

荣梦君, 刘惊雷

烟台大学 计算机与控制工程学院, 山东 烟台 264005

E-mail: yturmj@163.com (荣梦君), jinglei_liu@sina.com (刘惊雷)

论文摘要

图像检索在计算机领域是一项重要任务, 如何在图像数据库中快速找到查询项的相关图像是一个研究课题。近年来, 哈希方法以其存储成本低、检索效率高等优点在图像检索中得到广泛应用。现有的哈希方法大多采用紧凑的二进制码降低计算量, 然而没有最优的方法可以快速获得对原始特征空间的近似映射。针对该问题提出基于Haar小波投影的哈希方法, 实现对原始特征向量的快速投影。首先, 构造基于Haar小波变换的投影矩阵, 并对投影矩阵和二进制码进行迭代优化, 为了控制量化误差, 采用离散方法优化二进制码。然后, 原始特征向量与投影矩阵相乘, 将原始特征空间中的数据近似映射到低维特征空间。最后, 对低维特征空间中的数据点进行二进制嵌入处理, 利用得到的二进制码实现图像检索。该方法不仅将时间复杂度从传统方法的 $O(d \log d)$ 降低到 $O(d)$, 而且空间复杂度为 $O(d \log d)$ 。在图像数据集上的大量实验结果表明, 该方法的时间代价小, 并且可以保证可靠的检索精度。

论文简介

- 研究背景: 信息技术的飞速发展和图像资源的大量应用使图像检索成为计算机科学领域的重要课题, 哈希方法因其较高的检索效率和存储优势等优点在图像检索领域备受关注。哈希方法的目的是将原始特征空间中的数据点投影后嵌入到汉明空间, 在汉明空间中计算数据点之间的相似度。
- 问题: 在哈希方法中, 学习二进制码一般分为投影和二进制嵌入两个步骤。利用投影获得对原始特征数据的近似映射会带来计算成本过高和检索精度不足的问题。
- 方法提出: 本文提出一种基于Haar小波变换的投影方法, 其思想是利用Haar小波变换将投影矩阵分解为几个结构化矩阵的乘积, 实现快速投影。为了控制过高的精度损失, 采用离散方法优化二进制码, 即对投影后的特征向量中的每个元素独立地进行二进制嵌入。该方法具有显著的时间优势, 并且能够保持可靠的检索精度。

