# Label Embedding Online Hashing for Cross-Modal Retrieval

Yongxin Wang, Xin Luo, Xin-Shun Xu; School of Software, Shandong University, China

## CONTRIBUTIONS

- A novel supervised online cross-modal hashing framework is presented. It can not only fully exploit the semantic information by label embedding, but also consider the correlation between new and existing data.
- A discrete and efficient optimization algorithm is proposed for LEMON. The binary codes can be generated discretely without relaxation. And, it efficiently updates hash functions by keeping old hash codes unchangeable.
- Extensive experiential results demonstrate the superiority of LEMON.

## PROPOSED METHOD

### Label Embedding

– binary codes should preserve the semantic similarity

$$\min_{\mathbf{B}^{(t)}} \|\mathbf{B}^{(t)\top}\mathbf{B}^{(t)} - r\mathbf{S}^{(t)}\|^2, \quad s.t. \mathbf{B}^{(t)} \in \{-1,1\}^{r\times N_t},$$

– labels could be reconstructed from binary codes

$$\min_{\{\mathbf{B},\mathbf{P}\}^{(t)}} \|\mathbf{L}^{(t)} - \mathbf{P}^{(t)}\mathbf{B}^{(t)}\|^2 + \gamma\|\mathbf{P}^{(t)}\|^2, \quad s.t. \mathbf{B}^{(t)} \in \{-1,1\}^{r\times N_t},$$

– jointly considering above functions

$$\min_{\{\mathbf{B},\mathbf{P}\}^{(t)}} \alpha\|\mathbf{B}^{(t)\top}\mathbf{B}^{(t)} - r\mathbf{S}^{(t)}\|^2 + \beta\|\mathbf{L}^{(t)} - \mathbf{P}^{(t)}\mathbf{B}^{(t)}\|^2 + \beta\gamma\|\mathbf{P}^{(t)}\|^2, \quad s.t. \mathbf{B}^{(t)} \in \{-1,1\}^{r\times N_t},$$

### Efficient Discrete Optimization

– alternatively and iteratively update $\{\vec{\mathbf{B}}, \vec{\mathbf{V}}, \mathbf{P}, \mathbf{R}\}^{(t)}$

– auxiliary variables $\mathbf{C}_*^{(t-1)}$

E.g. $\mathbf{C}_1^{(t)} = \vec{\mathbf{L}}^{(t)}\vec{\mathbf{V}}^{(t)\top} + \mathbf{L}^{(t)}\mathbf{V}^{(t)\top}$

$\mathbf{C}_1^{(t-1)} \in \mathbb{R}^{c\times r}$ is calculated at the previous round. $\mathbf{C}_1^{(t)}$ is only relevant to new data, i.e., *nt*.

$$\mathbf{L}^{(t)}\mathbf{V}^{(t)\top} = [\mathbf{L}^{(t-1)}, \vec{\mathbf{L}}^{(t-1)}][\mathbf{V}^{(t-1)}, \vec{\mathbf{V}}^{(t-1)}]^\top = \mathbf{L}^{(t-1)}\mathbf{V}^{(t-1)\top} + \vec{\mathbf{L}}^{(t-1)}\vec{\mathbf{V}}^{(t-1)\top}$$

$$\mathbf{C}_1^{(t)} = \vec{\mathbf{L}}^{(t)}\vec{\mathbf{V}}^{(t)\top} + \mathbf{C}_1^{(t-1)}$$

### Online Learning

– define a block similarity matrix $\mathbf{S}^{(t)} = \begin{bmatrix} \mathbf{S}_{oo}^{(t)} & \mathbf{S}_{oc}^{(t)} \\ \mathbf{S}_{co}^{(t)} & \mathbf{S}_{cc}^{(t)} \end{bmatrix}$, $\mathbf{S}^{(t)} = 2\mathbf{U}^{(t)\top}\mathbf{U}^{(t)} - \mathbf{1}\mathbf{1}^\top$,

– keep $\mathbf{B}^{(t)}$ unchanged and only update $\vec{\mathbf{B}}^{(t)}$

$\min_{\{\vec{\mathbf{B}},\mathbf{P}\}^{(t)}} \alpha\|$ [redacted]

$+\beta\|$ [redacted]

