Label Self-Adaption Hashing for Image Retrieval

Jianglin Lu[a,b,c], Zhihui Lai[a,b], Jingxu Lin[a], Qinghong Lin[a,c], Jie Zhou[a,b]

[a]*Shenzhen University, Shenzhen, China*

[b]*Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen, China*

[c]*Kazan Federal University, Kazan, Russia*

Motivation

We attempt to **learn a self-adaptive cluster label matrix from the data** under the assumption that the nearest neighbor points should have a larger probability to be in the same cluster.

The hash codes are regressed to the learned cluster labels to make full use of **potential discriminative information** of data.

Based on anchor graph, the learning of hashing function **is conducted on the $n \times m$ similarity matrix** constructed by n sample points and m anchors. Therefore, it can perform efficient binary code learning for large-scale image retrieval.

Contributions

We propose a novel unsupervised hashing method called LSAH, which can **learn a self-adaptive cluster label matrix** to make full use of the **potential discriminative information of data** to guide the learning of binary codes.

An iterative algorithm based on **Augmented Lagrange Multiplier (ALM)** is elaborately designed to solve the resulting optimization problem.

Extensive experiments on three well-known data sets indicate the promising performance of the proposed LSAH.

Overall Framework

The proposed LSAH contains the following two parts, including the **effective hashing function learning** part and the **self-adaption label generation** part.

Effective Hashing Function Learning

$$\min_{F,B} \sum_{i}^{n} \sum_{j}^{m} ||\tilde{F}(x_i) - b_j||_2^2 S_{ij} + \lambda \mathcal{R}(\tilde{F})$$

$$s.t. \;\; B \in \{-1,1\}^{m \times r}$$

Self-Adaption Label Generation

$$\min_{B,K,M} \alpha\Omega(K) + \beta||K - BM^T||_F^2$$

$$s.t. \;\; B \in \{-1,1\}^{m \times r}, \; K^T K = I_c, \; K \geq 0, \; M^T M = I_c$$

Optimization

Algorithm 1: Optimization algorithm of LSAH

Input: Training set $X \in \Re^{n \times d}$, the number of classes c, bits r, iterations $maxT$, anchors m, nearest neighbor points k, the balance parameters λ, α, β, $\mu > 0$, $\rho > 1$.

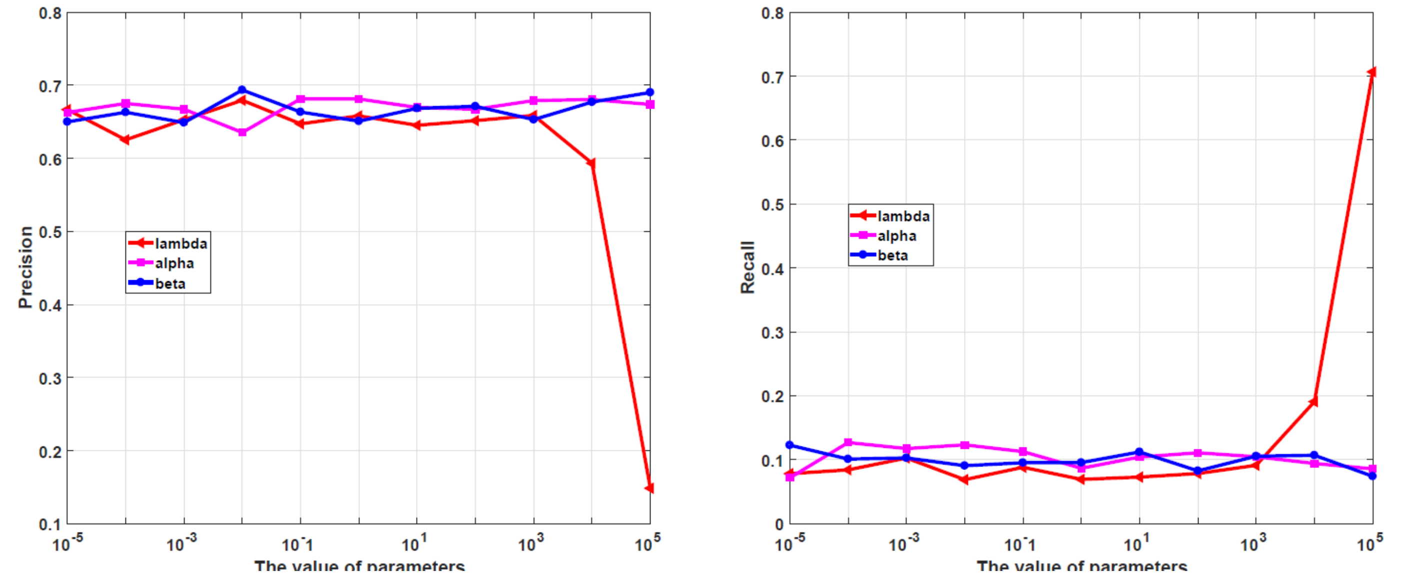
Output: Projection matrix P, regression matrix M, and label matrix K.

