

基于手势识别的人机交互技术的研究

董小龙, 赵斯衍

(福建省计量科学研究院, 福建福州, 350003)

摘要: 人体手势识别具有直观、方便的优点, 在人机交互技术领域具有重要的地位。本文提出了在复杂的背景环境下的手势识别, 通过 YCrCb 方法来提取出手势区域, 同时, 为了高效、准确的对手势进行分割, 利用基于全卷积神经网络算法来得到图像特征, 满足实时人机交互的需求, 使用户得到舒适的智能体验。

关键词: 手势识别; 分割算法; 卷积神经网络; 人机交互

DOI:10.16589/j.cnki.cn11-3571/tn.2020.19.023

0 引言

随着近些年人工智能的火热, 新技术也越来越趋于智能化和人性化^[1-3]。在人机交互领域, 从一开始的命令行界面, 到后来微软 windows 操作系统的 GUI 图形界面^[4,5], 都是朝着人性化的方向发展, 而技术的逐步提升, 带来了更加便捷的操作方式: 手势控制^[6-9]。本文提出的基于手势识别的人机交互技术能够在复杂的背景环境下对人体手势进行准确识别, 不会因为用户的多样性而给系统的实时性和精确性带来影响。同时, 利用手势识别和指尖检测来对人体手势进行分割, 并利用机器学习中的卷积神经网络来提高识别精度。

1 基于深度学习的手势分割法

深度学习中的卷积神经网络作为图像处理中的高效方法, 利用卷积、池化和全连接层来搭建网络结构, 其中, 卷积层用来提取特征, 通过权重和对比结果来得到相应的特征, 计算公式为:

$$x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l \right) \quad (1)$$

公式 (1) 中, x_j^l , x_i^{l-1} 和 M_j 作为卷积层中第 j 、 i 个特征映射和映射集合, b_j^l 、 k_{ij}^l 为偏置和权重。

池化层是用来对相似的特征进行融合, 从而大大降低了网络参数的个数, 通过公式:

$$x_j^l = \beta_j^l \text{down}(x_j^{l-1}) + b_j^l \quad (2)$$

公式 (2) 中, x_j^l , x_j^{l-1} 和 β_j^l 为池化层的特征映射和权重。之后, 利用全连接层将得到的二维特征图变换成一维的特征向量。本系统将基于深度学习的卷积神经网络来将手势进行分割, 可以解决在干扰情况下准确的判读出人体的手势。同时, 我们在系统后台数据库中利用全卷积网络来创建手势库, 计算 Softmax 归一化概率:

$$\begin{cases} x_i = x_i - \max(x_1, \dots, x_n) \\ p_i = e^{x_i} / \sum_{j=1}^n e^{x_j} \end{cases} \quad (3)$$

利用样本标签值 t 来计算损失值:

$$Loss = -\log p_t \quad (4)$$

在本论文中设计的全卷积神经网络的参数如表 1 所示。

表1 全卷积神经网络参数

层	尺寸k	补给像素	核滑动步长	图像特征
数据层	No	No	No	$4 \times (320 \times 320)$
卷积层	12×12	15	5	$96 \times (80 \times 80)$
池化层	4×4	0	2	$96 \times (40 \times 40)$
归一化层	No	No	No	$96 \times (40 \times 40)$

2 手势检测与分割

本文中通过肤色检测来对手和周围背景进行识别, 在 HSV 颜色空间下进行肤色检测, 对图像颜色空间进行转换, 利用 YCrCb 方法来选取阈值, 通过计算得到灰度直方图如图 1 所示。

