数学

形态学的开运算、闭运算、孔洞填充等步骤进行处理，得到最终处理结果，如图 2-c 所示。其中内部孔洞已被填充，噪点杂质已被清除，有利于图像下一步的操作。图像分割试验结果表明，采集图像经 Otsu 分割法分割并进行后期处理后，具有较好的实时性与识别率。图像的后期处理采用腐蚀膨胀的图像形态学操作，在操作中设定图像结构元素形状为平坦型圆盘结构元素(disk)，半径参数为 5，后续也会根据图像形状和质量的变化对结构元素的形状和大小进行动态调整。

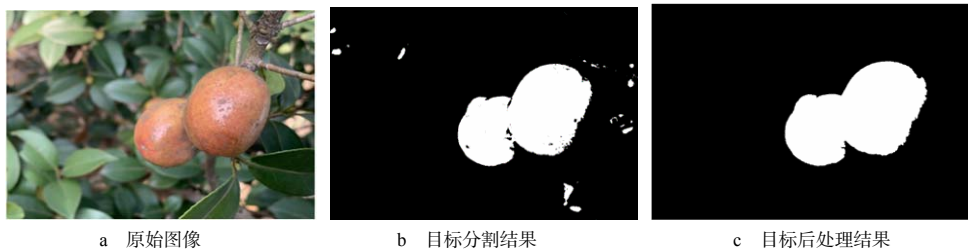


图 2 目标图像分割与后处理

Fig.2 Segmentation and post-processing of target images

1.3 特征提取

1.3.1 颜色特征参数提取

油茶果壳与籽粒在颜色与形态上具有差异化

特征。从颜色特征分析，油茶果壳颜色呈现淡青色，而籽粒颜色呈暗黑色，两者颜色特征差异使其适用于颜色特征的初步辨识，油茶果壳、籽粒的灰度直方图如图 3 所示。

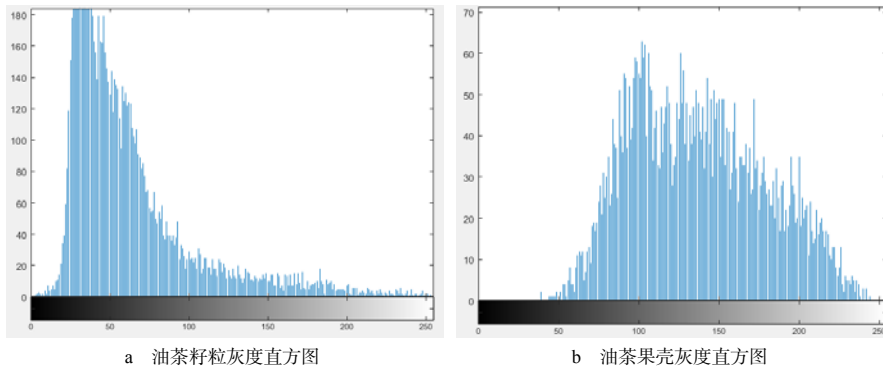


图 3 油茶果壳和籽粒的灰度直方图

Fig.3 Gray histogram differentiation of the shell and the seed for *Camellia*

基于油茶果壳、籽粒颜色特征的差异化特点,提取两者目标区域中的 RGB 色差图中 3 项颜色分量作为颜色特征数据。

1.3.2 形态特征参数提取

油茶果采摘季节为每年的秋季,受到南方秋季阴雨的气温气候影响,采摘后的油茶果壳颜色会发生一定的变化,尤其在晾晒 12 d 以后,其果壳外壁内壁颜色将由淡青色逐渐变化为暗黑色,进而与油茶籽粒颜色相接近。在这种情况下,仅依靠目标颜色分量进行目标辨识将出现很大的局限性,因此,必须对油茶果壳、籽粒的其他特征参数进行提取与综合分析。

