Hybrid Modality Metric Learning for Visible-Infrared Person

Re-Identification

基于混合模态度量学习的可见红外人再识别

LA ZHANG, HAIYUN GUO, KUAN ZHU, HONGLIN QIAO, GAOPAN HUANG,

SEN ZHANG,HUICHEN ZHANG,JIAN SUN,JINQIAO WANG

**摘要：**

本文将混合模态分为两种类型，模态内和跨模态。将VI-ReID任务中有影响的变量分为三类，：模态内变量、跨模态的模态相关变量和跨模态的模态无关变量。然后提出了四种基于配对的相似性约束的综合度量学习框架，以解决模态内和模态间的所有变化。该框架关注人物图像之间的类级和模态级相似关系。此外，分别将其与三元组损失和对比损失相结合详细实现，证明了此框架与任何基于配对的损失函数的居然较强的兼容性。

**1.引言**

对于三个图像样本，即查询图像、正样本、负样本，定义了以下变量：

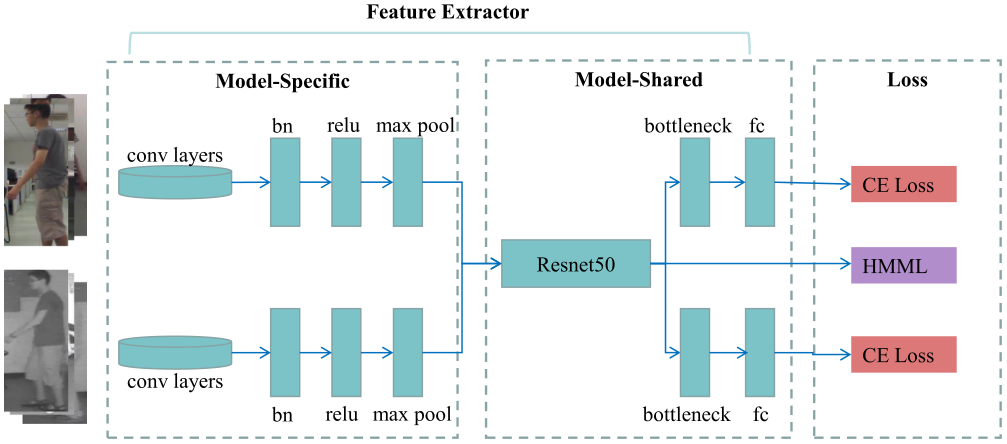
1. **模态内变量：**模态内是指所有样本都来自同一模态。模态内部的变化主要是由人体姿势的变化、摄像机视点的变化和光照的变化引起的，可能导致类内相似度小，类间相似度大。单模态人Re-ID只需要处理这种类型的差异。
2. **跨模态的模态相关变量：**对于来自不同模态的三个样本，我们将结果中可能干扰其相似性关系的变量定义为跨模态变量。根据变量是否由不同模态带来，我们进一步将其分为模态相关变量和模态无关变量。
3. **跨模态的模态无关变量：**除了模态差异之外，模态无关的变量也会破坏跨模态人物匹配的相似性排序结果。较大的姿态变化使得正样本RGB-RGB图像对之间的距离甚至大于负样本RGB-IR图像对的距离。这种变量对应于模态内变量的极端情况，其中由模态无关因素引起的变量甚至大于跨模态差异。我们将这种变量与模态内变量分开来定义，以强调这种极端情况。

**2.主要贡献**

提出了一个基于类级和模态级相似性约束的综合混合模态度量学习框架，以应对所有类型的变化，并将其分别与三元组损失和对比度损失相结合。

**3.方法介绍**

**3.1整体框架**



**3.1具有类级和模态级相似性约束的度量学习框架**

使用V和T来表示RGB图像集和IR图像集。让x∈V表示RGB图像，z∈ T表示红外图像。给定特定人的RGB图像x（或IR图像z），和表示正样本图像和负样本图像。我们将特征提取器表示为f(·)，所提取的x和z的特征由f(x)和f(z)表示。欧几里得距离D(·, ·)用于测量图像对的相似性，例如x和z之间的欧几里德距离D(x, z)表示为:



给定一组图像，定义以下四种类型的约束来解决模态内和模态间的所有变化：

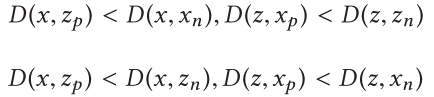
模态内相似性约束（WM约束）



跨模态的模态无关相似性约束（CM\_U约束）



跨模态的模态相关相似性约束（CM\_S约束和CM\_G约束）

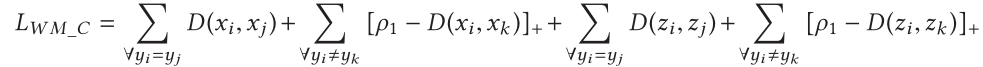


**3.2基于对应的损失函数的兼容性分析**

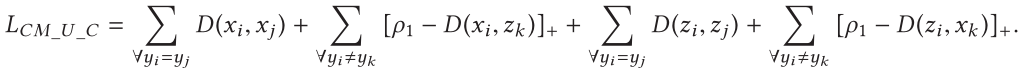
假设和的表示相同的行人，与表示不同的行人。

**3.2.1基于对比损失的HMML**

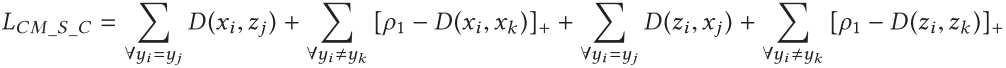
WM损耗用于模态内，其中正样本图像对，负样本对，代表阈值。

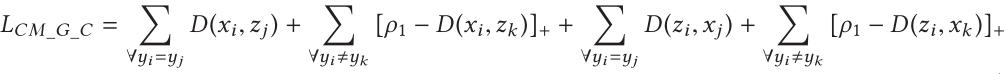


CM\_U约束解决模态内变化的极端条件。



CM\_S约束和CM\_G约束。





**3.2.2基于三元组损失的HMML**



