# Deep Position-Aware Hashing for Semantic Continuous Image Retrieval (Appendix)

Ruikui Wang[1,2], Ruiping Wang[1,2], Shishi Qiao[1,2], Shiguang Shan[1,2], Xilin Chen[1,2]
[1] Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing, 100190, China
[2] University of Chinese Academy of Sciences, Beijing, 100049, China
{ruikui.wang, shishi.qiao}@vipl.ict.ac.cn, {wangruiping, sgshan, xlchen}@ict.ac.cn

## A. Detailed derivation from Eq.(1) to Eq.(2)

The Eq.(2) can be derivated as follow:

$$
\begin{aligned}
L^i_{inter} &= max\{d_H(\boldsymbol{b_i}, \bar{\boldsymbol{\theta}}_{\boldsymbol{i}})(1+\alpha) - \min_{j \notin label\{i\}} d_H(\boldsymbol{b_i}, \theta_j),\ \ 0\} \\
&= max\{\max_{j \notin label\{i\}} (-d_H(\boldsymbol{b_i}, \boldsymbol{\theta_j})) + d_H(\boldsymbol{b_i}, \bar{\boldsymbol{\theta}}_{\boldsymbol{i}})(1+\alpha),\ \ 0\} \\
&\leq max\{log \sum_{j \notin label\{i\}} exp\{-d_H(\boldsymbol{b_i}, \boldsymbol{\theta_j})\} + d_H(\boldsymbol{b_i}, \bar{\boldsymbol{\theta}}_{\boldsymbol{i}})(1+\alpha),\ \ 0\} \\
&\leq log\{\sum_{j \notin label\{i\}} exp\{-d_H(\boldsymbol{b_i}, \boldsymbol{\theta_j})\} \cdot\ exp\{d_H(\boldsymbol{b_i}, \bar{\boldsymbol{\theta}}_{\boldsymbol{i}})(1+\alpha)\} + 1\} \\
&= log\{\frac{\sum_{j \notin label\{i\}} exp\{-d_H(\boldsymbol{b_i}, \boldsymbol{\theta_j})\} + exp\{-d_H(\boldsymbol{b_i}, \bar{\boldsymbol{\theta}}_{\boldsymbol{i}})(1+\alpha)}{exp\{-d_H(\boldsymbol{b_i}, \bar{\boldsymbol{\theta}}_{\boldsymbol{i}})(1+\alpha)}\} \\
&= -log\{\frac{exp\{-d_H(\boldsymbol{b_i}, \bar{\boldsymbol{\theta}}_{\boldsymbol{i}})(1+\alpha)}{exp\{-d_H(\boldsymbol{b_i}, \bar{\boldsymbol{\theta}}_{\boldsymbol{i}})(1+\alpha) + \sum_{j \notin label\{i\}} exp\{-d_H(\boldsymbol{b_i}, \boldsymbol{\theta_j})\}}\}
\end{aligned}
$$

## B. Supplementary experiments

| Method | CIFAR-10 | | | | ImageNet | | | |
|---|---|---|---|---|---|---|---|---|
| | 12-bit | 24-bit | 32-bit | 48-bit | 16-bit | 32-bit | 48-bit | 64-bit |
| DSDH[5] | 0.740 | 0.786 | 0.801 | 0.820 | - | - | - | - |
| TALR[2] | 0.732 | 0.789 | 0.800 | 0.826 | 0.5892 | 0.6689 | 0.6985 | 0.7053 |
| HBMP[1] | 0.799 | 0.804 | 0.830 | 0.831 | 0.574 | 0.692 | 0.712 | **0.742** |
| Greedy-Hash[7] | 0.774 | 0.795 | 0.810 | 0.822 | 0.625 | 0.662 | 0.682 | 0.688 |
| DPAH | **0.810** | **0.825** | **0.838** | **0.845** | **0.6517** | **0.7001** | **0.7149** | 0.7138 |

Table 1. Comparison of retrieval performance of our DPAH method and most recently published hashing methods on CIFAR-10 and ImageNet. Please note that DSDH does not report their result on ImageNet.

In order to compare with some most recently published deep hashing methods, we add the supplementary experiments on CIFAR-10 [3] which is also used widely. We follow the same data division protocol as [4, 6, 1, 7]. Specifically, we sample 100 images per category to form the query and the other 59K images to make up the database. Then 500 images per category is selected from the database to construct the training set. We also add the performance comparison of these recent methods with our method on ImageNet. The comparison result is listed as follows:

From Table 1, it can be seen that our DPAH outperforms other methods in most cases, and it is only inferior to HBMP under 64-bit on ImageNet100. Actually, HBMP is an advanced two stage hashing method, due to the matrix eigenvector solution and matrix inversion in its first stage, solving the ideal binary code will be difficult and inefficient when the data size is large. On the contrary, our method conducts binary encoding and hash function learning simultaneously in an end-to-end manner, so it is more suitable for real-world scenarios.

## References


[1] F. Cakir, K. He, and S. Sclaroff. Hashing with binary matrix pursuit. In *Proceedings of the European Conference on Computer Vision*, pages 332–348. Springer, 2018.

[2] K. He, F. Cakir, S. Adel Bargal, and S. Sclaroff. Hashing as tie-aware learning to rank. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4023–4032, 2018.

[3] A. Krizhevsky and G. Hinton. Learning multiple layers of features from tiny images. Technical report, Citeseer, 2009.

[4] H. Lai, Y. Pan, Y. Liu, and S. Yan. Simultaneous feature learning and hash coding with deep neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3270–3278, 2015.

[5] Q. Li, Z. Sun, R. He, and T. Tan. Deep supervised discrete hashing. In *Advances in Neural Information Processing Systems*, pages 2482–2491. 2017.

[6] W. J. Li, S. Wang, and W. C. Kang. Feature learning based deep supervised hashing with pairwise labels. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, 2016.

[7] S. Su, C. Zhang, K. Han, and Y. Tian. Greedy hash: Towards fast optimization for accurate hash coding in cnn. In *Advances in Neural Information Processing Systems*, pages 806–815. 2018.