

# 各向同性的迭代量化哈希算法

李来<sup>1,2</sup>, 刘光灿<sup>1,2</sup>, 孙玉宝<sup>1,2</sup>, 刘青山<sup>1,2</sup>

(1. 江苏省大数据分析技术重点实验室, 江苏南京 210044; 2. 南京信息工程大学信息与控制学院, 江苏南京 210044)

**摘要:** 准确有效的哈希算法是实现海量高维数据近邻检索的关键. 迭代量化哈希 (Iterative Quantization, ITQ) 和各向同性哈希 (Isotropic Hash, IsoHash) 是两种知名的编码方法. 但是 ITQ 算法对旋转矩阵施加的约束过于单薄, 容易导致过拟合; 而 IsoHash 算法缺乏对哈希编码的更新策略, 降低了编码质量. 针对上述问题, 提出了一种各向同性的迭代量化哈希算法. 该方法采用迭代的策略, 对编码矩阵和旋转矩阵交替更新, 并在正交约束的基础上增加各向同性约束来学习最优旋转矩阵, 最小化量化误差. 在 CIFAR-10、22K LabelMe 和 ANN\_GIST\_1M 基准库上与多种方法进行对比, 实验结果表明本文算法在查准率、查全率以及平均准确率均值等指标上均明显优于对比算法.

**关键词:** 哈希; 大规模图像检索; 各向同性; 迭代量化

**中图分类号:** TP311 **文献标识码:** A **文章编号:** 0372-2112 (2017)07-1707-08

**电子学报 URL:** <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2017.07.022

## Isotropic Iterative Quantization Hashing

LI Lai<sup>1,2</sup>, LIU Guang-can<sup>1,2</sup>, SUN Yu-bao<sup>1,2</sup>, LIU Qing-shan<sup>1,2</sup>

(1. Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing, Jiangsu 210044, China;

2. School of Information and Control, Nanjing University of Information Science & Technology, Nanjing, Jiangsu 210044, China)

**Abstract:** Hashing is a key technique to achieve fast nearest neighbor search in high-dimensional, massive datasets. Among various methods, Iterative Quantization (ITQ) and Isotropic Hash (IsoHash) are probably the most popular ones due to their high retrieval accuracy. However, as the constraints imposed on the rotation matrix are too weak, the optimization problem in ITQ is severely under-deterministic and therefore easy to cause over-fitting. In IsoHash, the isotropic projection matrix is updated in a manner that is completely independent of the binary hash codes, and thereby the quality of the produced hash codes may be depressed. To address these issues, this paper proposes an isotropic iterative quantization hashing method, which extends the formulation of ITQ by incorporating properly the isotropic prior adopted in IsoHash. In our method, the hash code matrix and rotation matrix are updated alternately in an iterative fashion. Experiments are conducted on three benchmark datasets, CIFAR-10, 22K LabelMe and ANN\_GIST\_1M. The results show that the proposed method performs better than the competing methods in terms of precision, recall and mAP.

**Key words:** hashing; large-scale image retrieval; isotropic prior; iterative quantization

## 1 引言

随着科技的发展, 获取到的图像数据也日益呈现出海量、高维的态势, 给图像检索、分类及其它相关视觉工作带来了严峻的挑战. 因此, 研究出一种能在大规模数据中进行快速近邻检索的方法是很有必要的. 传统线性查找以及基于树<sup>[1]</sup>的检索方法主要用于处理低维数据, 难以应对“维度灾难”的问题<sup>[2]</sup>. 近年来, 国内外

提出了基于哈希的检索方法<sup>[3-11]</sup>, 其原理是将原始高维数据映射到低维空间, 并保证量化后的哈希编码仍能保持数据在原始空间中的相对位置关系, 具有查询速度快, 存储和通信开销低的优点, 被认为是一种很有前景的检索方法.

哈希算法大体上可以分成两类: 独立于数据的算法<sup>[2]</sup>和基于数据的算法<sup>[12-15]</sup>. 为了避免算法中目标函数求解过程中的非凸性问题, 通用的做法是将哈希分

收稿日期: 2015-12-07; 修回日期: 2016-09-14; 责任编辑: 李勇锋

基金项目: 国家自然科学基金 (No. 61622305, No. 61502238, No. 61532009, No. 61672292); 江苏省自然科学基金 (No. BK20160040); 江苏省六大大人才高峰 (No. DZXX-037)

成两步:学习投影矩阵和量化编码.其中,有无投影矩阵的学习过程,是区分两类哈希算法的主要标准.第一类哈希算法中,通常采用随机生成投影矩阵的方式,不存在相应的学习过程,典型的算法包括 Indyk<sup>[2]</sup>等提出的局部敏感哈希(Local Sensitive Hashing, LSH), Broder<sup>[16]</sup>等提出的 Min-wise Hashing (MinHash) 以及 Raginsky<sup>[17-19]</sup>等提出的改进算法.由于忽略了数据的内部结构和使用随机投影对特征空间进行划分,致使这类哈希算法需要相对较长的编码来提升准确率<sup>[20]</sup>.第二类哈希算法从原始数据中学习投影矩阵,能够保证编码位数较少情况下的辨别力,节省了大量空间存储.因此,投影矩阵的学习是第二类哈希算法的关键所在,受到了广泛的关注.例如, Weiss<sup>[3]</sup>等提出的谱哈希(Spectral Hashing, SH),通过对数据进行图谱分割的方式学习投影矩阵; Kulis & Darrell<sup>[4]</sup>提出的二进制重建嵌入哈希(Binary Reconstruction Embedding, BRE),通过最小化原始距离和海明距离之间的重建误差,建立目标函数; Xu<sup>[9]</sup>等提出的互补哈希(Complementary Hashing, CH),采用建立多个互补哈希表的形式,降低单表误差.

