多模态及在线哈希算法

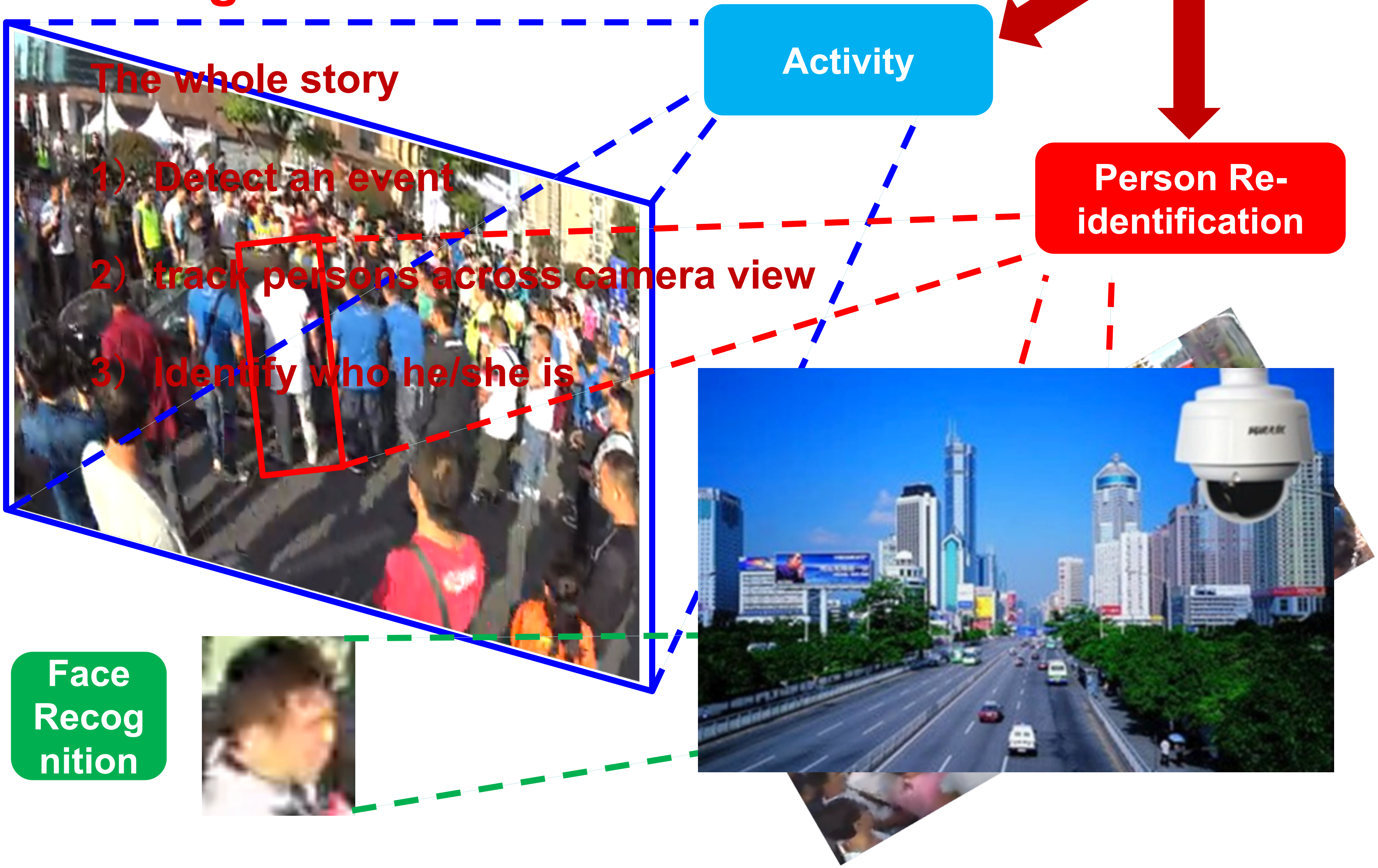
WEI-SHI ZHENG (郑伟诗)

