

基于树莓派的生物样本二维码识别系统的研究

葛 达, 张广腾, 王美晨, 苟奇发, 王纪鹏, 吴 畅, 陈亚东*

沈阳工业大学电气工程学院, 辽宁 沈阳

收稿日期: 2024年11月26日; 录用日期: 2025年1月9日; 发布日期: 2025年1月17日

摘 要

为了解决生物样本存储于密闭空间时, 由于光照不均匀而导致的二维码识别率不高的问题, 本文设计了一种基于树莓派的生物样本二维码识别系统。该系统硬件主要由计算机、树莓派、图像采集模块、设备外壳与平行光源组成, 软件采用多种图像处理算法, 实现对生物样本二维码的高精度识别。

关键词

树莓派, 生物样本二维码, 图像处理算法

Research on Two-Dimensional Bar Code Recognition System for Biological Samples Based on Raspberry Pi

Da Ge, Guangteng Zhang, Meichen Wang, Qifa Gou, Jipeng Wang, Chang Wu, Yadong Chen*

School of Electrical Engineering, Shenyang University of Technology, Shenyang Liaoning

Received: Nov. 26th, 2024; accepted: Jan. 9th, 2025; published: Jan. 17th, 2025

Abstract

In order to solve the problem that the two-dimensional bar code recognition of biological samples is difficult due to uneven illumination when stored in confined spaces, a two-dimensional bar code recognition system for biological samples based on Raspberry Pi was designed and manufactured in this paper. The hardware of the system is mainly composed of a personal computer, a raspberry pi, a camera, a device shell and a parallel light source, and some image processing algorithms are used to achieve high-precision recognition of the two-dimensional bar code of biological samples.

*通讯作者。

文章引用: 葛达, 张广腾, 王美晨, 苟奇发, 王纪鹏, 吴畅, 陈亚东. 基于树莓派的生物样本二维码识别系统的研究[J]. 人工智能与机器人研究, 2025, 14(1): 82-90. DOI: 10.12677/airr.2025.141008

Keywords

Raspberry Pi, Two-Dimensional Bar Code of Biological Sample, Image Processing Algorithms

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

生物样本库是一种集生物样本、信息和数据于一体的资源,对疾病预测、诊断、物种保护和作物改良等具有重要意义[1][2]。目前,转化医学和再生医学的发展都离不开生物样本资源库的理论基础和临床研究[3]-[5]。生物样本资源库发展于1970年代,欧洲国家首次开展与生物样本保存有关的工作,大约20年后,生物样本库在数量上出现了大幅增加的趋势[6][7]。从十二五期间开始,我国逐渐增加了对人类遗传资源样本库的投资,到现在为止,已经在全国范围内相继建立了一批大型的人类遗传资源样本库[8]。因此,实现对生物样本资源库的高效管理成为人们关注的焦点。

其中,二维码技术在生物样本存储与管理中得到广泛应用,成为机器视觉的热点[9]。目前,二维码的研究主要分为QR Code(二维码 Quick Response Code)和Data Matrix两种,生物样本存储主要采用Data Matrix二维码格式。针对Data Matrix二维码识别目前已经有多种算法可以识别,例如应用于方阵式DM码定位的图像分割技术[10],以及Matuszczyk等人提出在制造背景下直接识别DM二维码[11],Karrach Ladislav等人提出DM二维码识别方法[12]等。但一般情况下,由于光照不均匀或者二维码打印质量不高等原因,导致识别率不高,限制了其应用。

为了解决上述问题,本文以树莓派为核心,CSI摄像头为图像采集模块,采用传统的图像处理算法对二维码图像实现预处理,提高生物样本二维码识别的准确率。

2. 系统总体方案设计与主要硬件组成

针对生物样本库二维码识别,我们设计了树莓派采集系统,硬件系统如图1所示,当生物样本二维码放置在设备上后,在计算机的上位机软件界面点击拍照,树莓派会收到拍照指令,打开光源,并启动摄像头,通过CSI摄像头拍摄一张生物样本二维码图像,并通过http协议无线传输到上位机软件上并显示,最后点击识别图像,拍摄得到的二维码图像会经过一系列的图像处理算法,得到清晰的二维码灰度图像,将处理后的图像放入ZXing库中,识别得到生物样本二维码所对应的信息。