– solve update-imbalance problem

$\min_{\{\vec{\mathbf{B}},\vec{\mathbf{V}},\mathbf{P},\mathbf{R}\}^{(t)}} \alpha\|$ [redacted]

$$+\beta\|\vec{\mathbf{L}}^{(t)} - \mathbf{P}^{(t)}\vec{\mathbf{V}}^{(t)}\|^2 + \beta\|\mathbf{L}^{(t)} - \mathbf{P}^{(t)}\mathbf{V}^{(t)}\|^2 + \beta\gamma\|\mathbf{P}^{(t)}\|^2$$

$$+\|\vec{\mathbf{B}}^{(t)} - \mathbf{R}^{(t)}\vec{\mathbf{V}}^{(t)}\|^2 + \|\mathbf{B}^{(t)} - \mathbf{R}^{(t)}\mathbf{V}^{(t)}\|^2, \quad s.t. \vec{\mathbf{B}}^{(t)} \in \{-1,1\}^{r\times n_t}, \mathbf{R}^{(t)}\mathbf{R}^{(t)\top} = \mathbf{I}, \vec{\mathbf{V}}^{(t)}\vec{\mathbf{V}}^{(t)\top} = n_t\mathbf{I}, \vec{\mathbf{V}}^{(t)}\mathbf{1} = \mathbf{0}.$$

### Out-of-Sample

– Hash Mapping: linear regression

$$\min_{\mathbf{W}_m^{(t)}} \|\mathbf{B}^{(t)} - \mathbf{W}_m^{(t)}\mathbf{X}_m^{(t)}\|^2 + \xi\|\mathbf{W}_m^{(t)}\|^2,$$

– Online Learning: $\mathbf{B}^{(t)} = [\mathbf{B}^{(t)}, \vec{\mathbf{B}}^{(t)}]$

$$\min_{\mathbf{W}_m^{(t)}} \|\mathbf{B}^{(t)} - \mathbf{W}_m^{(t)}\mathbf{X}_m^{(t)}\|^2 + \|\vec{\mathbf{B}}^{(t)} - \mathbf{W}_m^{(t)}\vec{\mathbf{X}}_m^{(t)}\|^2 + \xi\|\mathbf{W}_m^{(t)}\|^2.$$

– Efficient Optimization

- auxiliary variables $\mathbf{H}_*^{(t-1)}$ and $\mathbf{F}_*^{(t-1)}$

– Out-of-Sample $H_m^{(t)}(\mathbf{x}_m) = \mathrm{sign}(\mathbf{W}_m^{(t)}\mathbf{x}_m)$.

## EXPERIMENTS

### MAP results on MIRFlickr-25K, and NUS-WIDE

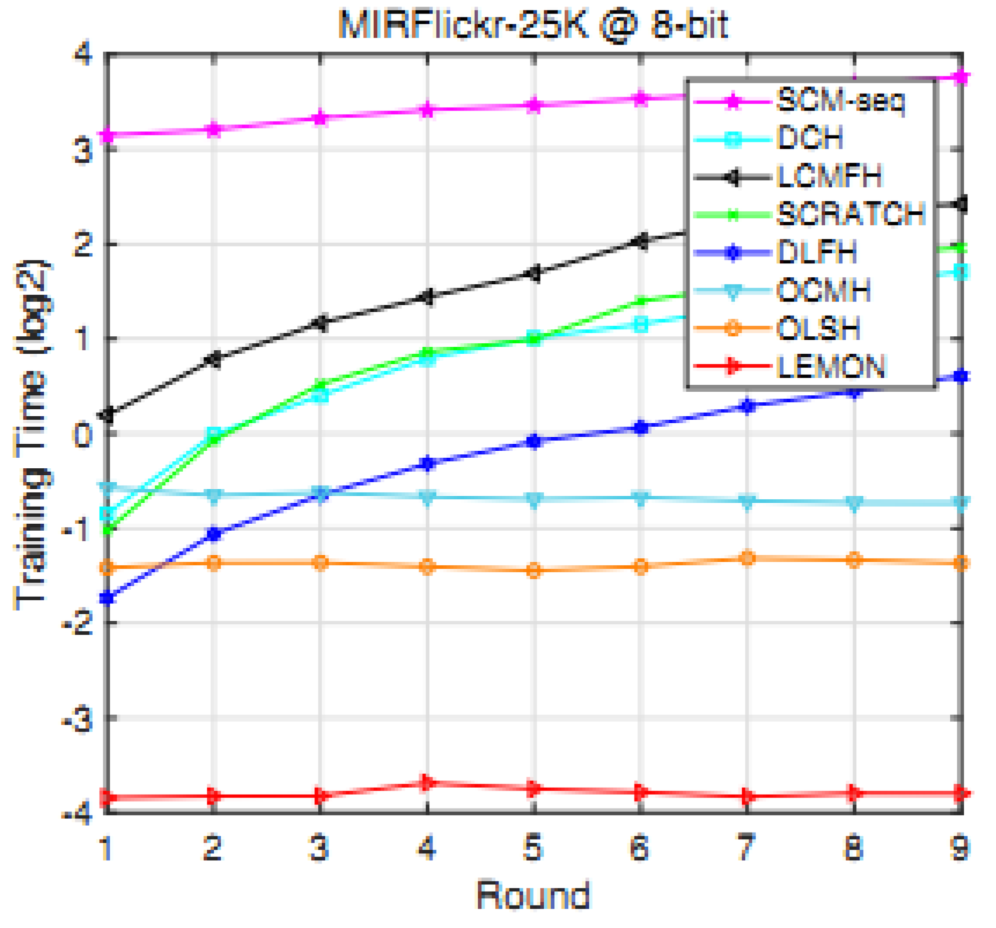
**Table 1: The MAP results of all methods on MIRFlickr-25K.**