从伸长率、圆形体、圆满度 3 项形态特征分析可以得出,油茶果壳、籽粒的 3 项形态特征区别较

为明显,其中,油茶籽粒形态较为圆满,近似于圆形;而果壳形状则呈现长条形的特征,长宽对比差异性较明显,因此提取目标区域图像中油茶果壳、籽粒的伸长率、圆形体、圆满度 3 项形态学分量作为形态特征数据在长度(L)、宽度(W)、周长(S)、面积(A)4 项基础数据的基础上,设定伸长率、圆形体、圆满度 3 项参数公式^[15]。为保证数据的范围与辨识多样化,设定了圆形体参数 $\epsilon=1.25$ 与圆满度参数 $\omega=1.98$,后续根据实际图像质量对参数进行调整,使输入特征数据具有更强的自适应性。

通过采集油茶果壳、籽粒目标区域图像的 R 分量、G 分量、B 分量、伸长率、圆形体、圆满度 6 项参数,得到颜色、形态特征参数范围如表 1 所示。

表 1 油茶果壳籽粒颜色和形态分量范围

Table 1 Color and morphological range of the shell and the seed for *Camellia*

材料	R 分量	G 分量	B 分量	伸长率	圆形体	圆满度
油茶籽粒	0~102	0~112	0~114	0.80~1.30	0.88~1.18	0.71~1.84
油茶果壳	140~240	140~240	140~240	0.33~0.70	0.37~0.45	0.12~0.63

2 改进的多特征免疫算法模型

2.1 多特征免疫算法

在得到油茶果壳、籽粒 6 项颜色、形态特征参数的基础上,使用多特征人工免疫网络算法作为多特征分选辨识算法,对油茶果壳、籽粒的多项特征进行综合辨识,以对油茶果壳、籽粒目标进行区分。

选择输入的特征参数为油茶果壳、籽粒的 RGB 颜色分量、以及目标延伸率、圆形体、圆满度的 6 项参数作为输入数据。计算免疫网络中数据的亲和力。

在免疫算法的多特征识别中,油茶果脱壳后的果壳与籽粒目标作为算法初始输入特征,将油茶果壳、籽粒的颜色、形态 6 项分量作为输入特征分量参数,并将这 6 项参数输入到免疫算法模型中进行运算。综合运算结果满足阈值要求时,定义为果壳目标,在生产线上使用并进行筛选;若目标定义为籽粒目标,则得到筛选结果。

多特征融合免疫算法的具体流程如图 4 所示。

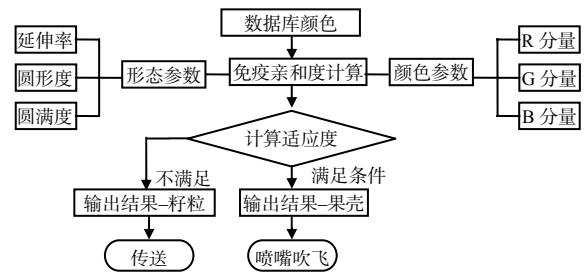


图 4 油茶果多形态特征免疫算法分选流程

Fig.4 Sorting process of *Camellia* based on multiple morphological features by immune algorithm

2.2 多特征降维模块

为降低相似的多维特征给分选识别造成的实时性影响,采用多特征降维处理,即根据不同的设备传送实时性要求,选择不同的分量作为输入参数进行运算。

首先评估颜色与形态特征数据相似性,以确定维度缩减阈值。从算法中运算时间上可以得到,6 特征同步计算将导致计算量过大,而 2 特征计算将导致分类效率过低,因此最好的策略是降低维度,而维度不能太低,宜选用 3~5 个维度特征进行计算。

将色差维度分量缩减阈值确定为 $\varepsilon=30$ ，当 R 分量与 G 分量的绝对值低于阈值，即 $\|R-G\| \leq \varepsilon$ 时，将缩减 G 分量作为输入分量，进一步将 R 分量与 B 分量进行绝对值运算，即 $\|R-B\| \leq \varepsilon$ 时，将缩减 B 分量作为输入分量，这样就将输入的分量进行了缩减。