http://isee.sysu.edu.cn/~zhwshi

SUN YAT-SEN University

Human Identification & Activity

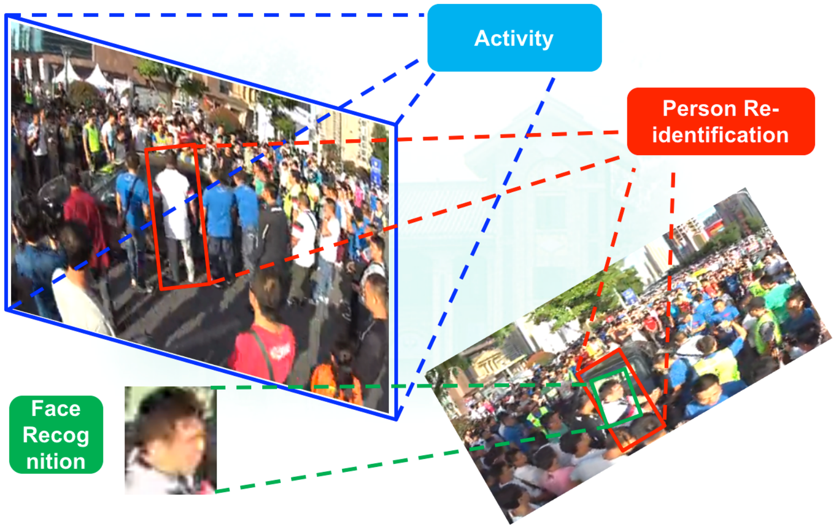
Background



Human Identification & Activity

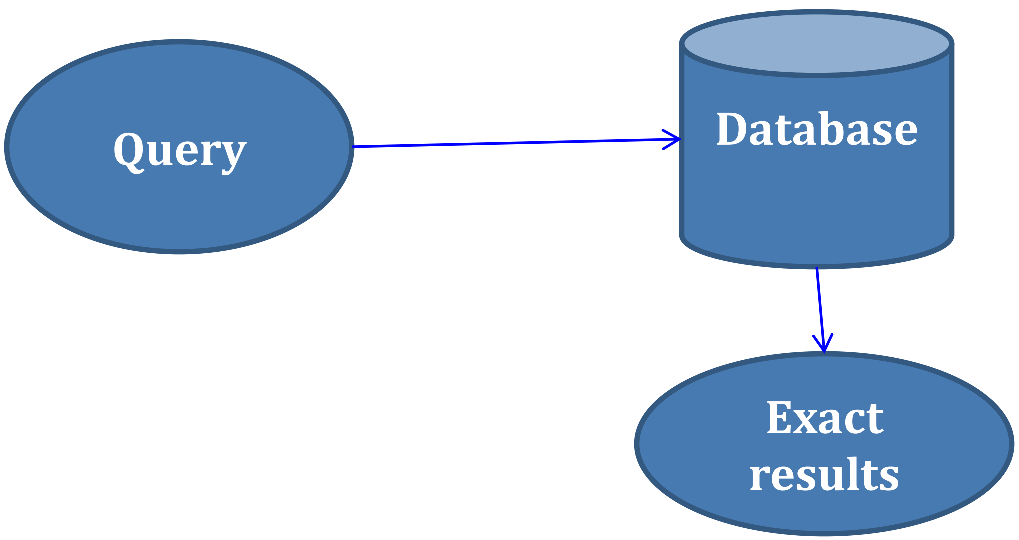
A Standard Processing

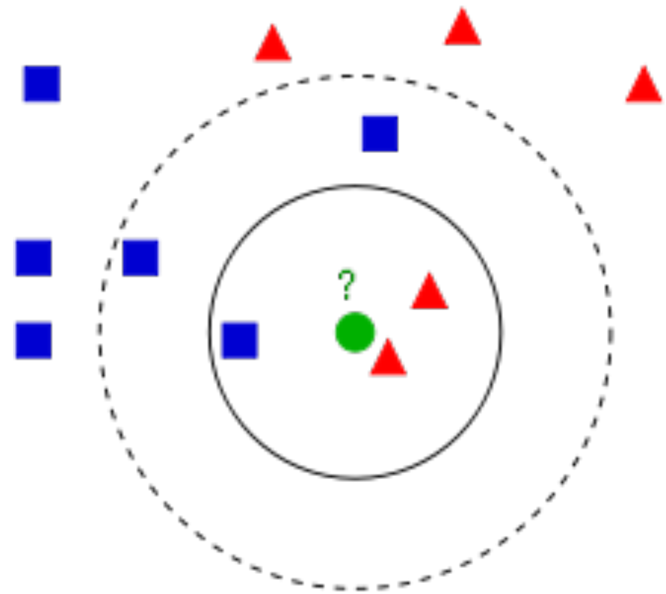
- Detect what is happening
- Who is the suspect?
- Track him/her across camera views
- Search him/her in large dataset
- Recognise him/her