Figure 1. Image acquisition system based on Raspberry Pi
图1. 基于树莓派的图像采集系统

树莓派负责执行上位机传达的指令，直接与硬件设备交互，执行具体的操作任务。在本文中，下位机是一个基于 Flask 框架的 PythonWeb 应用程序，它涉及树莓派 picamera2 库调用树莓派上的摄像头捕捉具有特定像素要求的二维码图像，以及通过 GPIO (通用输入输出)接口控制硬件。它还包括一个多线程任务队列来处理并发请求，并且使用了 UDP 广播来广播设备信息。硬件结构主要模块由四个部分组成，分别是：树莓派 4B 主控模块、图像采集模块、设备外壳、平行光源。

2.1. 树莓派 4B 主控模块

系统的核心硬件是树莓派 4B 型微型计算机，其提供了卓越的计算性能和高效的数据处理能力。树莓派主要通过 CSI 端口与图像采集模块连接。在系统发出采集图像指令时，树莓派通过控制图像采集模块采集外壳上的生物样本二维码图像，并将其通过 http 协议无线到上位机软件上，实现生物样本二维码图像的采集与传输。

2.2. 图像采集模块

该设计采用树莓派 Camera Module 3 摄像头模块，Camera Module 3 摄像头使用了索尼 IMX708 传感器，其具有 1200 万像素，最高可拍摄 1080P @50 帧的视频、新模块支持高动态范围(HDR)摄影和自动对焦，并且可以输出 RAW10 文件格式。采用该摄像头模块能够让摄像头采集图像更加精准。

2.3. 设备外壳与平行光源

设备外壳由实验室利用 SolidWorks 三维设计软件与 3D 打印机独立设计并打印出来，如图 2 所示。平行光源采用 24V 平行光源，放置于外壳底部，其相对于线性光源具有很好的光照均匀性与方向性，具体地，主要通过将光打在生物样本二维码基板上，使图像采集模块得到清晰的图像。



Figure 2. Device shell
图 2. 设备外壳

3. 图像处理算法

3.1. 二维码图像处理算法设计

在本文的设计中，图像处理算法在计算机上完成，采用 C#编程，计算机通过与树莓派建立连接，获取树莓派采集到的图像信息，并经过一系列的图像处理算法得到二维码及其轮廓，以便后续识别。在图像算法方面，采用边缘检测、二值化、形态学处理等算法，加强二维码图像信息，并使用 ZXing 库进行二维码识别，如图 3 为二维码识别算法流程图。

3.2. 图像二值化处理

图像二值化是一种图像处理技术，它将图像中的每个像素值转换为两种颜色之一：黑色或白色。这

个过程通常涉及选择一个阈值,将像素值与阈值进行比较,并根据比较结果将像素值设置为黑色或白色。图像二值化可以简化图像,提高对比度,减少数据量,并用于文字识别、物体检测和图像分割等应用。图像二值化处理方法主要分为全局阈值分割、自适应阈值分割、局部阈值分割等。在生物样本库二维码识别过程中,光照环境对识别结果影响较大,为了减少光照对识别结果的影响,本文选择自适应高斯阈值分割。

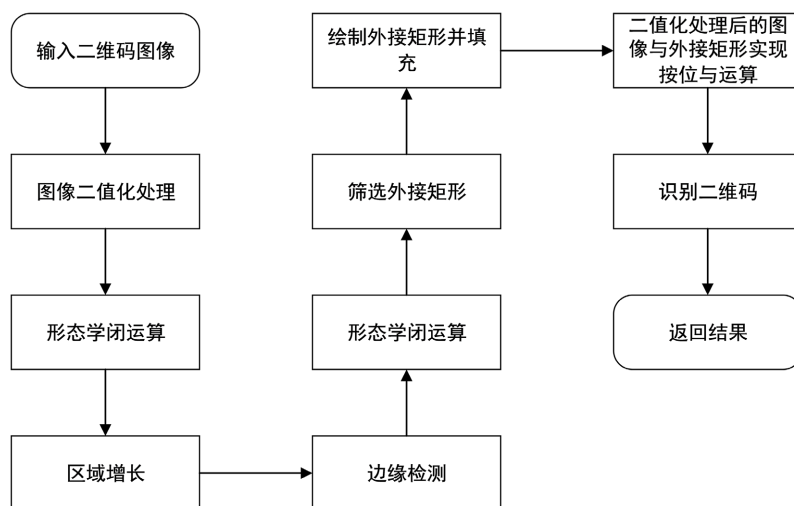


Figure 3. The flow chart of two-dimensional bar code recognition algorithms
图 3. 二维码识别算法流程图