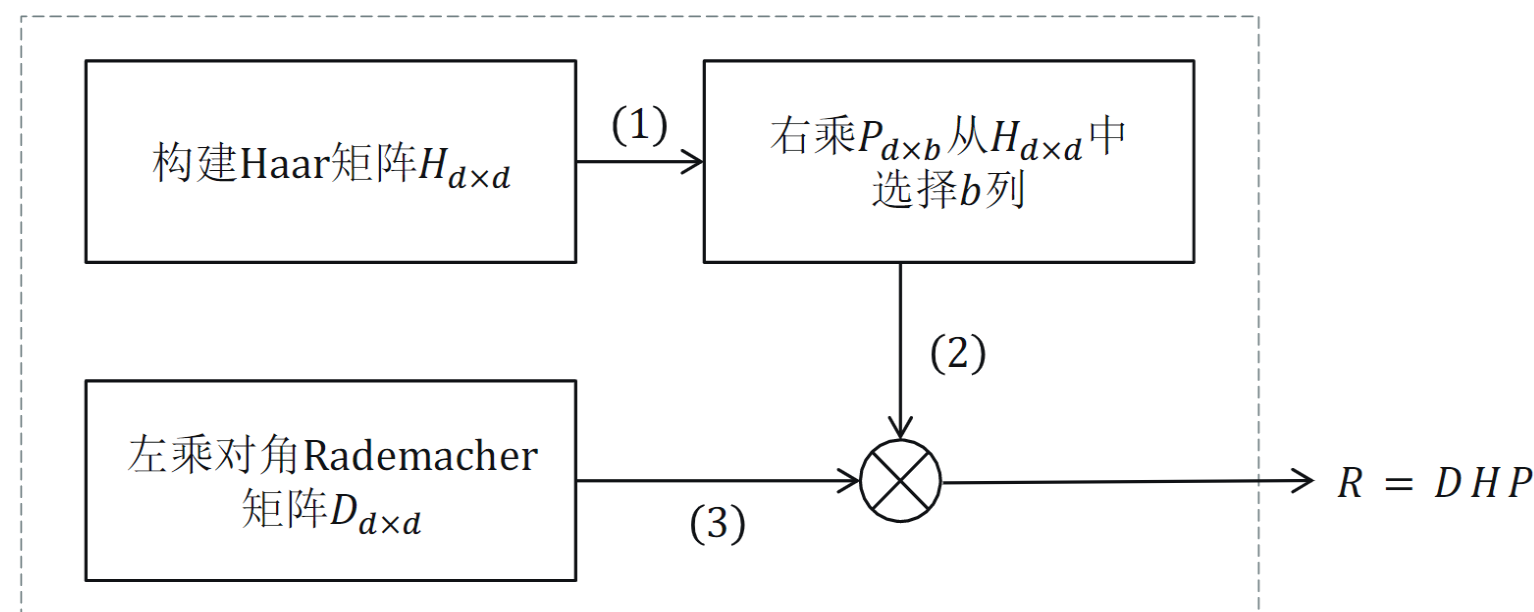


图1 Haar小波投影矩阵结构

方法框架

- 首先, 构造投影矩阵, 利用Haar小波变换将投影矩阵分解为几个矩阵的乘积, 即用几个小的矩阵表示一个大的投影矩阵, 并通过训练数据迭代优化得到最终的投影矩阵 R , 实现对高维数据的快速投影。
- 然后, 图像数据集原始特征矩阵与投影矩阵相乘, 将原始特征空间中的数据近似映射到低维特征空间, 低维特征矩阵二进制嵌入得到图像数据集的二进制码。
- 最后, 将查询图像通过投影和二进制嵌入得到查询图像的二进制码, 并与图像数据集的二进制码进行相似度比较, 得到检索结果。

