Hi-CMD:可见光-红外行人重识别的分层交叉解纠缠

**主要贡献：**

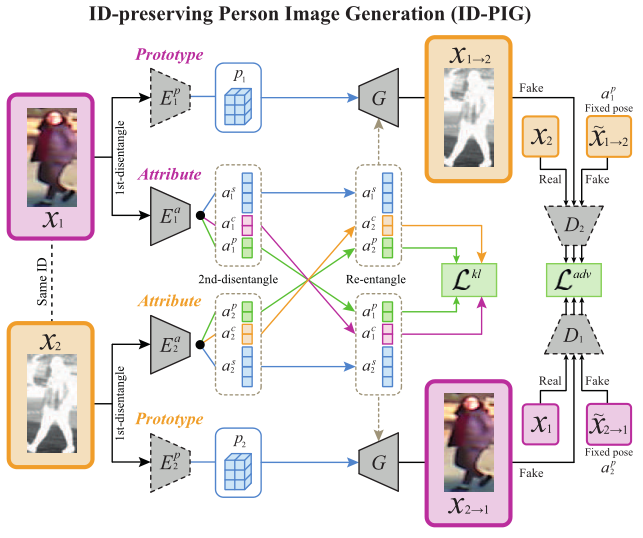
1. 提出了Hierarchical Cross-Modality Disentanglement (Hi-CMD) 方法，它是一种高效的学习结构，可以提取姿势不变和光照不变的特征进行交叉模态匹配。这是第一个从VI-ReID的跨模态图像中同时分离ID-discriminative factors和ID-excluded factors的工作。
2. 提出ID-preserving Person Image Generation(ID-PIG)网络，使改变姿势和照明属性同时保持特定人物的身份特征成为可能。

**方法概述：**

我们将可见光图像和红外图像分别表示为和，其中H和W分别为图像的高度和宽度。图像和中的每一个都对应一个标识标签y∈{1,2，…， N}，其中N是人的身份数。在训练阶段，使用多模态图像集和在特征提取网络φ(·)中训练。在测试阶段，给定一种模态的查询图像，计算另一种模态的图库集中的排序列表。两个特征向量φ(x1)和φ(x2)之间的距离由欧氏距离计算。

在VI-ReID任务中，最具挑战性的问题是可见光和红外图像之间同时存在跨模态和内模态差异。Hi-CMD方法旨在从跨模态图像中分离ID-discriminative factors和ID-excluded factors，同时减少跨模态和内模态的差异。提出了ID-preserving Person Image Generation (ID-PIG)网络和Hierarchical Feature Learning (HFL)模块

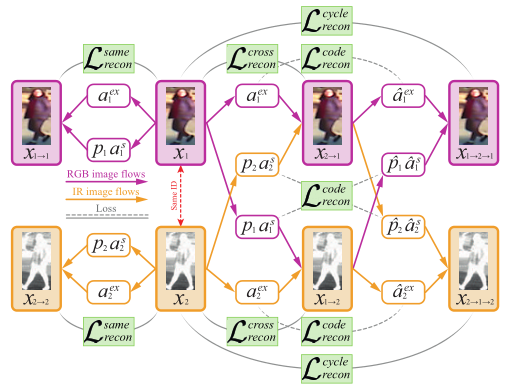
**ID-PIG网络**



ID-PIG 网络采用二阶层次化。第一层次采用原型编码器和属性编码器，原型编码器对图像中行人的体态等外貌特征进行编码，属性编码器对衣服风格、姿态、光照这类可变属性进行编码。第二层次属性编码划分成三类 ，具体为：风格属性编码、光照属性编码、姿态属性编码。这三种编码分别表示：风格属性是对行人的衣服结构进行编码；光照属性编码对应模态之间的差异，把不同RGB、IR摄像头的视觉差异定义为光照属性；姿态属性对应模态内的差异，理解为同一个模态内行人的多种姿态。

最终光照、姿态属性作为ID-excluded编码，而风格属性、原型编码作为ID-discriminative编码。

**损失函数说明**



1. **跨模态重构损失函数:**

在图像生成过程中，我们的主要策略是通过交换ID相同的两幅图像的ID-excluded来合成一对跨模态图像。保证原始图像的体态（原型编码）和衣服结构（风格属性），替换模态（光照属性）和姿态（姿态属性），重构生成的图像要与对应模态的样本图像尽可能接近。形式上，此跨模态重构损失表述如下:





1. **同模态重构损失函数：**

除了对不同模态图像的重构损失外，我们还应用了对相同模态图像的重构损失。对同模态的四个编码，重新生成原始图像。



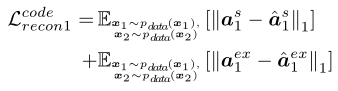
1. **cycle重构损失函数：**

两次跨模态重构



1. **编码损失函数:**

 同模态的编码需要尽可能接近



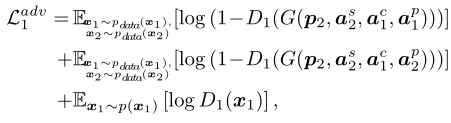
**总的重建损失函数**



1. **KL散度损失**



1. **对抗损失**



将ID-discriminative编码加权级联得到判别向量，再传入全连接层，得到最终的特征向量。ReID损失函数包含交叉熵损失和三元组损失，训练数据采用训练得到的生成器进行数据增强，生成相同ID但不同姿态、光照属性的行人图片