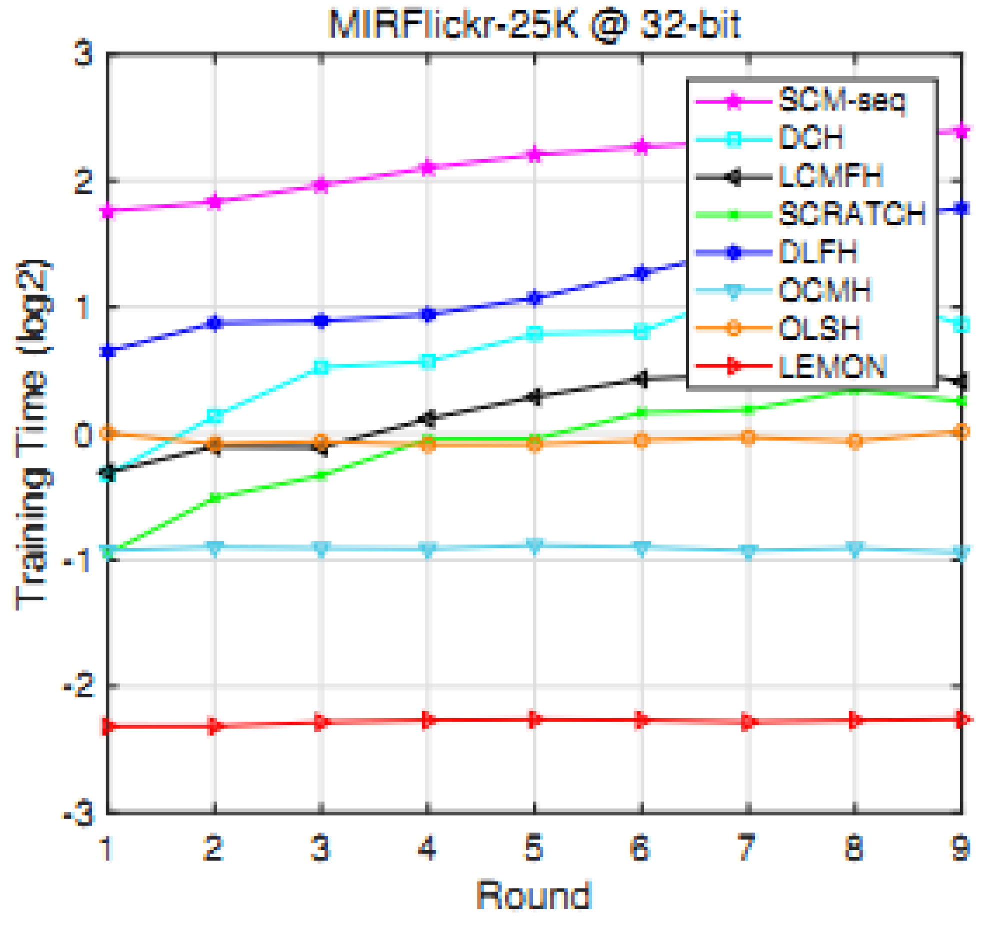
| Task | Method | 8-bit | 16-bit | 32-bit | 64-bit | 128-bit |
|---|---|---|---|---|---|---|
| $I \to T$ | SCM-seq | 0.6307 | 0.6457 | 0.6455 | 0.6629 | 0.6148 |
| | DCH | 0.6739 | 0.7093 | 0.6774 | 0.7219 | 0.7424 |
| | LCMFH | 0.6796 | 0.6750 | 0.6896 | 0.6898 | 0.7002 |
| | SCRATCH | 0.6870 | 0.7084 | 0.7136 | 0.7234 | 0.7256 |
| | DLFH | 0.7102 | 0.7076 | 0.7182 | 0.7188 | 0.7254 |
| | OCMH | 0.5484 | 0.5515 | 0.5578 | 0.5565 | 0.5536 |
| | OLSH | 0.5791 | 0.5778 | 0.6008 | 0.5971 | 0.5935 |
| | LEMON | **0.7272** | **0.7258** | **0.7476** | **0.7474** | **0.7485** |
| $T \to I$ | SCM-seq | 0.6151 | 0.6257 | 0.6245 | 0.6512 | 0.6010 |
| | DCH | 0.7388 | 0.7649 | 0.7410 | 0.7786 | 0.8010 |
| | LCMFH | 0.7332 | 0.7293 | 0.7528 | 0.7627 | 0.7740 |
| | SCRATCH | 0.7446 | 0.7692 | 0.7727 | 0.7810 | 0.7877 |
| | DLFH | 0.7235 | 0.7836 | 0.8066 | 0.8225 | 0.8285 |
| | OCMH | 0.5500 | 0.5530 | 0.5547 | 0.5565 | 0.5533 |
| | OLSH | 0.5801 | 0.5829 | 0.6094 | 0.6038 | 0.6024 |
| | LEMON | **0.7924** | **0.8166** | **0.8238** | **0.8298** | **0.8327** |