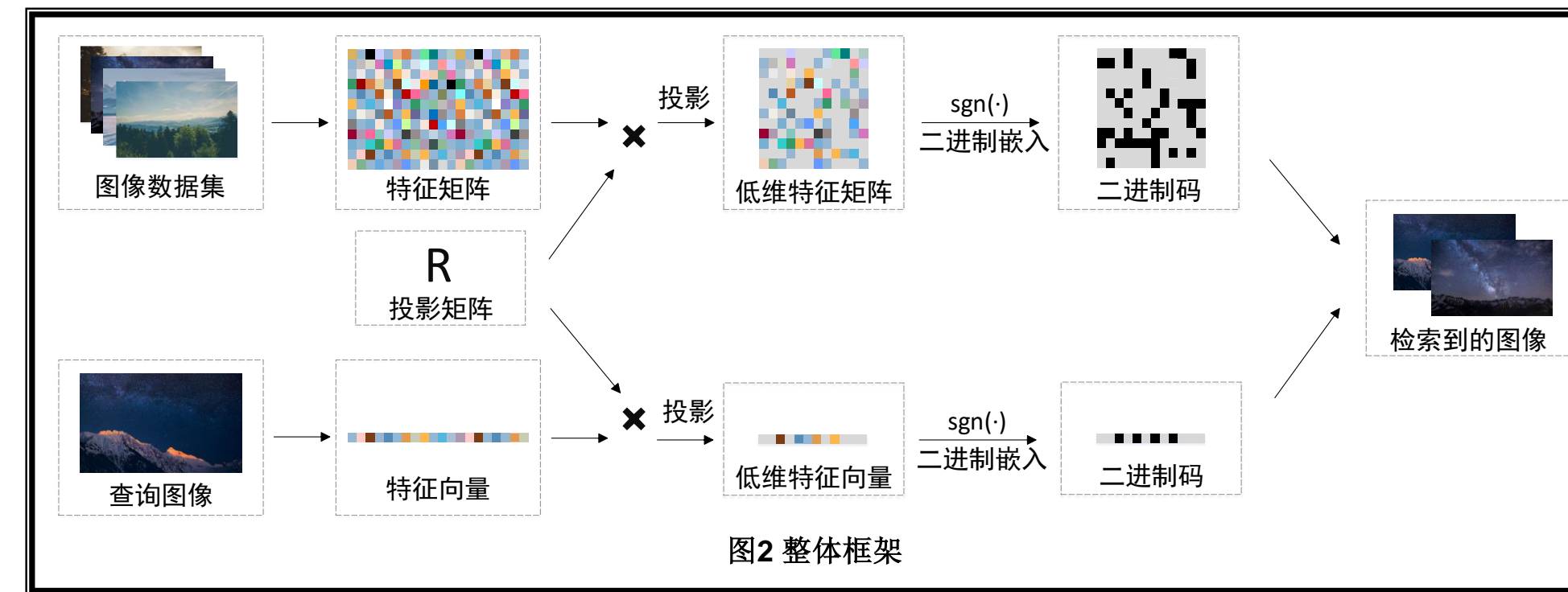


图2 整体框架

算法设计

将目标函数分解为3个子问题, 分别迭代求解, 设计基于Haar小波投影的哈希算法。

算法 1: 基于Haar小波投影的哈希算法

Input: 原始数据矩阵 $X \in \mathbb{R}^{n \times d}$

Output: 二进制码 B

初始化置换矩阵 P ;

初始化对角Rademacher矩阵 D ;

初始化Haar矩阵 H ;

初始化迭代数 $iteration$, 二进制码长度 bit 以及惩罚因子 β ;

$R = \frac{1}{\sqrt{b}} DHP$;

for 0 : $iteration$ do

$Y = (B + \beta XR)/(1 + \beta)$;

对 $Y^T X$ 进行奇异值分解得到 $Y^T X = U \Sigma V^T$;

$\bar{R} = V U^T$;

$B = \text{sgn}(X \bar{R})$;

$Z = X \bar{R}$;

$S = \frac{1}{\sqrt{b}} X, E = HP, K = \frac{1}{\sqrt{b}} H P Z^T X$;

$\text{diag}(D) = [(E E^T) \circ (S^T S)]^{-1} \times \text{diag}(K)$;

end

$B = \text{sgn}(XR)$;

return B

实验仿真

实验数据集

表1 数据集概述

数据集	数据量	特征维度
SIFT10K	25000	128
LabelMe	22019	512
CIFAR-10	60000	512
Caltech-256	29358	1024

编码时间比较

表2 不同哈希方法在4个数据集上的编码时间

方法	SIFT10K@64bits	LabelMe@64bits	CIFAR-10@64bits	Caltech-256@64bits
LSH	0.0457	0.0542	0.1452	0.1167
SKLSH	0.0731	0.0609	0.1443	0.1240
SH	0.8244	1.1472	2.6807	1.0296
ITQ	0.0399	0.0521	0.1404	0.1150
BP	0.0504	0.0553	0.1650	0.1110
CBE	0.8144	1.2900	5.3796	3.5275
Ours	0.0246	0.0390	0.0900	0.0909