从本质上来说,哈希是将原始空间中的样本映射到低维空间,并尽量保证在原始空间中近邻的样本投影到低维空间后依然维持近邻关系.主成份分析哈希(PCA Hashing, PCAH),将原始高维数据先进行 PCA 降维,然后量化编码.由于缺少约束条件,量化误差较大.针对这个问题,近年来提出了以下几种改进方法,包括 Gong<sup>[12,13]</sup>等提出的迭代量化哈希(Iterative Quantization, ITQ)、Xia<sup>[22]</sup>等提出的稀疏投影哈希(Sparse Projection, SP)以及 Kong & Li<sup>[21]</sup>提出的各向同性哈希(Isotropic Hash, IsoHash). ITQ 的基本原理是:先对数据进行 PCA 降维,然后在此基础上寻找一个正交旋转矩阵来减小空间映射时的量化误差.但该算法中未知量数目远远大于已知量,导致单一的正交约不能完全消除解的不确定性. SP 算法对 ITQ 进一步改进,增加稀疏性约束.但在编码位数相对较少时,稀疏约束则失去作用. PCA 降维得到的投影矩阵,其各个维度间携带信息量存在差异,拥有较大方差的维度携带更多的信息量.大部分哈希算法忽略了这一事实,并使用相同位数对各维度进行编码.针对这个问题, IsoHash 采用最小化投影矩阵维度间方差的差异性策略,实现各向同性.但由于该算法没有对各向同性矩阵做进一步的旋转优化,也没有最小化哈希编码和原始特征向量间的误差,所以检索效果的提升并不是很明显.

针对 ITQ 和 IsoHash 算法的问题,本文提出了各向同性迭代量化哈希(Isotropic Iterative Quantization Hashing, IITQ),通过学习一个各向同性的正交旋转矩阵来调整 PCA 得到的初始投影矩阵,使得量化误差和各个

维度间方差的差异最小,从而得到具有更好辨别力的哈希编码.算法中目标函数是非凸函数,常见的优化工具箱难以解决,本文基于 Attouch<sup>[23]</sup>等提出的逼近算法,设计出一种可行的迭代算法,有效地解决了迭代量化联合各向同性约束引起的非凸问题.在基准图像检索数据库上同多种算法进行了对比实验,结果表明本文算法在查准率、查全率以及平均准确率均值等评价指标方面都有较大提升,验证了本文算法的有效性.

## 2 相关工作

本文算法是建立在 PCA hashing, ITQ 和 IsoHash 三种算法基础之上,所以本章节先对哈希的基本术语和这三种算法进行简明的介绍.

$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$  是实值矩阵,  $n$  是样本数目,  $d$  是样本特征维度.不失一般性,假设数据已做过零中心化预处理,即  $\sum_{i=1}^n \mathbf{x}_i = 0$ . 哈希的目的就是学习  $k$  个哈希函数  $\{h_i(\cdot)\}_{i=1}^k$ , 将  $\forall \mathbf{x}_i \in \mathbf{X}$  映射成长度为  $k$  的二值编码  $b$ , 即  $\mathbb{R}^d \rightarrow \{-1, 1\}^k$ , 如式(1):

$$\mathbf{b} = [h_1(\mathbf{x}), h_2(\mathbf{x}), \dots, h_k(\mathbf{x})] \quad (1)$$

不难发现,对给定的样本矩阵直接求解最优的哈希函数是一个 NP-hard 问题<sup>[24]</sup>. 因此,通常采用先投影后量化的策略.

### 2.1 PCA Hashing

典型的 PCA Hashing 先对输入的  $\mathbf{X}$  降维,取协方差矩阵  $\mathbf{X}\mathbf{X}^T$  的前  $k$  特征向量  $\{\mathbf{w}_i\}_{i=1}^k$ , 组成投影矩阵  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k] \in \mathbb{R}^{d \times k}$ , 哈希函数、量化函数分别采用式(2)、式(3)中形式:

$$f_i(\mathbf{x}) = \mathbf{w}_i^T \mathbf{x}, i = 1, \dots, k \quad (2)$$

$$h_i(\mathbf{x}) = \text{sgn}(f_i(\mathbf{x})) = \begin{cases} 1, & f_i(\mathbf{x}) \geq 0 \\ -1, & \text{otherwise} \end{cases} \quad (3)$$

在投影的基础上,根据其符号进行量化得到样本  $\mathbf{x}$  的哈希编码:

$$\begin{aligned} \mathbf{b} &= [\text{sgn}(\mathbf{w}_1^T \mathbf{x}), \text{sgn}(\mathbf{w}_2^T \mathbf{x}), \dots, \text{sgn}(\mathbf{w}_k^T \mathbf{x})] \\ &\triangleq \text{sgn}(\mathbf{W}^T \mathbf{x}) \end{aligned} \quad (4)$$

### 2.2 Iterative Quantization

ITQ 算法在 PCA Hashing 基础上寻求一个正交旋转矩阵,目标函数如式(5)所示:

$$\begin{aligned} \min_{\mathbf{B}, \mathbf{Q}} \|\mathbf{B} - \mathbf{Q}^T \mathbf{V}\|_F^2 \\ \text{s. t. } \mathbf{Q}^T \mathbf{Q} = \mathbf{I}, \mathbf{B} \in \{-1, 1\}^{k \times n} \end{aligned} \quad (5)$$

其中,  $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n] \in \{-1, 1\}^{k \times n}$  是哈希编码,  $\mathbf{b}_{i \in (1, 2, \dots, n)}$  是  $\mathbf{X}$  中第  $i$  个样本的第  $k$  位二值编码,  $\mathbf{V} = \mathbf{W}^T \mathbf{X}$  是经过 PCA 投影后的矩阵,  $\mathbf{Q} \in \mathbb{R}^{k \times k}$  是待求的正交矩阵.