自适应高斯阈值分割是基于图像局部区域的统计特性来计算阈值,该方法假设图像中的每个局部区域服从高斯分布,并利用这一假设来动态确定每个像素点的阈值。自适应高斯阈值分割特别适用于图像中存在不同光照条件或不同灰度级别的区域。自适应高斯阈值的计算公式如下:

$$T(x, y) = \mu(x, y) + k \cdot \sigma(x, y) \quad (1)$$

其中, $T(x, y)$ 是在位置 (x, y) 处计算出的自适应高斯阈值, $\mu(x, y)$ 是位置 (x, y) 处邻域内像素的灰度均值, $\sigma(x, y)$ 是位置 (x, y) 处邻域内像素的灰度标准差, k 是一个常数,用于调整阈值相对于均值和标准差的敏感度。

最后将每个像素的灰度值与其自适应高斯阈值 $T(x, y)$ 进行比较,如果像素值大于或等于阈值,则将该像素设置为前景(通常是白色或 255),否则设置为背景(通常是黑色或 0)。

本设计中,将二维码图像二值化是二维码图像识别的重要步骤,图像自适应高斯阈值分割可以去除图像中的噪声、提高图像对比度并去除二维码图像中的干扰信息,提高二维码识别的准确率。如图 4 为二维码原图像与图像自适应高斯阈值分割后的图像。

3.3. 形态学闭运算

图像形态学闭运算通过先进行膨胀操作,再进行腐蚀操作,来填充图像中的小孔和裂缝,并平滑图像的边缘。此过程可以增强图像中的连通区域,并减少图像中不必要的细节信息,从而提高图像的鲁棒性。形态学闭运算计算过程如下:

先膨胀:

$$A \oplus B = \{Z | (B)_Z \cap A \neq \emptyset\} \quad (2)$$

再腐蚀:

$$A \ominus B = \{Z | (B)_Z \subseteq A\} \quad (3)$$

最后进行闭运算:

$$A \cdot B = (A \oplus B) \ominus B \quad (4)$$

其中, $A \oplus B$ 表示 A 被结构元素 B 膨胀的结果, $A \ominus B$ 表示 A 被结构元素 B 腐蚀的结果, $A \cdot B$ 表示 A 经过闭运算后的结果。

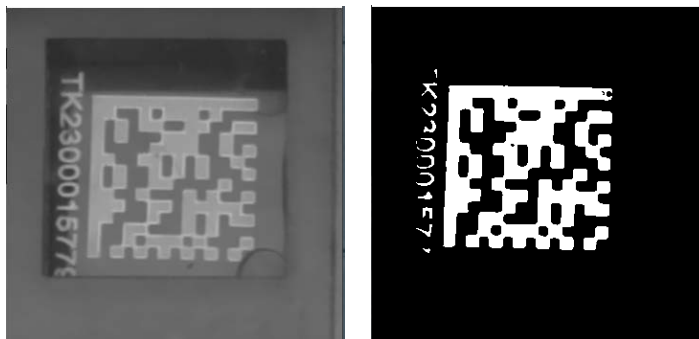


Figure 4. The original image of the two-dimensional bar code and the binary image