同样，将形态学分量缩减阈值确定为 $\phi=0.15$ ，当伸长率(S)参数与圆形度参数的绝对值小于阈值时，缩小圆形度分量(Y)；当伸长度参数与圆满度参数的绝对值小于设定阈值时，将缩减圆满度分量(M)。维度缩减公式如式(1)~式(4)：

$$\|R-G\| \leq \varepsilon \quad (1)$$

$$\|R-B\| \leq \varepsilon \quad (2)$$

$$\|S-Y\| \leq \phi \quad (3)$$

$$\|S-M\| \leq \phi \quad (4)$$

考虑到降低维度与维度缩减对于算法实时性的提高，根据降维模块对算法模型进行调整，根据设定的缩减阈值($\varepsilon=30$ 、 $\phi=0.15$)进一步将算法的输入维度进行收缩。

3 模型评估与验证

3.1 算法辨识时间与效率

首先测试 10 张晾晒 3 d 的油茶果图片。选取经过降维运算的综合颜色形态特征的 2 特征分量、4 特征分量和 6 特征分量进行算法的分选识别，选取涵盖颜色与形态特征的 2 特征分量、4 特征分量以及 6 特征分量图片进行特征训练与运算识别，2 特征分量为 R 分量与伸长率分量输入组合；4 特征分量为 R 分量， G 分量与伸长率分量，圆形度分量输入组合；6 分量为 R 分量， G 分量， B 分量与伸长率分量、圆形度分量、圆满度分量输入组合。通过对 3 个特征分量进行运算，得到晾晒 3 d 时不同分量样本图片下的识别率与识别时间。

晾晒 12 d 后，油茶果壳与籽粒的颜色变得相近，给颜色识别带来了很大的困难，颜色识别率相对降低，而形态特征识别率基本不变，因此，2 分量、4 分量、6 分量的整体识别率相对降低，而整体识别时间基本不变。

通过多维度识别效率与识别时间的分析可以得出，多特征分量可提高识别率，但是增加了算法的复杂度，增加了识别时间。而使用降维运算后大幅度降低了识别时间，同时也保证了一定的识别效率。

晾晒 3 d 的油菜果，2 分量、4 分量与 6 分量样本识别率均值分别达到了 70%、80%、90%；12 d 时，由于颜色特征的变化，识别率均值降低到了 50%、60%、75%，识别时间均值分别为 60 ms、350 ms 与 450 ms，识别时间最低在 2 分量时达到了 50 ms。考虑到 MATLAB 程序的复杂度与试验计算机的效率，当程序移植到 Visual C++6.0 上时，可以满足现有设备对实时率的要求。

经过降维多特征免疫算法识别后得到的目标油茶果壳、籽粒的分选效果如图 5 所示。经过分选辨识后，油茶果壳、籽粒目标已经在视觉上进行了标记和分选，有利于后续在设备上分离操作。



图 5 分选标记后的油茶果壳和籽粒

Fig.5 The seeds and the shells of *Camellia* after sorting and labeling

3.2 降维算法与其他算法效率对比

选用晾晒 3 d 和 12 d 的油茶果实籽粒样本图片各 30 张，提取脱壳后的油茶果壳、籽粒目标的形态、颜色参数进行降维免疫分选试验，在测试中分别选用降维人工神经网络算法、多维 SVM 算法^[16]、单阈值分选算法进行对比，不同分选算法的识别率和耗时列于表 2。

表 2 不同分选算法的时间和识别率

Table 2 Comparison of the sorting time and the recognition rates for different algorithms

算法	耗时/ms		识别率/%	
	晾晒 3 d	晾晒 12 d	晾晒 3 d	晾晒 12 d
降维免疫算法	450	480	90	78
多维 SVM 算法	350	420	88	60
单一色选算法	60	80	89	15