**Table 3: The MAP results of all methods on NUS-WIDE.**

| Task | Method | 8-bit | 16-bit | 32-bit | 64-bit | 128-bit |
|---|---|---|---|---|---|---|
| $I \to T$ | SCM-seq | 0.4636 | 0.4324 | 0.5045 | 0.4941 | 0.4941 |
| | DCH | 0.6309 | 0.6076 | 0.5714 | 0.5835 | 0.5937 |
| | LCMFH | 0.5903 | 0.6047 | 0.6353 | 0.6354 | 0.6371 |
| | SCRATCH | 0.6116 | 0.6184 | 0.6413 | 0.6487 | 0.6468 |
| | DLFH | 0.5789 | 0.6211 | 0.6519 | 0.6617 | 0.6684 |
| | OCMH | 0.3447 | 0.3411 | 0.3491 | 0.3429 | 0.3439 |
| | OLSH | 0.4872 | 0.5162 | 0.5243 | 0.5368 | 0.5270 |
| | LEMON | **0.6389** | **0.6579** | **0.6672** | **0.6711** | **0.6700** |
| $T \to I$ | SCM-seq | 0.4812 | 0.4536 | 0.5276 | 0.5253 | 0.5286 |
| | DCH | 0.7615 | 0.7419 | 0.6868 | 0.7110 | 0.7349 |
| | LCMFH | 0.6807 | 0.7275 | 0.7455 | 0.7549 | 0.7552 |
| | SCRATCH | 0.7241 | 0.7567 | 0.7673 | 0.7853 | 0.7893 |
| | DLFH | 0.6779 | 0.7736 | 0.8066 | 0.8071 | 0.8140 |
| | OCMH | 0.3454 | 0.3420 | 0.3539 | 0.3417 | 0.3436 |
| | OLSH | 0.5167 | 0.5594 | 0.5648 | 0.5920 | 0.5804 |
| | LEMON | **0.7778** | **0.7946** | **0.8122** | **0.8288** | **0.8333** |

### Training time-round curves on MIRFlickr-25K




## CONTRACT

Yongxin Wang: yxinwang@hotmail.com
Xin-Shun Xu: xuxinshun@sdu.edu.cn
MIMA Group: http://mima.sdu.edu.cn/

*0436