图 4. 二维码原图像与图像二值化后的图像

本设计中, 图像形态学闭运算主要填充二维码二值化图像中的小孔和裂缝, 并平滑图像, 消除图像噪声点对图像识别的影响, 如图 5 为图像形态学闭运算后的图像。

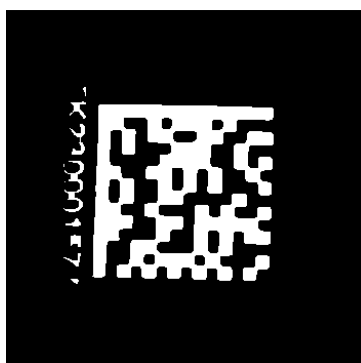


Figure 5. Image after morphology closing operation

图 5. 图像形态学闭运算后的图像

3.4. 图像区域增长

图像区域增长是一种基于种子点的图像分割方法, 它从种子点开始, 逐像素地判断周围像素是否满足特定的条件, 并将满足条件的像素添加到区域中, 从而逐步扩大区域范围, 直至达到图像分割目标。增长过程如下:

设 S 是种子点的集合, T 是当前区域, N 是像素 p 的邻域。

初始化: $T = S$ (初始区域包含所有种子点)

增长过程: 对于 T 中的每个像素 p , 计算其邻域 $N_{(p)}$ 中的每个像素 q 的相似性:

$$f(q) = |I(q) - I(p)| \quad (5)$$

其中, $I(q)$ 和 $I(p)$ 分别是像素 q 和 p 的灰度值。如果 $f(q)$ 小于某个阈值 θ , 则将 q 添加到区域 T 中:

$$T = T \cup \{q | f(q) < \theta\} \quad (6)$$

重复上述过程, 直到没有新的像素可以添加到区域 T 中。

本设计中, 图像区域增长主要用于消除图像中二维码以外的区域与比较分散的噪声点, 提高图像识别率, 本文种子点选取 20 个种子点, 分布于图像中心区域, 如图 6 为图像区域增长后的图像。

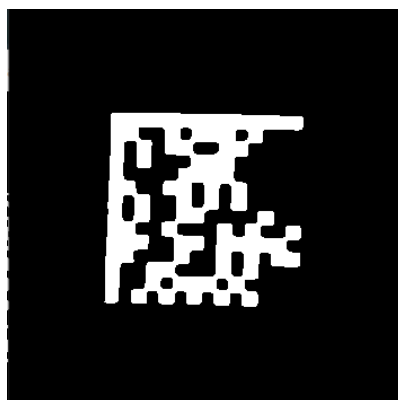


Figure 6. Image after the image region growth algorithm
图 6. 图像区域增长后的图像

3.5. 图像边缘检测

图像边缘检测通过识别图像中亮度或颜色发生显著变化的像素点, 来提取图像的边缘信息。边缘检测可以帮助我们识别图像中的物体轮廓、纹理边界等特征, 从而进行图像分割、目标识别等后续处理。本设计中, 图像边缘检测使用 Canny 边缘检测算法, 主要用于提取二维码轮廓信息, 便于后续的图像填充与图像运算。如图 7 为边缘检测后的图像。

Canny 边缘检测算法通常包括五个步骤: 滤除噪声、计算梯度、非极大值抑制、双阈值、边缘跟踪和滞后阈值。

滤除噪声, 使用高斯滤波器平滑图像以减少噪声的影响。高斯滤波器是一种加权平均滤波器, 其权重由高斯函数确定:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (7)$$

其中, σ 是高斯分布的标准差, 它控制着平滑的程度。

计算梯度, 计算图像中每个像素点的梯度大小和方向。通常使用 Sobel 算子来近似图像的梯度:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (8)$$

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (9)$$

梯度的大小 M 和方向 θ 可以通过以下公式计算:

$$M = \sqrt{G_x^2 + G_y^2} \quad (10)$$

$$\theta = \arctan 2(G_y, G_x) \quad (11)$$

其中, $\arctan 2$ 是计算从 x 轴到点 (G_y, G_x) 的向量角度的函数。

非极大值抑制, 每个像素点与其梯度方向上相邻的像素点进行比较。如果该像素点的梯度不是局部最大值, 则将其抑制(即设置为 0)。这是通过比较当前像素点的梯度大小与其梯度方向上的相邻像素点的梯度大小来完成的。

双阈值, 经过非极大抑制后图像中仍然有很多噪声点。Canny 算法中应用了一种叫双阈值的技术。即设定一个阈值上界和阈值下界, 图像中的像素点如果大于阈值上界则认为必然是边界(称为强边界), 小于阈值下界则认为必然不是边界, 两者之间的则认为是候选项(称为弱边界), 需进行进一步处理。