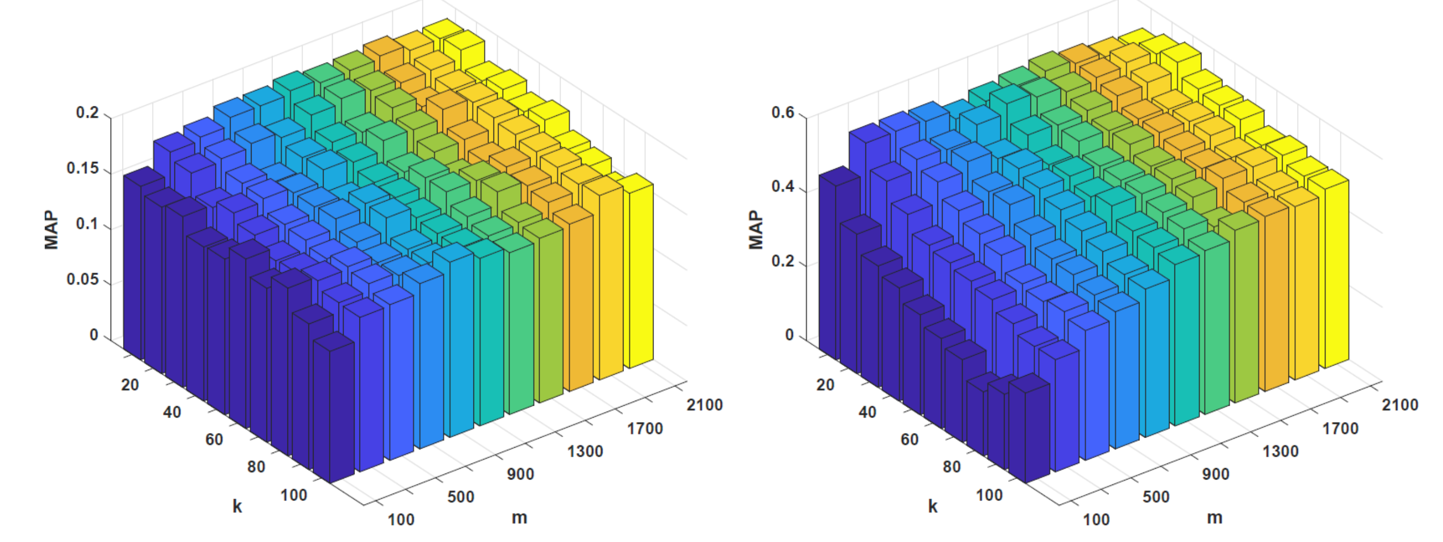
Generate m anchor points by using K-Means clustering method.
Construct anchor matrix Z and Laplacian matrix L.
Initialize B, P randomly.
Initialize U and K as two identity matrices, Θ with each element equal to 1.
for $iter = 1 : maxT$ **do**
  Update M: $M = \Gamma_2 \Psi_2^T$
  Update P: $P = (X^T X + \lambda U)^{-1} X^T S B$
  Update K: $K = \Gamma_1 \Psi_1^T$
  Update R: $r_{ij} = \max(0, \hat{r}_{ij})$
  Update B: $B = sign(S^T X P + \beta K M)$
  for $i = 1 : d$ **do**
    $U_{ii} = 1/(2||P_i||_2)$
  end
  $\Theta = \Theta + \mu(K - R)$
  $\mu = \min(10^6, \rho\mu)$.
end