### 2.3 Isotropic Hashing

使用相同编码位数对不同维度进行编码缺乏合理

性. 针对这一问题, IsoHash 期望最小化各维度间方差的差异. 假设  $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$  是协方差矩阵  $\mathbf{X}\mathbf{X}^T$  特征值分解得到的前  $k$  个特征值,  $\mathbf{a}$  是  $k$  维向量, 每个元素都相同, 都等于特征值之和的均值, 即  $a_i = \frac{1}{k} * \left( \sum_{j=1}^k \lambda_j \right)$ . 算法中期望的各向同性问题最终转化为求解下式中的旋转矩阵  $\mathbf{Q}$ :

$$\begin{aligned} \min_{\mathbf{Q} \in \mathbf{R}^{n \times k}} & \|\text{diag}(\mathbf{Q}^T \mathbf{A} \mathbf{Q}) - \text{diag}(\mathbf{a})\|_F^2 \\ \text{s. t.} & \mathbf{Q}^T \mathbf{Q} = \mathbf{I} \end{aligned} \quad (6)$$

其中,  $\mathbf{A} = \mathbf{W}^T \mathbf{X} \mathbf{X}^T \mathbf{W}$ .

### 3 各向同性迭代量化哈希

IITQ 算法不能消除解的不确定性; IsoHash 算法对初始投影矩阵没有进一步约束, 同时没有最小化哈希编码和原始特征向量间的量化误差. 针对上述问题, 本文将编码更新和各向同性约束统一到目标函数(7)中, 提出了各向同性迭代量化哈希算法 (Isotropic Iterative Quantization Hashing, IITQ).

$$\begin{aligned} \min_{\mathbf{B}, \mathbf{Q}} & \frac{1}{2} \|\mathbf{B} - \mathbf{Q}^T \mathbf{V}\|_F^2 + \frac{\alpha}{2} \|\text{diag}(\mathbf{Q}^T \mathbf{A} \mathbf{Q}) - \text{diag}(\mathbf{a})\|_F^2 \\ \text{s. t.} & \mathbf{Q}^T \mathbf{Q} = \mathbf{I}, \mathbf{B} \in \{-1, 1\}^{k \times n} \end{aligned} \quad (7)$$

其中, 参数  $\alpha \geq 0$  是平衡量化损失和各向同性之间的权重参数,  $\mathbf{B}$  是哈希编码,  $\mathbf{Q}$  是待求的正交旋转矩阵. 显然, 目标函数是非凸的, 直接求解极其复杂, 利用 *At-touch*<sup>[23]</sup> 等提出的逼近算法, 令:

$$\beta(\mathbf{Q}) = \text{diag}(\mathbf{Q}^T \mathbf{A} \mathbf{Q}) - \text{diag}(\mathbf{a}) \quad (8)$$

$$g(\mathbf{Q}) = \frac{1}{2} \alpha \|\beta(\mathbf{Q})\|_F^2 \quad (9)$$

通过求解式(10)和式(11)来更新式(7)中的哈希编码  $\mathbf{B}$  及正交旋转矩阵  $\mathbf{Q}$ .

$$\mathbf{B}_{k+1} = \arg \min_{\mathbf{B} \in \{-1, 1\}^{k \times n}} \frac{1}{2} \|\mathbf{B} - \mathbf{Q}_k^T \mathbf{V}\|_F^2 \quad (10)$$

$$\begin{aligned} \mathbf{Q}_{k+1} = \arg \min_{\mathbf{Q} \in \mathbf{R}^{n \times k}} & \frac{1}{2} \|\mathbf{B}_{k+1} - \mathbf{Q}^T \mathbf{V}\|_F^2 \\ & + \langle \mathbf{Q} - \mathbf{Q}_k, \Delta_{\mathbf{Q}} g(\mathbf{Q}_k) \rangle + \frac{\rho_k}{2} \|\mathbf{Q} - \mathbf{Q}_k\|_F^2 \end{aligned} \quad (11)$$

其中,  $\mathbf{B}_k$  和  $\mathbf{Q}_k$  是第  $k$  次迭代的估计值,  $\rho_k > 0$  是惩罚项,  $\Delta_{\mathbf{Q}} g(\mathbf{Q}_k)$  是函数  $g(\mathbf{Q})$  在  $\mathbf{Q}_k$  处的导数. 根据 *Chu*<sup>[25]</sup> 提出的利用对角元素和特征值构建埃尔米特矩阵 (Hermitian Matrix) 方法,  $\Delta_{\mathbf{Q}} g(\mathbf{Q}_k)$  和  $\rho_k$  可以转化成式(12)、(13):

$$\Delta_{\mathbf{Q}} g(\mathbf{Q}_k) = \alpha \mathbf{A} \mathbf{Q}_k \beta(\mathbf{Q}_k) \quad (12)$$

$$\rho_k = k \alpha \|\mathbf{A}\| \quad (13)$$

其中,  $k > 1$  是一个参数, 本文中设置为 10,  $\|\cdot\|$  表示矩阵的算子范数, 即最大的奇异值. 通过上述转换, 式(10)、(11)中的优化问题都有了相应的闭式解. 其中,

$\mathbf{B}_{k+1}$  的解如式(14):

$$\mathbf{B}_{k+1} = \text{sgn}(\mathbf{Q}_k^T \mathbf{V}) \quad (14)$$

求解  $\mathbf{Q}_{k+1}$  还需要在式(11)基础上做进一步变换:

$$\mathbf{Q}_{k+1} = \arg \max_{\mathbf{Q} \in \mathbf{R}^{n \times k}} \langle \mathbf{Q}, \mathbf{V} \mathbf{B}_{k+1}^T + \rho_k \mathbf{Q}_k - \Delta_{\mathbf{Q}} g(\mathbf{Q}_k) \rangle \quad (15)$$