边缘跟踪和滞后阈值, 只有当弱边缘连接到强边缘时, 弱边缘才被认为是真正的边缘。这个过程可以确保边缘是连续的, 并且减少了假边缘的出现。

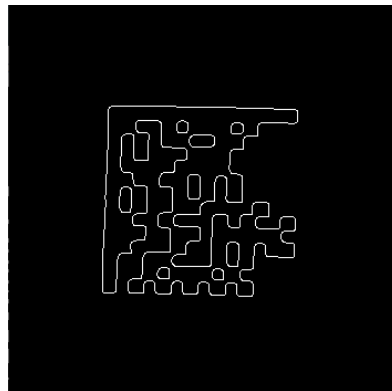


Figure 7. Image after edge detection algorithm canny

图 7. 边缘检测后的图像

3.6. 图像填充与图像运算

图像填充涉及将特定颜色或图案填充到图像的指定区域中, 这个过程通常通过算法来确定填充区域的边界, 并确保填充不会超出这些边界, 常用于图像编辑、增强和预处理步骤中, 以改善视觉效果或为后续处理做准备。

图像运算是通过对图像中的像素值执行数学运算来生成新图像或修改现有图像的过程, 这些运算包括加法、减法、乘法、除法以及更复杂的算术和逻辑运算, 用于实现图像增强、滤波、合成、对比度调整等多种图像处理效果。其中, 图像按位与运算过程如下:

对于图像中的每个像素 P_A 和 P_B , 按位与运算的公式可以表示为:

$$P_{out} = P_A \& P_B \quad (12)$$

其中, P_{out} 是输出图像的像素值, $\&$ 是按位与运算符。

本设计中图像填充主要用于填充整个二维码图像便于后续提取二维码轮廓信息; 图像运算主要用于将填充后的图像与形态学闭运算后的图像实现按位与运算, 得到完整的二维码图像, 如图 8 为填充后的图像与形态学闭运算后的图像。

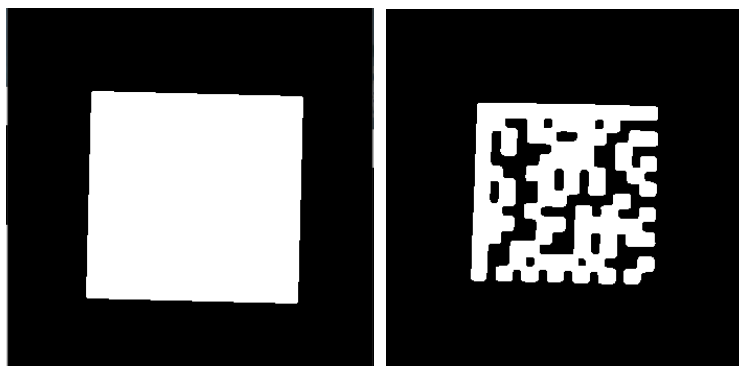


Figure 8. The image after filling and the image after morphological closure operation
图 8. 填充后的图像与形态学闭运算后的图像

3.7. 二维码识别结果

ZXing (“Zebra Crossing”的缩写)是一个开源的、多语言支持的条码处理库,它提供了在多种平台上识别和生成一维码(如 UPC、Code 128)和二维码(如 QR Code、Data Matrix)的功能,通过其核心 API 进行图像预处理、条码检测和解码,同时支持多种编程语言,包括 Java、C#、PHP、Python 等,使得开发者能够在应用程序中轻松集成条码扫描和生成功能。ZXing 广泛应用于移动支付、产品追踪、信息获取、文档管理等领域。如图 9 为经过图像处理与 ZXing 库识别后的二维码识别结果图。由图 9 可以看出,算法可以实现 48 孔生物样本二维码识别。

