Binary Hashing with Semidefinite Relaxation and Augmented Lagrangian

Thanh-Toan Do, Anh-Dzung Doan, Duc-Thanh Nguyen, Ngai-Man Cheung

Singapore University of Technology and Design

{thanhtoan_do, dung_doan, ngaiman_cheung}@sutd.edu.sg

Binary Hashing for Image Search

- Binary hashing
  - Efficient storage and fast searching
    → Attractive approach for large scale visual search
- Traditional approach for hashing
  - Minimizing the specific loss function under the binary constraint on the codes
  - → coupling of the hash function and the binary constraint → very challenge to solve
- Two-step approach for hashing [TSH[15], FastHash[16]]
  - inference binary codes
  - learn hash function, given binary codes
  - → reduce the complexity of the coupled problem; flexible using of different of hash functions

Contributions

- Unified formulation for both supervised and unsupervised hashing
- ***Two novel approaches for inferencing binary codes***
  - Semidefinite Programming
  - Augmented Lagrangian

Unified formulation

- Input: $\mathbf{S}$: similarity matrix between samples, i.e., pairwise distance matrix for unsupervised or pairwise label matrix for supervised; L: code length; n: number of training samples
- Target: learning binary codes $\mathbf{Z}$ s.t. the similarity matrix $\mathbf{S}$ is preserved in Hamming space, i.e., solving

$$\min_{\mathbf{Z}\in\{-1,1\}^{L\times n}} \left\| \mathbf{Z}^T\mathbf{Z} - \mathbf{Y} \right\|^2 \quad (1)$$

  - unsupervised: $\mathbf{Y} = \mathbf{L} - \frac{L\mathbf{S}}{2}$, where $\mathbf{S}$ is pairwise distance matrix
  - supervised: $\mathbf{Y} = \mathbf{L} - 2L\mathbf{S}$, where S is label pairwise matrix
- Using coordinate descent approach for solving above NP-hard, i.e., solving one row of $\mathbf{Z}$ at a time. Let $\mathbf{x} = [x_1, ..., x_n]^T = \mathbf{z}^{(k)}$, we solve BQP

$$\min_{\mathbf{x}} \mathbf{x}^T\mathbf{A}\mathbf{x} \quad s.t.\ {x_i}^2 = 1, \forall i = 1, ..., n. \quad (2)$$

where $\mathbf{A} = \{a_{ij}\} \in \mathbb{R}^{n\times n}$; $a_{ij} = \bar{\mathbf{z}}_i^T\bar{\mathbf{z}}_j - y_{ij}$.

Semidefinite Relaxation (SDR) approach

Let $\mathbf{B} = \mathbf{A} - \lambda_1\mathbf{I}$, where λ_1 is the largest eigenvalue of $\mathbf{A}$; $\mathbf{X} = \mathbf{x}\mathbf{x}^T$
→ **solve equivalent problem**

$$\min_{\mathbf{X}} \ trace(\mathbf{B}\mathbf{X}) \quad s.t.\ diag(\mathbf{X}) = \mathbf{1}; \mathbf{X} \succeq 0; rank(\mathbf{X}) = 1 \quad (3)$$

Two-step solution:

- Drop rank one constraint → solving Semidefinite Program

$$\min_{\mathbf{X}} \ trace(\mathbf{B}\mathbf{X}) \quad s.t.\ diag(\mathbf{X}) = \mathbf{1}; \mathbf{X} \succeq 0 \quad (4)$$

  Solving: Using Convex OPT packages: SeDuMi, SDPT3 → achieving the global optimal solution $\mathbf{X}^*$
- Recover binary solution $\hat{\mathbf{x}}$ from $\mathbf{X}^*$
  Solving: apply randomized rounding process several times and select the best solution
  - generate ξ by $\xi \sim \mathcal{N}(0, \mathbf{X}^*)$
  - get feasible point:

$$\hat{\mathbf{x}} = sgn(\xi) \quad (5)$$

- **Bound on objective value at** $\hat{\mathbf{x}}$: Let f_{opt} be global optimum objective value of (3) and $f_{SDR-round} = \hat{\mathbf{x}}^T\mathbf{B}\hat{\mathbf{x}}$, we have

$$f_{opt} \le E[f_{SDR-round}] \le \frac{2}{\pi} f_{opt} \quad (6)$$

Augmented Lagrangian (AL) approach

Let $\Phi(\mathbf{x}) = \left[(x_1)^2 - 1, ..., (x_n)^2 - 1\right]^T$; $\Lambda = [\lambda_1, ..., \lambda_n]^T$: Lagrange multipliers, by using AL for (2), we minimize the unconstrained augmented Lagrangian function

$$\mathcal{L}(\mathbf{x}, \Lambda; \mu) = \mathbf{x}^T\mathbf{A}\mathbf{x} - \Lambda^T\Phi(\mathbf{x}) + \frac{\mu}{2}\|\Phi(\mathbf{x})\|^2 \quad (7)$$

- When μ is large → penalize the binary constraint violation severely → force the minimizer of the AL function (7) closer to the feasible region of the original problem (2)
- Theoretically, not necessary to take $\mu \to \infty$ in order to achieve a local optimum of (2)

