

Jointly Multiple Hash Learning

Xingbo Liu,¹ Xiushan Nie,^{2,*} Yingxin Wang,¹ Yilong Yin^{3,*}

¹School of Computer Science and Technology, Shandong University, Jinan, P.R. China

²School of Computer Science and Technology, Shandong Jianzhu University, Jinan, P.R. China

³School of Software, Shandong University, Jinan, P.R. China

sclxb@mail.sdu.edu.cn, wyxwalker@163.com, niexiushan@163.com, ylyin@sdu.edu.cn

Motivation

Hashing can compress heterogeneous high-dimensional data into compact binary codes while preserving the similarity to facilitate efficient retrieval and storage, and thus hashing has recently received much attention from information retrieval researchers. Most of the existing hashing methods first predefine a fixed length (e.g., 32, 64, or 128 bit) for the hash codes, before learning them with this fixed length. When the length of the hash code changes, these methods need to be retrained and learn the hash codes again. However, we know that hash codes are compact representations of the original samples, and one sample can be represented by various hash codes with different lengths. Intuitively, hash codes with different lengths that represent the same sample would reflect different types of specific information from the original sample. If they are treated as different views of the original sample, some relationships and associations should exist among different views that could improve the performance of hash learning according to multi-viewing theory (Xu, Tao, and Xu 2013). Therefore, harnessing these relationships will boost the performance of hashing methods.

Inspired by this possibility, in this study, we propose a new model called jointly multiple hash learning (JMH), which can learn hash codes with multiple lengths simultaneously. In the proposed JMH method, three types of information are used for hash learning, which come from hash codes with different lengths and the original features of the samples and label. In contrast to the existing hashing methods, JMH can learn hash codes with different lengths in one step. To the best of our knowledge, JMH is one of the first attempts to learn multi-length hash codes simultaneously. In addition, JMH can obtain a closed-form and discrete solution by using cyclic coordinate descent without relaxation, which is faster and more precise than the existing methods.

Formulation

It is assumed that we have a training set \mathbf{X} comprising n samples (i.e., $\mathbf{X} = (\mathbf{x}_i)_{i=1}^n$), where $\mathbf{x}_i \in R^d$ is the feature vector. A semantic label matrix $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^n$ is also available, where $\mathbf{y}_i = \{y_{ij}\} \in R^c$ is the label vector of the i_{th} instance

*Xiushan Nie and Yilong Yin are the corresponding authors.
Copyright © 2019, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

and c is the number of categories in the training set. If the i_{th} instance belongs to the j_{th} semantic category, $y_{ij} = 1$; otherwise, -1 . In addition, it is assumed that we want to obtain K different hash codes (vectors) with each length L_k . The hash matrix of all n samples for the k_{th} ($1 \leq k \leq K$) hash codes is defined as $\mathbf{H}_k = \{\mathbf{h}_i^k\}_{i=1}^n$, where \mathbf{h}_i^k is the k_{th} hash code of the i_{th} sample with length L_k . \mathbf{P}_k , \mathbf{U}_k , and \mathbf{W}_k are the matrices that describe the linear projections.

In the proposed JMH, three types of information are used, which come from hash codes with different lengths and the original features of the samples and labels. They are represented by the projections \mathbf{P}_k , \mathbf{U}_k , and \mathbf{W}_k . To learn these three projections and multi-length hash codes, we propose three loss terms comprising Q -Loss, L -Loss, and M -Loss in the objective function of the JMH, where Q -Loss represents the loss between the k_{th} hash code \mathbf{H}_k and the other hash codes, L -Loss represents the loss between the hash codes and original features and M -Loss describes the classification error, respectively.

$$\begin{aligned} Q &= \text{Loss}(\mathbf{H}_k, \mathbf{P}_k^T \mathbf{H}_{k+1}) + \lambda_k \|\mathbf{P}_k\|^2 \\ &= \sum_{k=1}^{K-1} \gamma_k \|\mathbf{H}_k - \mathbf{P}_k^T \mathbf{H}_{k+1}\|^2 + \sum_{k=1}^{K-1} \lambda_k \|\mathbf{P}_k\|^2, \end{aligned} \quad (1)$$

$$L = \text{Loss}(\mathbf{H}, \mathbf{X}) = \sum_{k=1}^K \alpha_k \|\mathbf{H}_k - \mathbf{U}_k^T \mathbf{X}\|^2, \quad (2)$$

$$M = \text{Loss}(\mathbf{Y}, \mathbf{H}) = \sum_{k=1}^K \mu_k \|\mathbf{Y} - \mathbf{W}_{k-1}^T \mathbf{H}_k\|^2, \quad (3)$$

where $\text{Loss}(\cdot)$ is a loss function, and λ_k and γ_k are the parameters.

By combining Equations (1), (2), and (3), the final formulation for the JMH can be defined as:

$$\begin{aligned} &\min_{\mathbf{H}_k, \mathbf{W}_k, \mathbf{P}_k, \mathbf{U}_k} \sum_{k=1}^K \mu_k \|\mathbf{Y} - \mathbf{W}_{k-1}^T \mathbf{H}_k\|^2 \\ &+ \sum_{k=1}^{K-1} \gamma_k \|\mathbf{H}_k - \mathbf{P}_k^T \mathbf{H}_{k+1}\|^2 + \sum_{k=1}^K \alpha_k \|\mathbf{H}_k - \mathbf{U}_k^T \mathbf{X}\|^2 \\ &+ \sum_{k=1}^{K-1} \lambda_k \|\mathbf{P}_k\|^2 + \sum_{k=1}^K \beta_k \|\mathbf{H}_k \mathbf{1}\|^2 \\ \text{s.t. } &\mathbf{H}_k \in \{-1, 1\}^{L_k \times n}, \end{aligned} \quad (4)$$

where μ_k , α_k , λ_k and γ_k are the parameters.