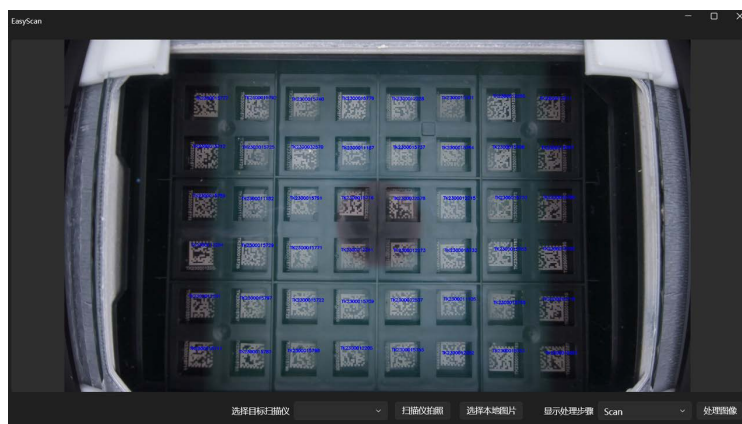


Figure 9. Diagram of two-dimensional bar code's recognition result
图 9. 二维码识别结果图

4. 结果讨论与分析

本文构建了基于树莓派的二维码识别系统,将树莓派与个人计算机连接在同一个局域网上,并通过 http 协议连接。通过上位机控制树莓派采集生物样本二维码图像,并传输到上位机,接着经过一系列的图像处理算法,得到二维码轮廓并将其传输到 ZXing 库中进行识别,最后得到生物样本二维码图像信息。

为了解决光照等因素对二维码识别率的影响,在图像处理算法方面,选择了图像二值化处理、形态学运算、区域增长等算法,消除了图像中的干扰信息,其中形态学运算可以提取图像边缘,区域增长算法从图像中心区域选取多个种子点,提取了二维码区域,而将文字部分排除。然后通过图像边缘检测找到二维码的两个长边,通过填充图像区域,得到了二维码的准确位置。最终将提取得到的二维码图像导

入 ZXing 库, 实现了生物样本二维码的识别通过上述的图像处理方法, 大大提高了二维码图像识别的准确率。测试了多个生物样本 SBS 板架, 均可识别所有的二维码。

基金项目

辽宁省教育厅面上项目(JYTMS20231205)。

参考文献

- [1] 赵小燕, 裴宇盛, 高华, 等. 生物安全样本库的发展、应用现状与探讨[J]. 中国医药生物技术, 2021, 16(4): 378-382.
- [2] 于继海, 高大鹏, 王肇光, 等. 中国生物样本库的现状与建设思考[J]. 中国高新科技, 2023(4): 148-150.
- [3] 陈涵, 李静. 芬兰红十字血液中心生物样本库伦理管理经验及启示[J]. 中国医学伦理学, 2023, 36(4): 428-433.
- [4] 蔡莉, 张玉霞. 科技创新引领下医院生物样本库可持续发展问题探讨[J]. 科技管理研究, 2022, 42(6): 83-88.
- [5] 张雨桐, 刘健, 陈诗雨, 等. 转化医学发展对科研管理工作的启示[J]. 经济师, 2022(9): 249-250.
- [6] 芦慧敏, 王岩, 陈聪, 等. 生物样本库信息管理系统的建设与应用[J]. 重庆医学, 2023, 52(14): 2211-2214.
- [7] 张雪莹, 朱承超. 自动化生物样本库建设与信息化管理[J]. 海军医学杂志, 2023, 44(4): 430-432.
- [8] 董晨章, 朱晓燕, 姜来, 等. 我国人类遗传资源应用与管理现状[J]. 上海预防医学, 2023, 35(5): 513-517.
- [9] 金龙, 潘军威. 二维码识别方法、装置、计算机设备和可读存储介质[P]. 中国专利, CN202010321936.X. 2020-09-01.
- [10] 康三顺. 应用于方阵式 DM 码定位的图像分割技术研究[J]. 现代信息科技, 2020, 4(22): 91-94.
- [11] Matuszczyk, D. and Weichert, F. (2023) Reading Direct-Part Marking Data Matrix Code in the Context of Polymer-Based Additive Manufacturing. *Sensors*, **23**, Article 1619. <https://doi.org/10.3390/s23031619>
- [12] Karrach, L. and Pivarčiová, E. (2021) Comparative Study of Data Matrix Codes Localization and Recognition Methods. *Journal of Imaging*, **7**, Article 163. <https://doi.org/10.3390/jimaging7090163>