对上式中的  $\mathbf{V} \mathbf{B}_{k+1}^T + \rho_k \mathbf{Q}_k - \Delta_{\mathbf{Q}} g(\mathbf{Q}_k)$  进行奇异值分解, 若用  $\mathbf{U}_k \sum_k \mathbf{P}_k^T$  表示分解结果, 根据 *Gower*<sup>[26]</sup> 提出的谱问题, 则  $\mathbf{Q}_{k+1}$  的闭式解为:

$$\mathbf{Q}_{k+1} = \mathbf{U}_k \mathbf{P}_k^T \quad (16)$$

最终得到的  $\{(\mathbf{B}_k, \mathbf{Q}_k)\}$  序列能够保证目标函数收敛, 文献[23]对其收敛性有严格的证明. IITQ 算法的计算复杂度如式(17),  $d$  为数据维度,  $n$  为样本数,  $k$  为编码长度, 其算法流程如算法 1.

$$O(d^{2 \times n}) + O(d^3) + O(k^{2 \times n}) + O(k^3) \quad (17)$$

为了直观地表述 IITQ 算法, 采用 toy 示例进行解释. 假定立方体的顶点采用二进制编码, 哈希就是把每一个数据点量化成与它距离最近的立方体的顶点所表示的编码<sup>[12,13]</sup>, 本文用二维平面代替立方体, 对数据点的映射就转变为根据其所在象限位置进行量化. 图 1 中的数据点可量化成  $[\pm 1 \pm 1]$  四种编码, 其中横轴和纵轴 (红色交叉实线) 表示数据的 PCA 方向. 图 1(a) 中采用原始 PCA 对齐方式, 横轴将点簇分割成上下两部分. 以右边点簇为例, 同一点簇被量化成  $(+1, -1)$  和  $(+1, +1)$  两部分, 所以量化误差约为 0.5, 最终整体的误差大约为 1. 图 1(b) 对投影矩阵进行了旋转优化, 只有少部分属于同一簇的点被分隔开, 所以旋转后的量化误差远小于旋转前的 PCA 对齐方式. 图 2(a) 中维度 2 比维度 1 明显携带更多的信息, 使用相同位数对维度 1、2 进行编码, 缺乏合理性. 图 2(b) 采取旋转策略, 使其满足各向同性. 将数据点向两个维度上量化, 误差明显降低. 本文提出的算法就是结合优化旋转和各向同性这两种策略来降低误差.

#### 算法 1 IITQ 算法流程

输入: 训练集  $\mathbf{X}$ , 编码长度  $k$

输出: 哈希编码  $\mathbf{Y}$

步骤 1: 初始化

1. 利用 PCA 初始化投影矩阵  $\mathbf{W}$
2. 初始化  $\mathbf{Q}$  为单位阵

步骤 2: 迭代优化

Repeat

1. 通过式(14), 更新编码矩阵  $\mathbf{B}$
2. 通过式(15), 更新旋转矩阵  $\mathbf{Q}$
3. 判断收敛性, 若收敛, 跳出循环, 返回旋转矩阵  $\mathbf{Q}$

步骤 3: 量化编码

$$\mathbf{Y} = \text{sgn}(\mathbf{Q}^T \mathbf{W}^T \mathbf{X})$$

结束

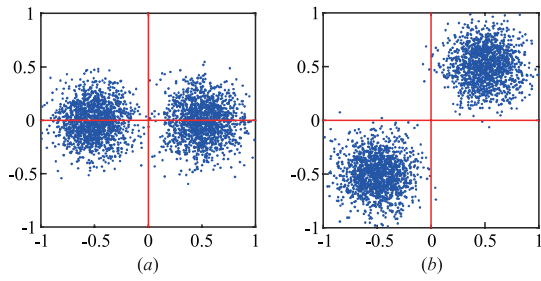


图1 旋转前后量化误差

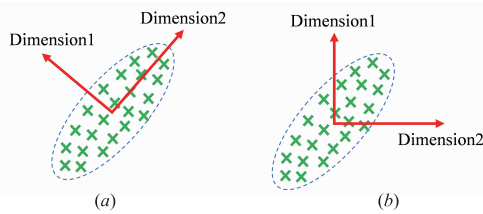


图2 各向异性、同性量化误差

### 4 实验

为验证本文提出算法的有效性,选取 8 种目前最具代表性的哈希算法:ITQ<sup>[12,13]</sup>, IsoHash<sup>[21]</sup>, SP<sup>[22]</sup>, PCAH, LSH<sup>[2]</sup>, SH<sup>[3]</sup>, SPH<sup>[14]</sup> 和 USPLH<sup>[27]</sup> 与 IITQ 算法进行比较,其参数设置与原文献保持一致.

#### 4.1 数据集与评价标准

本文选取大规模图像检索中使用最广泛的 3 个数据集作为基础数据集:

CIFAR-10<sup>[28]</sup>:60000 张 32 × 32 大小的 RGB 图像,数据库被标记成 10 类,每类 6000 张. 实验中随机选取 59000 张作为训练样本,剩下作为测试样本,每幅图像用 512 维 GIST 特征表示.

22K LabelMe<sup>[29]</sup>:由 22019 幅图像组成,实验中选取 20000 张作为训练,剩下的作为测试,每幅图像同样用 512 维 GIST 特征表示.

ANN\_GIST\_1M<sup>[30]</sup>:样本均采用 960 维的 GIST 特征描述子. 实验中选取 100K 样本作训练,1K 样本作测试.

针对样本有无类别标签,本文对查询样本的 Ground truth 做出如下两种定义.

①有标签:用类别标签定义查询样本的真正近邻,即拥有相同类别标签的认定为近邻样本<sup>[20]</sup>.