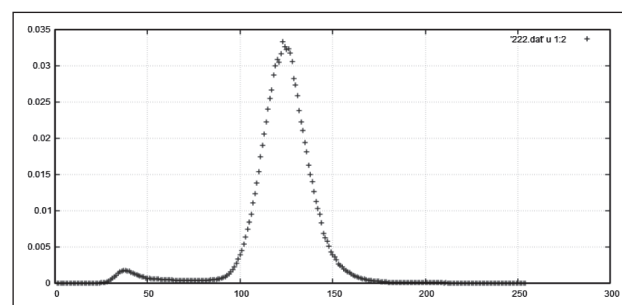


图1 灰度直方图

从图 1 中可以看出, 在峰值处, 其点是分散且急剧变化的部分是背景色 (100-150 之间), 峰值很小, 即点很集中的是前景色, 我们将图中的均值记为 M 。我们将这个图像分为两部分, 分别记为 X 和 Y , 对应的就是前景色和背景

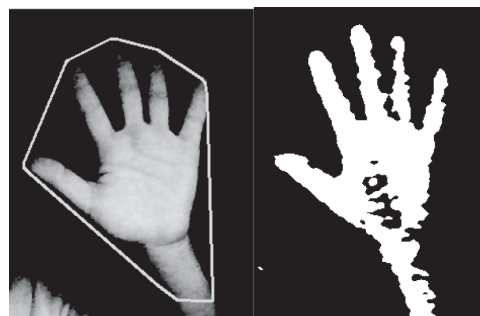


图2 Otsu 阈值分割结果图

色, M_x 和 M_y 为这两个部分的均值。P 是像素的数量占总的像素数量的比例, P_x 就是 X 部分的比例, 而 Y 部分的相应比例记作 P_y 。在 Otsu 中, $ICV = P_x \cdot (M_x - M)^2 + P_y \cdot (M_y - M)^2$, ICV 取最大值时, 阈值此时为最优解。图 2 为 Otsu 阈值分割结果。

再用单指尖检测和多指尖检测来对人体手势进行检测, 其中单指尖检测选择重心距离法, 计算手区域的重心位置, 然后在整个手型边缘中寻找满足距离重心最远的点, 并判断其是否是指尖, 图 3 为单指尖检测图像, 蓝色点为重心, 红色点为指尖点。

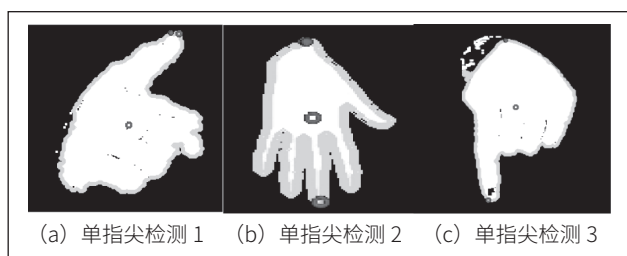


图 3 单指尖检测图像

多指尖检测是将手势图像进行细化, 转换为像素点, 对这些点进行识别筛选, 找到指尖点。本文中利用深度学习中的卷积神经网络结构进行实现, 在这之前, 扩充手势库, 通过傅里叶变换得到类标签, 增加图像的旋转平移不变性, 用来提高识别准确率, 基于全卷积神经网络手势分割算法的分割结果图如图 4 所示。



图 4 手势分割结果图

3 手势识别和分析

在进行手势识别之前, 应该建立手势数据库, 通过摄像头拍摄一系列规定的手势, 并将图像保存到文件中, 利用傅里叶变换具有旋转、平移不变性, 先对图像轮廓建立坐标系, 将组成轮廓边缘的每个点强制转换成复数表示形式, 对于轮廓图像, 用一个一维方程来表示为:

$$c(t) = x(t) + j \cdot y \quad (5)$$

然后, 对公式 (5) 进行傅里叶级数展开:

$$c(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(k\omega t) + b_k \sin(k\omega t) \quad (6)$$

通过傅里叶描述子对手势库中每个手势图像进行傅里叶变换得到类别标签, 存放在相对应的 txt 文件中, 具有识别准确率高, 受角度影响小的优势。将存储的 txt 文本转换成向量格式, 将类别标签作为 SVM 模型的输入, 如图 5 所示, SVM 模型在测试时的运行结果。

```
{'C': 1, 'gamma': 0.125, 'kernel': 'linear'}
SVM Model save...
总共错了0个数据
错误率为0.000000%
```

