**公开(公告)号:** [CN110019652B](https://analytics.zhihuiya.com/patent-view/abst?patentId=dd66f9a4-6c47-4722-a184-7505ecaeb106&shareId=4E00E206-F06F-7872-B147-71GD79EF4E4D&from=EXPORT&signature=Bkjrky27gp4aVEqQnlvrQaxrza2b7LdgtEBGeZdBEp8%3D&expire=94608000&date=20240403T004109Z&version=1.0)

**标题:** 一种基于深度学习的跨模态哈希检索方法

**法律状态/事件:** 授权

**当前申请(专利权)人:** 九江学院

**应用领域分类:** 静止图像数据索引|文本数据库索引|特殊数据处理应用

**申请日:** 2019-03-14

**公开(公告)日:** 2022-06-03

**摘要:** 一种基于深度学习的跨模态哈希检索方法，假设个对象的图像模态的像素特征向量集为，其特征是该方法包括以下步骤：（1）使用基于深度学习技术设计的目标函数得到图像模态和文本模态共享的二进制哈希编码，图像模态和文本模态的深度神经网络参数和，以及图像模态和文本模态的投影矩阵和；（2）使用交替更新的方式求解目标函数中的未知变量、、、和；（3）基于求解得到的图像模态和文本模态的深度神经网络参数和，以及投影矩阵和；（4）基于生成的二进制哈希编码计算查询样本到检索样本集中各个样本的汉明距离；（5）使用基于近似最近邻搜索的跨模态检索器完成对查询样本的检索。该方法有效地提升了跨模态哈希检索的性能。

**第一权利要求:** 1.一种基于深度学习的跨模态哈希检索方法，假设n个对象的图像模态的像素特征向量集为其中，vi表示第i个对象在图像模态的像素特征向量；令表示这n个对象在文本模态的特征向量，其中，ti表示第i个对象在文本模态的特征向量；将n个对象的类别标记向量表示为其中，c表示对象类别的数量；对于向量yi来说，如果第i个对象属于第k类，则令向量yi的第k个元素为1，否则，向量yi的第k个元素为0；其特征在于，该方法包括以下步骤：(1)使用基于深度学习技术设计的目标函数得到图像模态和文本模态共享的二进制哈希编码B，图像模态和文本模态的深度神经网络参数θv和θt，以及图像模态和文本模态的投影矩阵Pv和Pt；(2)使用交替求解的方式求解目标函数中的未知变量B、θv、θt、Pv和Pt，即交替的求解如下三个子问题：固定B、Pv和Pt，求解θv和θt；固定B、θv和θt，求解Pv和Pt；固定θv、θt、Pv和Pt，求解B；(3)基于求解得到的图像模态和文本模态的深度神经网络参数θv和θt，以及投影矩阵Pv和Pt，为查询样本和检索样本集中的样本生成二进制哈希编码；(4)基于生成的二进制哈希编码计算查询样本到检索样本集中各个样本的汉明距离；(5)使用基于近似最近邻搜索的跨模态检索器完成对查询样本的检索；所述步骤(1)中的基于深度学习技术设计的目标函数形式如下：其中，γ1和γ2为非负平衡因子，B＝[b1,b2,…,bn]T∈{-1,+1}n×k，和为投影矩阵，θv和θt为深度神经网络参数，和分别为n个对象在图像模态和文本模态的深度特征，并且矩阵F和矩阵G的第i列的向量分别为f(vi；θv)和g(ti；θt)，为拉普拉斯矩阵用于保持模态内一致性和模态间的一致性，1为全部元素为1的列向量，||·||F表示矩阵的Frobenius范数，tr(·)表示矩阵的迹，(·)T表示矩阵的转置。

**IPC分类号:** G06F16/31 | G06F16/51