②无标签:在原始空间中,与查询样本欧式距离最近的前  $k\%$  样本认定为近邻样本<sup>[9,27]</sup>. 本文中  $k$  的值在 3 个数据库中分别设定为 0.02、0.015 和 0.02.

针对第一类 Ground truth,文中采用两种策略来评价哈希的性能. ①Hash lookup:检索返回与查询样本距离小于等于  $r$  的样本,计算此时的查准率;②Hamming ranking:计算查询样本与每个样本的海明距离并排序,检索返回指定数目的样本,计算此时的 mAP.

针对第二类 Ground truth,采用查准率-查全率曲线 (precision-recall curves, P-R) 以及查全率随检索样本变化曲线 (recall-returned samples curves, R-N) 来衡量算法的好坏.

#### 4.2 结果分析

图 3 是哈希算法在不同编码位下的查准-查全对比图. P-R 曲线不仅能反映各种哈希算法的查准率随查全率变化的情况,同时也能反映各种算法的 mAP,曲线位置越高,mAP 越大. 从图 3 中可以看出,各向同性迭代量化哈希的 P-R 曲线位置最高,表明 mAP 值最大,意味着 IITQ 算法检索性能最优.

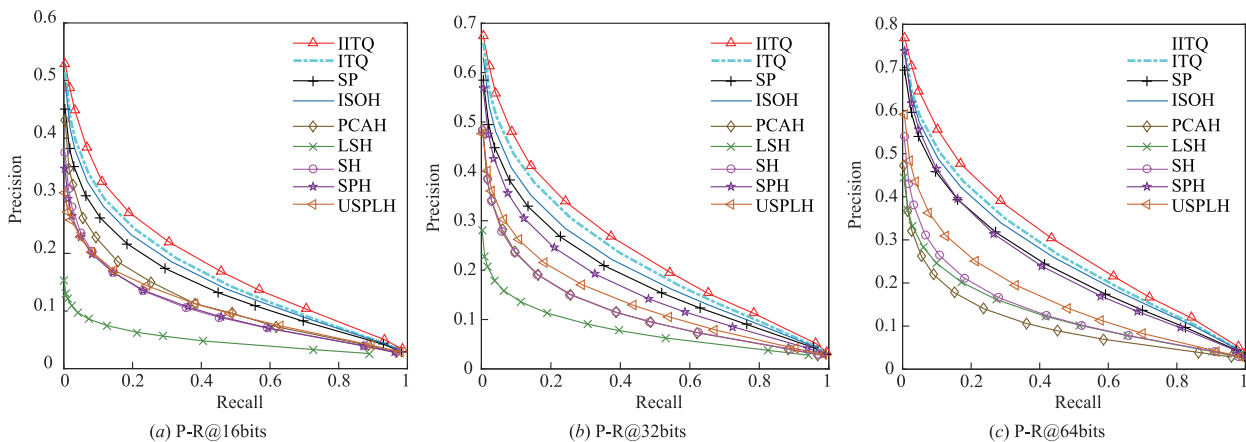


图3 CIFAR-10不同编码位下查准率-查全率对比

图 4 反映的是随着检索样本数目发生变化时查全率的变化情况,IITQ 仍具有最高的查全率. 图 5 是在 LabelMe 数据集上的对比实验. 本文算法在各项评价的指标上均优于其他算法.

表 1 是本文算法与最相关的 4 种哈希方法的平均准确率均值 (mAP) 在 CIFAR-10 数据库上对比的实验结果. 表 2 为采用 hash lookup 评价策略的结果. 实验中查询半径和编码位数分别设置为 2 和 16. 结果表明本

文算法的平均准确率均值和查准率都要优于其他四种 算法,这与图 3、图 4 中得出的结论保持一致。

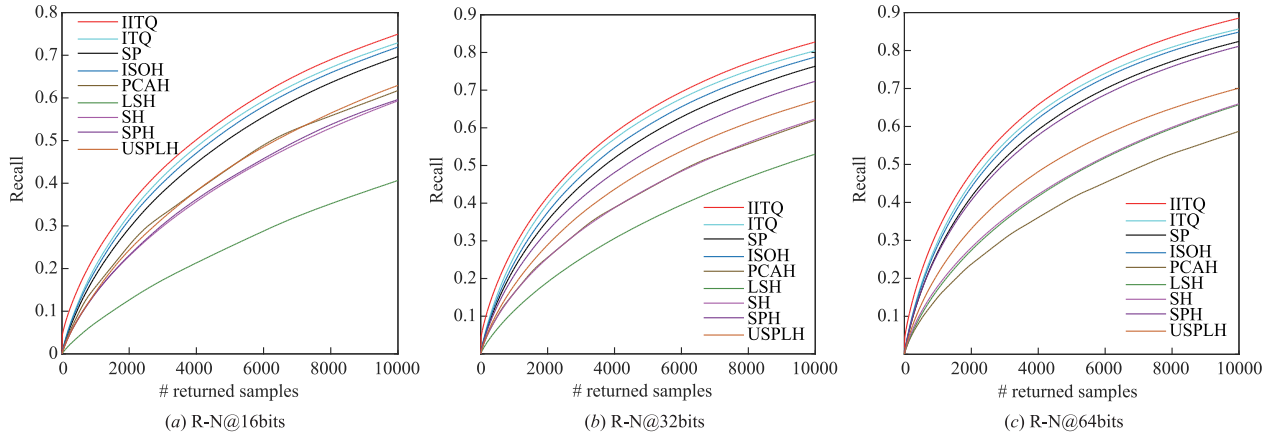


图4 CIFAR-10不同编码位下查全率-检索样本数目对比

表 1 CIFAR-10 海明距离为 2 时的 mAP

位数 \ 算法	16 位	32 位	64 位	96 位
IITQ	0.1625	0.1720	0.1801	0.1821
ITQ	0.1612	0.1692	0.1784	0.1796
SP	0.1568	0.1565	0.1598	0.1622
ISOH	0.1536	0.1611	0.1685	0.1726
PCAH	0.1395	0.1365	0.1345	0.1323