Table 1: Performance in terms of MAP score.

Method	CIFAR-10				MS-COCO				NUS-WIDE			
	12 bits	24 bits	36 bits	48 bits	12 bits	24 bits	36 bits	48 bits	12 bits	24 bits	36 bits	48 bits
SH	0.2704	0.2908	0.2898	0.2961	0.6490	0.6501	0.6616	0.6800	0.5958	0.5986	0.6032	0.6070
PCA-ITQ	0.2767	0.3554	0.3398	0.3562	0.5328	0.6281	0.6661	0.6907	0.3181	0.4050	0.4767	0.6016
PCA-RR	0.2903	0.3054	0.2938	0.3172	0.5590	0.5677	0.6481	0.6594	0.5519	0.5649	0.5794	0.6282
MFH	0.2991	0.3345	0.3473	0.3623	0.6171	0.6330	0.6470	0.6500	0.5820	0.6088	0.6305	0.6315
SDH	0.3473	0.3830	0.4111	0.4984	0.5482	0.6037	0.6489	0.6531	0.4978	0.5775	0.7120	0.7350
COSDISH	0.3820	0.4366	0.4854	0.5295	0.5230	0.5369	0.5390	0.6164	0.3099	0.3135	0.3890	0.4274
FSDH	0.5370	0.6218	0.6526	0.6632	0.7287	0.7780	0.7966	0.8174	0.6808	0.7663	0.7765	0.7778
JMH	0.4576	0.6517	0.6624	0.6695	0.7741	0.8096	0.8032	0.8437	0.7048	0.8110	0.8095	0.8205

Optimizing Equation (4) directly is a challenge because it is both non-convex and non-smooth. However, obtaining a solution becomes easier when we take one variable while another is fixed. In the JMH, Equation (4) can be solved by an iterative framework and cyclic coordinate descent until convergence, and we can directly obtain a discrete solution for the hash codes without relaxation.

Experiments

We used three datasets comprising CIFAR-10, MS-COCO, and NUS-WIDE in the experiments, which have been employed widely in image retrieval studies. To evaluate the proposed method, we used the evaluation metric comprising the mean average precision (MAP), which is used widely in image retrieval.

We compared the proposed JMH with the following methods: Spectral Hashing (SH) (Weiss, Torralba, and Fergus 2009), PCA-Iterative Quantization (PCA-ITQ) (Gong and Lazebnik 2011), PCA-Random Rotation (Gong and Lazebnik 2011), Collective Matrix Factorization Hashing (MFH) (Ding, Guo, and Zhou 2014), Supervised Discrete Hashing (SDH) (Shen et al. 2015), Column Sampling based Discrete Supervised Hashing (COSDISH) (Kang, Li, and Zhou 2016), Fast Supervised Discrete Hashing (FSDH) (Gui et al. 2018),

During training, four lengths of hash codes (i.e., $K=4$) ranging from 12 to 48 bits were employed in comparisons with several state-of-the-art methods. In addition, five types of hyperparameters were used with our proposed JMH method. We empirically set $\{\mu\}_{k=1}^K = \{1, 0.8, 0.7, 0.6\}$, $\{\alpha\}_{k=1}^{K-1}$, $\{\gamma_k\}_{k=2}^K$ to $1e-6$, $\{\beta_k\}_{k=1}^K$ and $\{\lambda_k\}_{k=0}^K$ to 0.1.

The MAP results obtained for all of the methods based on CIFAR-10, MS-COCO and NUS-WIDE using hash codes with different lengths are listed in Table 1. Table 1 shows that the proposed JMH method performed substantially better than the other hashing methods, where the MAP results showed that our method improved the performance by 3–10% in most cases. The results demonstrated that the relationships among multi-length hashes could boost the retrieval performance.

Conclusion

In this study, we propose a novel supervised discrete hashing method for learning multi-length hash codes in a united framework. In the proposed model, every hash code is treated

as a view of the corresponding sample. After training using the cyclic coordinate descent method, discrete multi-length hashes are obtained based on the associations among different hash codes. Experiments conducted based on three benchmark datasets confirmed the superior performance of our method in various retrieval scenarios. One challenge with the proposed method is that the number of parameters increases with the number of hash codes in the framework. We will try to address this challenge in future research.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (61876098, 61671274, 61573219), and the Fostering Project of Dominant Discipline and Talent Team of Shandong Province Higher Education Institutions.

References

- Ding, G.; Guo, Y.; and Zhou, J. 2014. Collective matrix factorization hashing for multimodal data. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2083–2090.
- Gong, Y., and Lazebnik, S. 2011. Iterative quantization: A procrustean approach to learning binary codes. In *IEEE Conference on Computer Vision and Pattern Recognition*, 817–824.
- Gui, J.; Liu, T.; Sun, Z.; Tao, D.; and Tan, T. 2018. Fast supervised discrete hashing. *IEEE transactions on pattern analysis and machine intelligence* 40(2):490–496.
- Kang, W.-C.; Li, W.-J.; and Zhou, Z.-H. 2016. Column sampling based discrete supervised hashing. In *AAAI*, 1230–1236.
- Shen, F.; Shen, C.; Liu, W.; and Tao Shen, H. 2015. Supervised discrete hashing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 37–45.
- Weiss, Y.; Torralba, A.; and Fergus, R. 2009. Spectral hashing. In *Advances in neural information processing systems*, 1753–1760.
- Xu, C.; Tao, D.; and Xu, C. 2013. A survey on multi-view learning. *arXiv preprint arXiv:1304.5634*.