结果表明,3种算法在晾晒3d和晾晒12d的分选识别时间基本一致。单特征色差分选方法虽然分选时间较短,但不具备自适应性,导致其分选效率较低。当油茶果晾晒12d时,果壳与籽粒目标的颜色特征区分已经不明显,因而单特征颜色阈值分选算法的整体识别率大幅度下滑,降低到了15.2%,已经不适用于分选识别系统;SVM多特征识别算法由于其算法特征的二分性特点,导致其算法结构不适用于复杂的多特征分选要求,同时晾晒12d的油茶果壳、籽粒目标由于其颜色特征趋于相似,造成识别率下降,也导致其不适用于油茶目标的多特征分选识别。而采用6分量降维人工免疫算法在晾晒3d和12d的识别时间与效率比另2种算法均达到了较好的效果。

4 结论

为保证油茶果壳、籽粒分选视觉系统的速度和效率,采用了Otsu图像分割法,与一种基于油茶果壳、籽粒颜色形态分量的辨识方法,进行目标多特征辨识。

算法中为保证运算效率,使用了多特征降维方法,即设定了相应的颜色与形态相似性阈值,通过阈值缩减相应的输入特征维度,并将降维模块植入到多特征辨识算法中。

Matlab程序试验结果表明,所采取的方法可以有效地提高算法的运行效率,缩减了算法的运算时间,多特征降维免疫算法识别时间最低在2分量时达到了50ms,保证了算法运行过程中分选的效率 and 速率。

参考文献:

[1] 李立君,胡文字,高自成,等.曲柄摇杆振动破壳油茶果粗选机运动学分析及试验[J].农业工程学报,2016,32(7):28-35.
[2] 李立君,刘涛,高自成,等.基于旋量理论的六自由度林果采摘混联机械臂运动学逆解[J].农业工程学报,

2019,35(8):75-82.

- [3] 高自成,赵凯杰,李立君,等.悬挂振动式油茶果采摘执行机构设计与试验[J].农业工程学报,2019,35(21):9-17.
[4] 高自成,李立君,李昕,等.齿梳式油茶果采摘机采摘执行机构的研制与试验[J].农业工程学报,2013,29(10):19-25.
[5] 康地,陈泽君,李阳,等.油茶果分级滚筒筛设计[J].湖南林业科技,2017,44(6):95-98.
[6] 李阳,王勇,邓腊云,等.揉搓型油茶果分类脱壳分选机的脱壳和清选效果研究[J].湖南林业科技,2015,42(2):38-42.
[7] 苗中华,沈一筹,王小华,等.自然环境下重叠果实图像识别算法与试验[J].农业机械学报,2016,47(6):21-26.
[8] 周志宇,刘迎春,张建新.基于自适应Canny算子的柑橘边缘检测[J].农业工程学报,2008,24(3):21-24.
[9] WANG Q,WANG H,XIE L J,et al. Outdoor color rating of sweet cherries using computer vision[J]. Computers and Electronics in Agriculture, 2012, 87: 113-120.
[10] PEARSON T, DAN M, PEARSON J. A machine vision system for high speed sorting of small spots on grains[J]. Journal of Food Measurement & Characterization, 2012, 6(1/2/3/4): 27-34.
[11] 王丹丹,何东健.基于R-FCN深度卷积神经网络的机器人疏果前苹果目标的识别[J].农业工程学报,2019,35(3):156-163.
[12] 周敬东,李敏慧,周明刚,等.油茶果的模糊聚类色选算法[J].中国农机化学报,2015,36(4):94-99.
[13] 刘浩,王焱清,袁博.油茶果脱壳与清选关键技术及工艺研究[J].湖北工业大学学报,2015,30(1):65-68.
[14] OTSU N. A thresholding selection method from gray-level histogram[J]. IEEE Transction on Systems Man and Cybernetics, 2007, 9(1): 62-66.
[15] 张晗,王成,董宠图,等.基于机器视觉的白菜种子精选方法研究[J].农机化研究,2021(12):31-35.
[16] 崔鹏宇.多维特征融合与Adaboost-SVM的车辆识别算法[J].控制工程,2019,26(3):608-612.

责任编辑:罗慧敏

英文编辑:吴志立