表 2 CIFAR-10 海明距离为 2、编码位数为 16 时的查准率

算法	IITQ	ITQ	SP	ISOH	PCAH
查准率	0.275	0.250	0.2472	0.2486	0.216

为了模拟实际的应用场景,采用了百万量级库 ANN\_GIST\_1M 进行实验. 通过图 3 ~ 图 6, 不难发现,

在不同数据集、不同编码位数情况下,本文提出的 IITQ 算法均能取得较好的指标,验证了本文算法在大规模数据集上有效性和可用性. 主要原因可以归结为: IITQ 充分的利用各向同性和正交旋转的条件,有效的降低了哈希投影后的量化误差.

最后,本文对各个算法在 CIFAR-10、LabelMe 和 ANN\_GIST\_1M 三个数据集上的训练、编码时间进行了对比. 结果记录为表 3、表 4,表中数据为 5 次实验的平均结果,单位为秒. 从表中不难看出,本文算法的训练时间相对较长,其主要归因于目标函数求解的复杂性. 但是耗时较多的参数训练是离线进行的,对样本编码基本没有影响,所以 IITQ 算法能够在大规模数据库上应用.

表 3 CIFAR-10、LabelMe 训练、编码时间

数据库	CIFAR-10						22K LabelMe					
	16 位		32 位		64 位		16 位		32 位		64 位	
算法	训练	编码	训练	编码	训练	编码	训练	编码	训练	编码	训练	编码
IITQ	74.0821	0.0242	259.4942	0.0293	228.3874	0.0598	30.7102	0.0096	119.2830	0.0127	66.7117	0.0208
ITQ	1.4517	0.0221	2.7320	0.0356	4.9357	0.0519	0.4766	0.0122	0.8384	0.0165	1.6530	0.0247
SP	5.2410	0.0233	6.8550	0.0334	10.0892	0.0546	1.7709	0.0095	2.3034	0.0135	3.4350	0.0198
ISOH	2.6208	0.1248	3.2292	0.1716	3.7284	0.2496	1.1544	0.0624	1.6068	0.0780	1.8876	0.1716
PCAH	0.3517	0.0272	0.3783	0.0306	0.3789	0.0450	0.1270	0.0100	0.1391	0.0133	0.1430	0.0185
LSH	0	0.0336	0	0.0355	0	0.0634	0	0.0101	0	0.0121	0	0.0204
SH	6.2972	0.1292	7.3508	0.3656	9.1792	1.4844	4.0231	0.0431	4.8718	0.1278	6.5564	0.5034
SPH	5.5623	0.2534	8.1594	0.2914	13.9523	0.3662	2.5780	0.0944	4.2309	0.1131	7.1750	0.1409
USPLH	13.4787	0.0231	27.8338	0.0324	56.3461	0.0554	11.2219	0.0110	23.1823	0.0143	46.5492	0.0209