图 5 SVM 模型训练结果

另外, 由于背景中可能出现物体颜色干扰识别结果, 还有一些噪声干扰, 因此需要进行去噪处理和形态学处理, 增强图像的效果。本系统利用 OpenCV 中提供的低通滤波器来对图像的区域进行检查, 改变该像素的亮度, 由于肤色区域设定为白色, 因此, 将噪点转为黑色。

形态学处理的目的是与之类似, 也是为了去除一些干扰因素, 在分割出的黑白图像中, 经常会在图像的不对应区域出现杂质点, 而去除这些点的方法是通过腐蚀和膨胀两种操作。腐蚀操作即是对图像进行“瘦身”, 计算图像周围的极大值, 并给参考点像素赋值, 这样可以使白色的部分增长; 膨胀操作正好相反, 即是对图像进行“增肥”, 计算局部的最小值, 并赋给参考点, 这样白色区域将减少。本系统通过形态学的开运算和闭运算来对图像进行处理, 开运算即是先腐蚀后膨胀, 适用于去除图像中孤立的小点, 闭运算即是先膨胀后腐蚀, 去除图像中一些裂缝形状的噪点。

4 结论

本文主要提出了针对人机交互背景下的手势分割、识别和手指的单 / 多指尖检测方法, 在复杂的外界环境下能够对人体手势进行精确识别。利用 YCrCb 方法来提取出手势区域, 之后再利用单指尖检测和多指尖检测方法来对手势进行分割, 并利用深度学习中的卷积神经网络算法精准的得到图像特征, 使人体的手势表现出丰富的数据信息。

参考文献

- * [1] 易生, 梁华刚, 茹锋. 基于多列深度 3D 卷积神经网络的手势识别 [J]. 计算机工程, 2017, 43(08): 243-248.
- * [2] 张勋, 陈亮, 朱雪婷, 胡诚. 基于区域卷积神经网络 Faster R-CNN 的手势识别方法 [J]. 东华大学学报 (自然科学版), 2019, 45(04): 559-563.
- * [3] 姜洋洋. 基于卷积神经网络与 CUDA 加速计算的手势识别算法应用研究 [J]. 系统仿真技术, 2020, 16(01): 22-26.
- * [4] 刘于嘉, 李传江, 杜芸菲, 黄晴. 基于手势识别的教育机器人 (下转第 54 页)

到 C 的信息传送。而节点 C 发往 A 的信息则是由 S2/P1 光纤的前半时隙来携带,经节点 B 穿通到达节点 A,完成由 C 到 A 的信息传送。

而如果链路出现故障,如图 4 所示,AB 节点间链路中断,此时故障点相连的节点 A 和 B 进行桥接倒换,利用时隙交换技术,将 S1/P2 光纤和 S2/P1 光纤上的工作时隙业务信号移到另一根光纤上的保护时隙,则 A 发往 C 的信息由 S2/P1 光纤的后一半时隙 P1 携带,经节点 D 和 C 到达 B,然后节点 B 进行倒换,将 S2/P1 光纤的 P1 上的业务倒换到 S1/P2 光纤的 S1 上,再由 S1 携带传送到 C,完成由 A 到 C 的信息传送。而对于 CA 间的业务,由 B 节点进行倒换,将 S2/P1 光纤上 S2 携带的信息转换到 S1/P2 光纤上的后半时隙 P2 上,再经 C、D 传送到 A,完成由 C 到 A 的信息传送。这就是发生故障后,二纤双向复用段共享保护环的工作过程,网络业务得以维持,不会中断。

