

个人工作

1. 问题

目前已经存在很多用于人体姿态检测的特征，但是特征的选择依旧是目前计算机视觉，模式识别等领域的重点课题，我所解决的问题是通过对原有的特征进行改进从而形成一种新的特征，并利用现有的模型来证实这种新特征能够取得很好地效果。

2. 技术路线

对于一张图片，先计算梯度图得到 X 方向的梯度图和 Y 方向的梯度图，再通过计算得到图片的梯度图，梯度图包括图中每一个像素的梯度大小和梯度方向，接着设立一个固定大小的块，在实验中我们设定块大小为 3×3 ，用固定块依次扫描梯度图提取出多个特征块，提取的数量为图片的宽减去二与高减去二的积，我们假设提取了 N 个块，然后计算每一块的 HOG，与之前的 HOG 不同，我们不是计算包含四个小块的大块的 HOG 而仅仅计算每个固定块的 HOG，因此每一块的 HOG 维度为 9，即将 3×3 大小的块映射在 9 个方向的直方图中，然后我们把 N 个块的 HOG 图从左到右排列起来形成整张图的 HOG，由上可知 HOG 大小为 $9 \times N$ ，用来作为整张图的特征图。

接着我们将整张图的 HOG 当做输入来计算它的稀疏表达，我们提前训练好一个字典，因为要配合之前的 HOG，因此字典的高为 9，宽我们自己设定为 M ，这个 M 代表的是字典的完备性， M 越大意味着字典越完备，我们将字典与 HOG 作正交匹配追踪得到各个块的稀疏表达，每一个块的稀疏表达大小为 $M \times 1$ ，于是最后我们得到了整张图的梯度特征的稀疏表示图，大小为 $M \times N$ ，我们将其称之为梯度稀疏特征图。

接着我们将梯度稀疏特征图作为输入，使用目前最流行的融合多姿态（FMP）模型对特征图进行处理，其中包括了卷积，动态规划等过程，经过一系列的计算得到图中人物的最优姿态估计，接着我们将人物的最优姿态估计与图中本身具有的姿态标准比较来计算正确点的百分比（PCK）。对于不同的数据集，我们设置了不同的训练集测试集，由训练集来调整模型的参数，如滤波器等，再由测试集来验证最终得到的效果。

3. 工作

实验中，HOG 的概念是 Dalal 等人在 2005 年的 CVPR 会议上提出的，稀疏编码直方图由 Ramanan 在 2013 的 CVPR 会议上提出，融合多姿态(fmp)模型由 yi yang 在 2013 年提出，关键点百分比(pck)也是由 yi yang 在 2013 年的论文中一并提出，所使用的正样本数据集是别人提供的 PARSE 数据集，LSP 数据集和 MPII 数据集，负样本数据集是 INRIA 数据集

我们的工作是基于上述的工作提出了一个新的特征，在此称为梯度稀疏直方图，并使用了现有的模型进行实验，这个特征的好处在于：相比于 HOG，我们的特征能解决 HOG 所不能解决的问题，HOG 虽然对变化明显的边表现的很敏感，但是 HOG 在复杂场景下，尤其在遮挡问题比较严重的图片中就难以体现身体部位的信息，新特征因为稀疏编码的存在，使其不仅仅是受梯度信息的约束，因此能够解决遮挡问题，而与之之前的稀疏编码直方

图相比，由于原来的稀疏编码直方图所提取的底层特征是颜色特征，而很明显颜色特征在目前的效果是不如梯度特征的，因为颜色特征有时会形成混淆或者产生多余信息，对于同样一幅图，基于图的梯度信息进行稀疏表达比直接对颜色信息进行稀疏表更能提取有用的信息，这就能去除一些没有意义的局部图像。而且，HOG 和稀疏编码的结合，对降维也有好处，HOG 和稀疏编码减少非零数字的出现的特性，能提高模型的性能。总的来说，我们的工作就是先以小块为单位计算最初的 HOG，去掉后面的合并过程，然后分别对 HOG 进行稀疏表达，将得到的特征排列来得到最终特征，这种特征结合了 HOG 和稀疏表达的优势。

4. 工作量

实验的工作量主要是写梯度稀疏直方图的代码，并调整模型的相关参数使得模型能够以我们的梯度系数直方图作为输入，此外就是更改数据集的图片姿态标准的表达方式，使我们能够计算 PCK 从而得到实验效果。