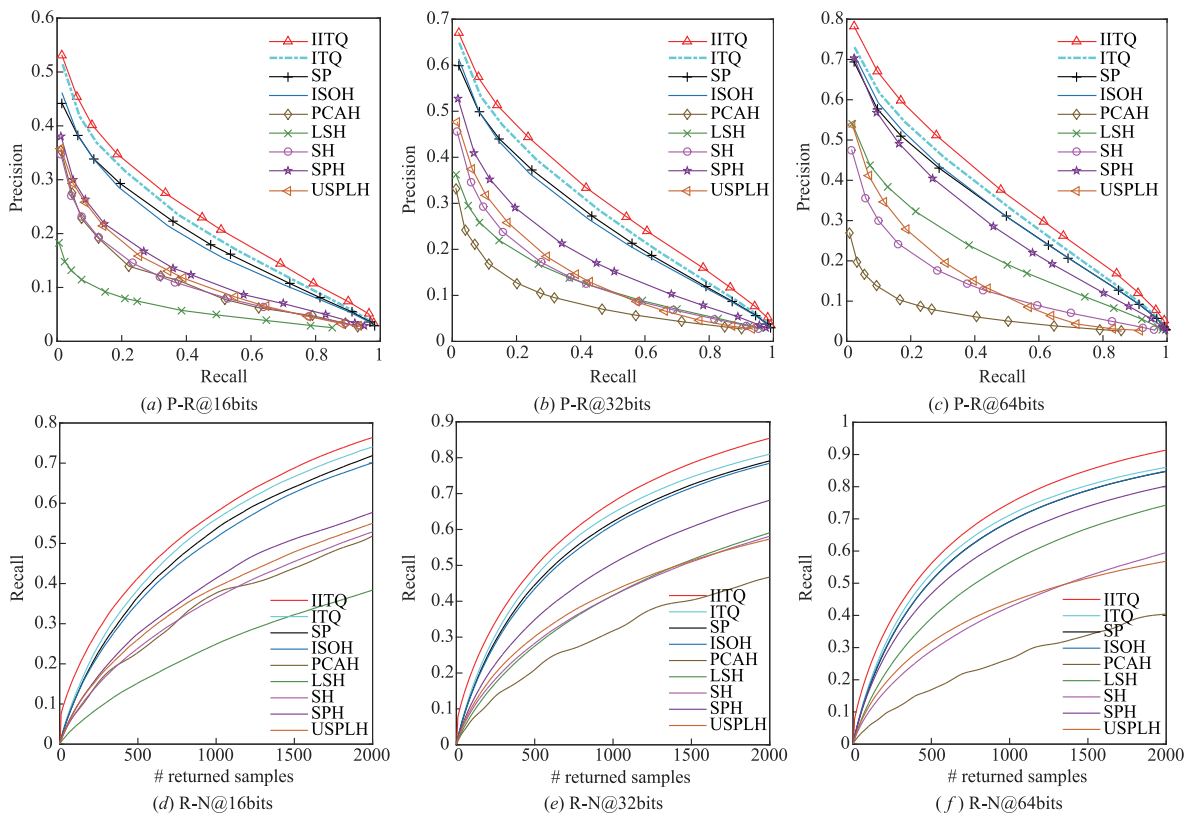


图5 22K LabelMe不同编码位下P-R、R-N对比

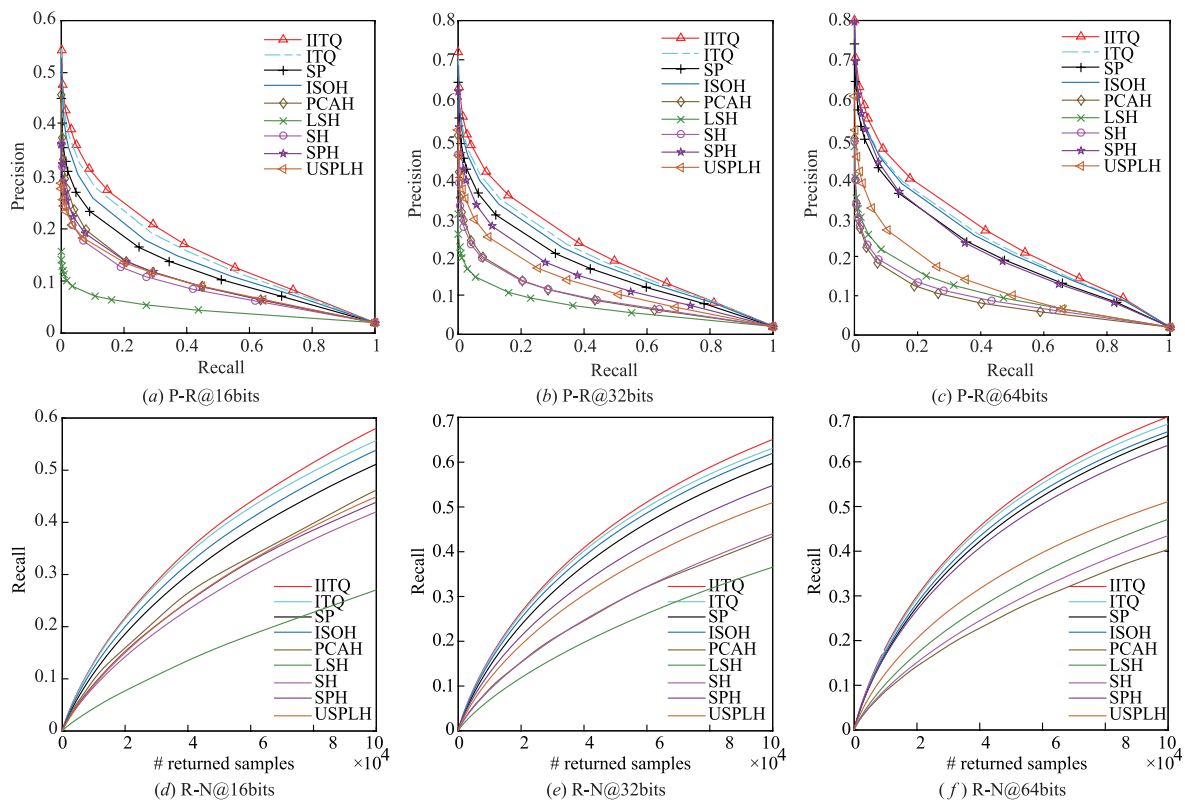


图6 ANN\_GIST\_1M不同编码位下P-R、R-N对比

表 4 ANN\_GIST\_1M 训练、编码时间

数据库	ANN_GIST_1M					
	16 位		32 位		64 位	
编码位	训练	编码	训练	编码	训练	编码
IITQ	4292	0.568	8070	0.695	13298	1.174
ITQ	32.61	0.571	48.60	0.673	84.959	1.219
SP	143.3	0.468	167.7	0.674	229.58	0.972
ISOH	18.92	0.388	17.46	0.671	18.564	1.183
PCAH	8.684	0.505	8.765	0.598	8.4356	0.985
LSH	0.001	0.492	0.001	0.678	0.0023	1.154
SH	9.934	1.501	11.09	3.842	13.226	11.55
SPH	42.09	2.398	75.84	2.703	13.952	3.439
USPLH	132.1	0.519	27.83	0.686	56.346	1.215

## 5 结语

针对迭代量化哈希 (ITQ) 寻找最优旋转矩阵时缺少充分先验, 以及各向同性哈希 (IsoHash) 没有考虑最小化编码和原始特征向量间的量化误差的问题, 提出了一种基于各向同性的迭代量化哈希 (IITQ)。本文算法的主要思想是利用旋转优化结合各向同性约束, 进一步降低投影后每个维度的量化误差。相比其它经典方法, IITQ 具有更加有效的检索性能。在 CIFAR-10、22K LabelMe 和 ANN\_GIST\_1M 基准数据库上的对比实验表明本文算法可以显著提高检索的查准率、查全率以及平均准确率均值。