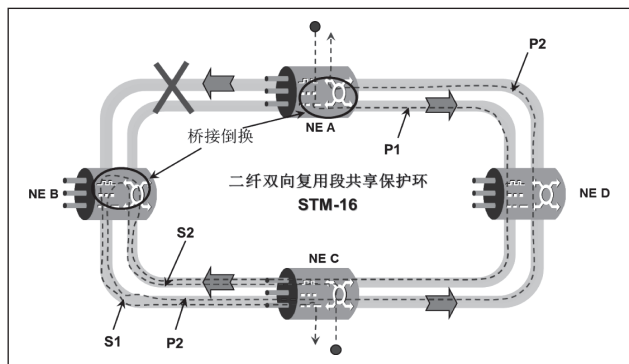


图4 桥接倒换后的业务流向图

二纤双向复用段共享保护环还有一个特点,就是在正常情况下,工作时隙传送工作业务,而保护时隙可以传送可靠性要求不高的额外业务,一旦发生故障,保护时隙将中断额外业务,通过故障两端的网元的环倒换功能,由保护时隙完成工作业务的传送,这样,时隙利用率较高,多用于业务量大,业务分散的网络。

2 二纤单向复用段保护

如图 5 (a) 所示,二纤单向复用段保护环,在正常情况下,

低速支路信号仅仅从 S1 光纤进行分插,保护光纤 P1 是空闲的。如果 BC 节点间光缆被切断,与光缆切断点相邻的两个节点 B 和 C 的保护倒换开关将利用 APS 协议转向环回功能,如图 5(b) 所示,对于 AC 间的业务:在 B 节点, S1 光纤上的业务信号 (AC) 经倒换开关从 P1 光纤返回,沿逆时针方向经 A 节点和 D 节点仍然可以到达 C 节点,并经 C 节点倒换开关环回到 S1 光纤并落地分路。对于 CA 间的业务:由于业务是经过 D 点在 S1 光纤上进行传输的,不受断纤的影响,与正常时传输情况相同。

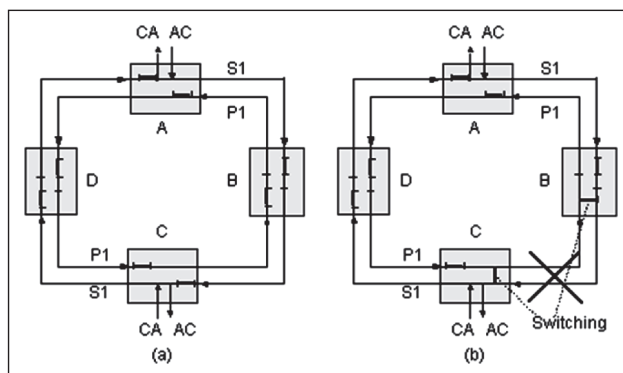


图5 二纤单向复用段保护环示意图

3 结束语

本文详细分析了光传输 SDH 网络环形网的二纤双向复用段共享保护方法以及二纤单向复用段保护环的工作原理,图文并茂分析了正常情况和故障情况下各业务的流向。希望能给读者有所启发,也希望各位专家批评指正。

参考文献

- * [1] 华为技术有限公司 .SDH 原理手册 [Z].2007
- * [2] 胡庆,等 . 光纤通信系统与网络 [M]. 北京: 电子工业出版社, 2019.
- * [3] 孙桂芝,等 . 光传输网络组建与维护案例教程 [M]. 北京: 机械工业出版社, 2014.

(上接第 74 页)

- 人机交互技术研究 [J]. 电子测试 ,2019(23):17-19.
- * [5] 斯琴其木格,田英爱,高丹阳,朱慧敏 . 基于 Arduino 的多重手势识别传感器的设计与实现 [J]. 数字技术与应用 ,2019,37(11):134-137+139.
- * [6] 曹书豪,许成哲 . 基于深度学习模型的表面肌电信号手势动作识别算法研究 [J]. 科技视界 ,2019(01):118-120.

- * [7] 惠丹 . 基于深度学习算法的可穿戴设备手势识别系统设计 [J]. 信息技术与网络安全 ,2019,38(09):30-33.
- * [8] 张勋,陈亮,胡诚,孙韶媛 . 一种基于深度学习的静态手势实时识别方法 [J]. 现代计算机 (专业版) ,2017(34):6-11.
- * [9] 杨益平,闵啸 . 基于计算机视觉的手势识别人机交互技术 [J]. 电子技术与软件工程 ,2018(12):138-139.