MAP RESULTS

Dataset	CIFAR-10					MNIST				
Code Length (bits)	16	32	64	96	128	16	32	64	96	128
LSH	0.1210	0.1357	0.1437	0.1521	0.1568	0.2102	0.2620	0.3155	0.3701	0.3768
SH	0.1266	0.1252	0.1254	0.1250	0.1249	0.2661	0.2606	0.2400	0.2367	0.2330
AGH	0.1623	0.1624	0.1466	0.1406	0.1381	0.5321	0.4363	0.3659	0.3254	0.2993
ITQ	0.1526	0.1634	0.1669	0.1680	0.1705	0.3378	0.3505	0.3895	0.4008	0.4090
IMH	0.1736	0.1742	0.1775	0.1926	0.1866	0.5016	0.4581	0.3982	0.3774	0.3453
SP	0.1593	0.1653	0.1733	0.1760	0.1774	0.3931	0.4166	0.4376	0.4513	0.4610
DPLM	0.1660	0.1901	0.1928	**0.2000**	0.1996	0.3673	0.4448	**0.5417**	0.5338	0.5419
SHSR	**0.1763**	0.1675	0.1648	0.1634	0.1617	0.5379	0.4527	0.3957	0.3726	0.3603
LSAH	0.1751	**0.1923**	**0.1987**	0.1983	**0.2005**	**0.5423**	**0.5423**	0.5353	**0.5649**	**0.5520**

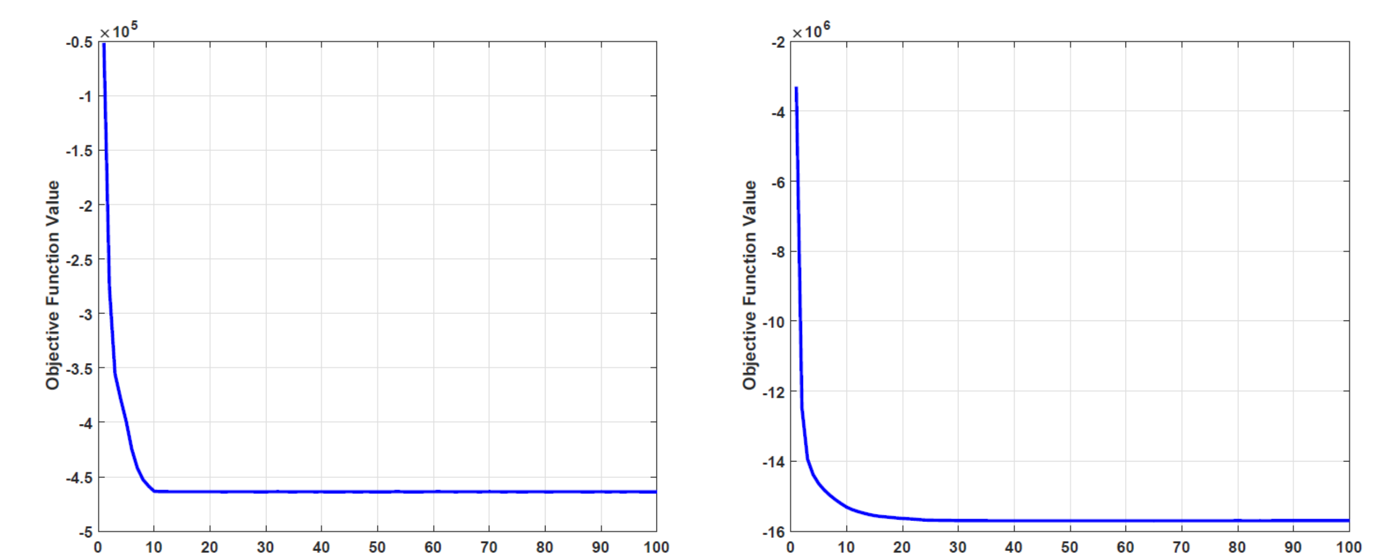
Parameter Sensitivity





(a) (b)

Convergence



(a) (b)