### 参考文献

- [1] Arya S, Mount D M. Approximate nearest neighbor queries in fixed dimensions [A]. Acm-Siam Symposium on Discrete Algorithms [C]. Society for Industrial and Applied Mathematics, 1996. 271 – 280.
- [2] Gionis A, Indyk P, Motwani R. Similarity search in high dimensions via Hashing [A]. International Conference on Very Large Data Bases [C]. Morgan Kaufmann Publishers Inc, 2000. 518 – 529.
- [3] Weiss Y, Torralba A, Fergus R. Spectral Hashing [A]. Advances in Neural Information Processing Systems [C]. MIT, 2009. 1753 – 1760.
- [4] Kulis B, Darrell T. Learning to Hash with binary reconstructive embeddings [A]. Advances in Neural Information Processing Systems [C]. MIT, 2009. 1042 – 1050.
- [5] Kulis B, Jain P, Grauman K. Fast similarity search for learned metrics [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(12): 2143 – 2157.
- [6] Weiss Y, Fergus R, Torralba A. Multidimensional spectral Hashing [A]. Computer Vision-ECCV 2012 [C]. Berlin, Heidelberg: Springer, 2012. 340 – 353.
- [7] Andoni A, Indyk P. Near-optimal Hashing algorithms for approximate nearest neighbor in high dimensions [A]. 47th Annual IEEE Symposium on Foundations of Computer Science [C]. IEEE, 2006. 459 – 468.
- [8] Norouzi M, Blei D M. Minimal loss Hashing for compact binary codes [A]. Proceedings of the 28th International Conference on Machine Learning [C]. ACM, 2011. 353 – 360.
- [9] Xu H, Wang J, Li Z, et al. Complementary Hashing for approximate nearest neighbor search [A]. 2011 International Conference on Computer Vision [C]. IEEE, 2011. 1631 – 1638.
- [10] Chang S F. Supervised Hashing with kernels [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. IEEE, 2012. 2074 – 2081.
- [11] Liu W, Wang J, Kumar S, et al. Hashing with Graphs [A]. Proceedings of the 28th International Conference on Machine Learning [C]. ACM, 2011. 1 – 8.
- [12] Gong Y, Lazebnik S. Iterative quantization: A procrustean approach to learning binary codes [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. IEEE, 2011. 817 – 824.
- [13] Gong Y, Lazebnik S, Gordo A, et al. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(12): 2916 – 2929.
- [14] Heo J P, Lee Y, He J, et al. Spherical Hashing [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. IEEE, 2012. 2957 – 2964.
- [15] He K, Wen F, Sun J. K-means Hashing: An affinity-preserving quantization method for learning binary compact codes [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. IEEE, 2013. 2938 – 2945.
- [16] Broder A Z, Charikar M, Frieze A M, et al. Min-wise independent permutations [J]. Journal of Computer and System Sciences, 2000, 60(3): 630 – 659.
- [17] Raginsky M, Lazebnik S. Locality-sensitive binary codes from shift-invariant kernels [C]. Advances in Neural Information Processing Systems [C]. MIT, 2009. 1509 – 1517.
- [18] Datar M, Immorlica N, Indyk P, et al. Locality-sensitive Hashing scheme based on P-stable distributions [A]. Proceedings of the Twentieth Annual Symposium on Computational Geometry [C]. ACM, 2004. 253 – 262.
- [19] Kulis B, Grauman K. Kernelized locality-sensitive Hashing for scalable image search [A]. IEEE 12th International Conference on Computer Vision [C]. IEEE, 2009. 2130 – 2137.
- [20] Shen F, Shen C, Liu W, et al. Supervised discrete Hashing

- [A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. IEEE, 2015. 37 – 45.
- [21] Kong W, Li W J. Isotropic Hashing[A]. Advances in Neural Information Processing Systems[C]. MIT, 2012. 1646 – 1654.
- [22] Xia Y, He K, Kohli P, et al. Sparse projections for high-dimensional binary codes[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. IEEE, 2015. 3332 – 3339.
- [23] Attouch H, Bolte J. On the convergence of the proximal algorithm for non-smooth functions involving analytic features[J]. Mathematical Programming, 2009, 116(1 – 2): 5 – 16.
- [24] Hochba D S. Approximation algorithms for NP-Hard problems[J]. ACM SIGACT News, 1997, 28(2): 40 – 52.
- [25] Chu M T. Constructing a Hermitian matrix from its diagonal entries and eigenvalues[J]. Siam Journal on Matrix Analysis & Applications, 1995, 16(1): 207 – 217.
- [26] Gower J C, Dijksterhuis G B. Procrustes problems[J]. Oxford Statistical Science, 2004, 168(2): 233.
- [27] Wang J, Kumar S, Chang S F. Sequential projection learning for Hashing with compact codes[A]. Proceedings of 27th International Conference on Machine Learning[C]. ACM, 2010. 1127 – 1134.
- [28] Krizhevsky A, Hinton G. Learning Multiple Layers of Features from Tiny Images[R]. University of Toronto, 2009.
- [29] Torralba A, Fergus R, Weiss Y. Small codes and large image databases for recognition[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. IEEE, 2008. 1 – 8.
- [30] Yuan Y, Lu X, Li X. Learning Hash functions using sparse reconstruction[A]. Proceedings of International Conference on Internet Multimedia Computing and Service[C]. ACM, 2014. 14.

#### 作者简介



**李来** 男, 1990年11月出生, 江苏徐州人, 南京信息工程大学硕士研究生在读, 研究方向为模式识别与图像处理.

E-mail: laili@nuist.edu.cn



**刘光灿** 男, 1982年5月出生, 湖南邵阳人, 南京信息工程大学教授、硕士生导师, 研究方向为模式识别、图像处理、数据挖掘.

E-mail: gcliu@nuist.edu.cn

**孙玉宝** 男, 1983年10月出生, 江苏连云港人, 南京信息工程大学副教授、硕士生导师, 研究方向为高光谱图像处理.

E-mail: sunyub@nuist.edu.cn

**刘青山** 男, 1975年4月出生, 安徽合肥人, 南京信息工程大学教授、博士生导师, 研究方向为图像与视频分析、机器学习.

E-mail: qslu@nuist.edu.cn