检索性能比较

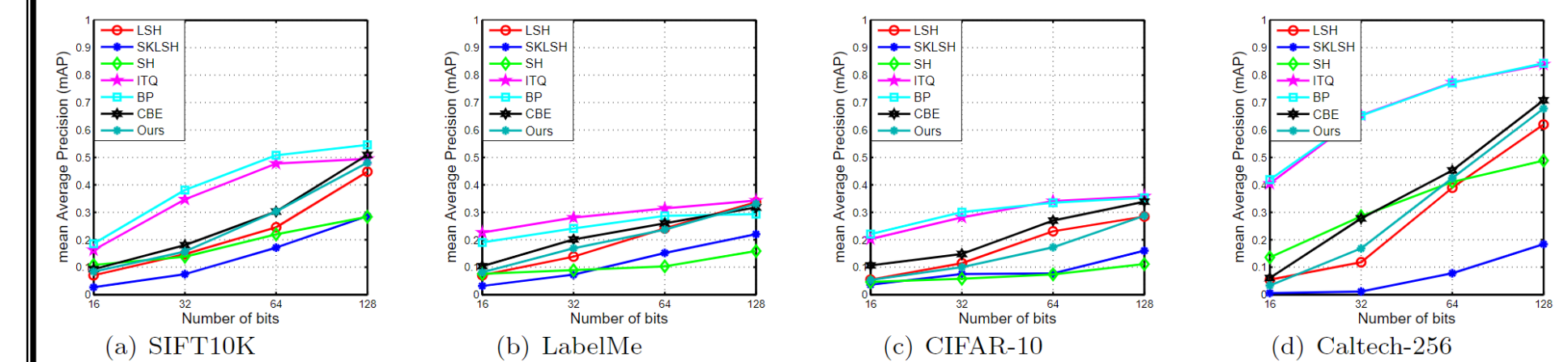


图3 数据集SIFT10K, LabelMe, CIFAR-10和Caltech-256上不同哈希方法的mAP曲线

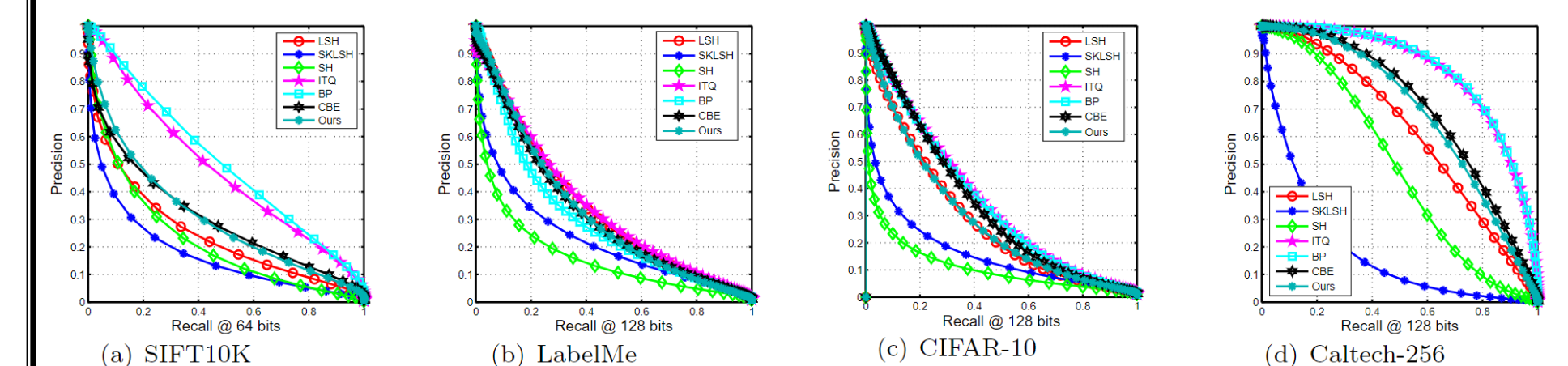


图4 数据集SIFT10K, LabelMe, CIFAR-10和Caltech-256上不同哈希方法的PR曲线

论文结论

本文将哈希方法应用于图像检索, 提出了一个基于Haar小波投影的哈希算法框架, 通过由Haar小波变换构造的投影矩阵实现对高维图像数据的维度约简, 再对投影后的数据进行二进制嵌入, 最后利用得到的二进制码实现图像检索, 该框架的核心是利用Haar小波变换矩阵进行快速投影。实验结果表明该方法的时间复杂度和存储性能具有明显的优势。 $d \times d$ 阶投影矩阵与向量相乘的时间复杂度为 $O(d)$, 最大存储代价为 $O(d \log d)$ 。与几种主流的哈希方法相比, 该方法最显著的优势是缩短了投影时间, 编码效率得到大幅提升, 并且在检索精度方面的表现是可靠的。