Compare SDR and AL approaches

	Computational	Memory
SDR	$\mathcal{O}(n^{4.5})$	$\mathcal{O}(n^2)$
AL	$\mathcal{O}(tt_1n^2)$; $t_1 \le 50$; $t \le 10$	$\mathcal{O}(n)$

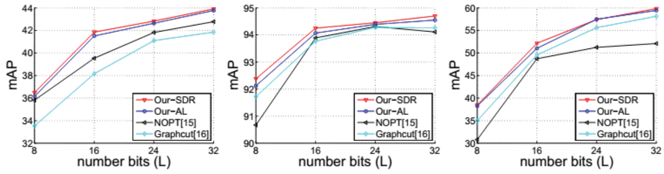
Table 1: Memory and computational complexity of SDR and AL

Dataset

- CIFAR10: GIST features; database: 50K; training: 5K; testing: 10K
- MNIST: raw (intensity) features; database: 60K; training: 5K; testing: 10K
- SUN397: AlexNet features; database: 31K; training: ∼5K; testing: 4.2K

Comparison with state-of-the-art binary inference methods

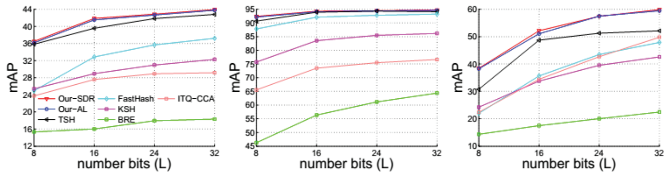
- Given inferenced binary codes, SVM-RBF is used as hash functions for all compared methods.



(a) CIFAR10 (b) MNIST (c) SUN397

Figure 1: mAP comparison of different binary inference methods

Evaluation of Supervised Hashing



(a) CIFAR10 (b) MNIST (c) SUN397

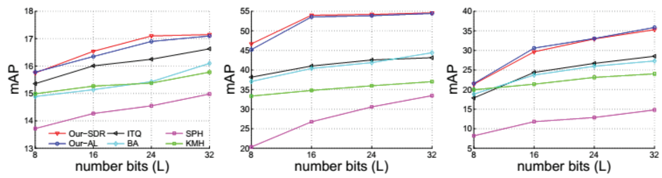
Figure 2: mAP comparison with state-of-the-art supervised hashing methods

	CIFAR10				MNIST				SUN397			
L	8	16	24	32	8	16	24	32	8	16	24	32
Our-SDR	30.57	46.61	48.22	48.43	86.33	93.86	94.26	94.56	12.11	59.19	62.98	61.78
Our-AL	30.07	46.33	47.95	48.02	85.96	93.49	94.09	94.36	12.03	57.20	63.14	61.45
TSH[15]	29.09	45.69	47.41	47.66	81.14	93.41	93.88	93.84	10.08	55.70	60.28	59.30
FastHash[16]	22.85	40.81	42.25	32.49	66.22	92.14	92.79	91.41	8.91	46.84	51.84	39.40
KSH[14]	24.26	37.26	40.95	36.52	54.29	86.94	89.31	88.33	11.79	39.41	51.28	46.48
BRE[13]	16.19	22.74	28.87	18.41	36.67	70.59	81.45	82.83	9.62	27.93	39.42	30.39
ITQ-CCA[8]	22.66	35.36	38.39	39.13	53.46	79.70	82.98	83.43	11.67	36.35	49.19	46.81

Table 2: Precision at Hamming distance r = 2 comparison with state-of-the-art supervised hashing methods

- SDR slightly outperforms AL
- SDR and AL outperform most of compared methods.

Evaluation of Unsupervised Hashing



(a) CIFAR10 (b) MNIST (c) SUN397

Figure 3: mAP comparison with state-of-the-art unsupervised hashing

	CIFAR10				MNIST				SUN397			
L	8	16	24	32	8	16	24	32	8	16	24	32
Our-SDR	17.19	22.82	27.40	25.87	43.08	73.72	81.34	82.17	12.17	32.15	44.28	45.38
Our-AL	17.34	23.23	27.26	25.21	42.09	74.36	81.50	82.29	11.99	33.34	44.13	45.60
ITQ[8]	15.55	22.49	26.69	15.36	33.40	69.96	81.36	74.70	9.75	30.80	42.07	34.70
BA[35]	15.62	22.65	26.55	11.42	32.62	69.03	79.11	74.00	10.15	31.61	42.52	31.97
SPH[10]	14.66	20.32	24.67	12.32	20.77	51.74	72.20	63.38	6.38	20.66	30.10	19.97
KMH[9]	15.11	22.57	27.25	10.36	32.45	64.42	79.97	65.79	9.88	31.04	43.67	28.85

Table 3: Precision at Hamming distance $r = 2$ comparison with state-of-the-art unsupervised hashing

- SDR outperform AL
- SDR and AL outperform all compared methods, i.e., 5%-10% mAP on MNIST and SUN397.

Conclusion

- Unified framework for both unsupervised and supervised hashing
- Two novel approaches, i.e. Semidefinite Relaxation and Augmented Lagrangian, for binary code inference