About Hashing based Search

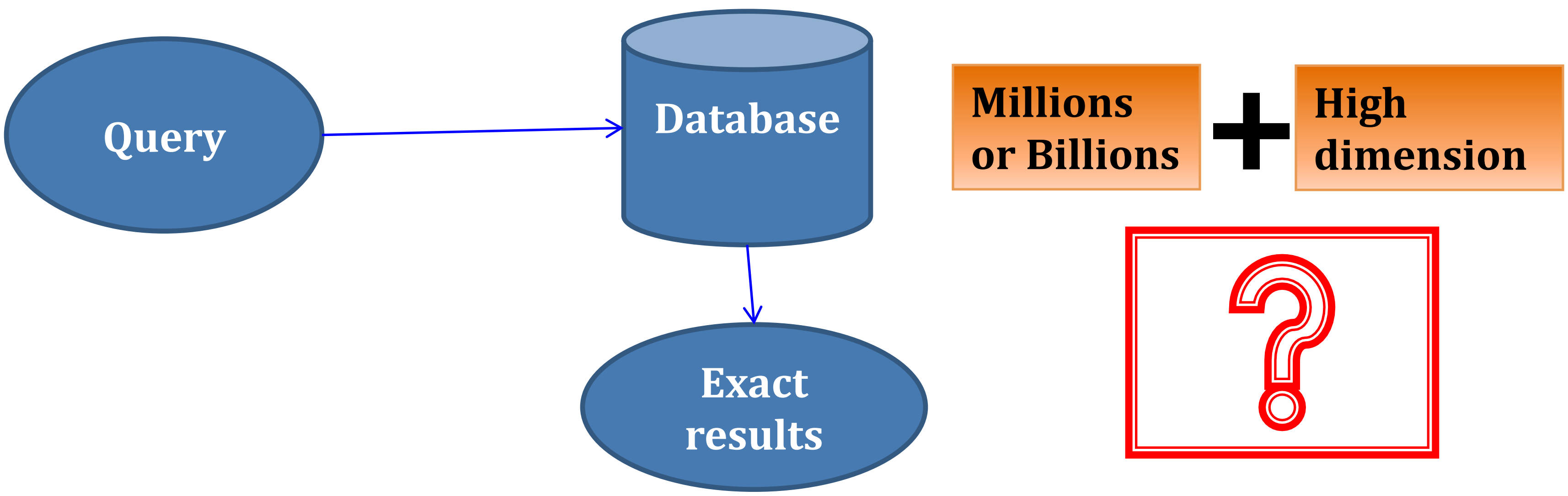
- K Near Neighbour Search
  - Find the nearest point for the query (green point)





About Hashing based Search

- **Problems traditional NN search faced**
  - Large data size, e.g., millions or billions;
  - High dimension
  - Exhaust time search
  - Large storage space

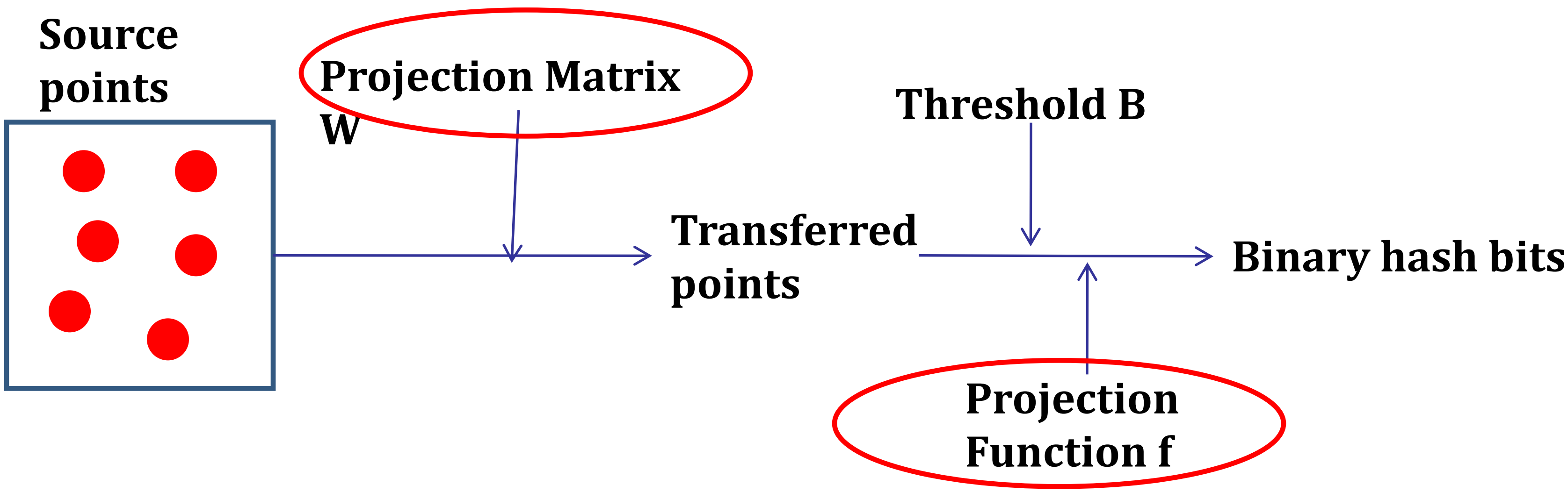


About Hashing based Search

- A solution using Hashing

$$h_k(x) = \mathrm{sgn}(f(w_k^T x + b_k))$$

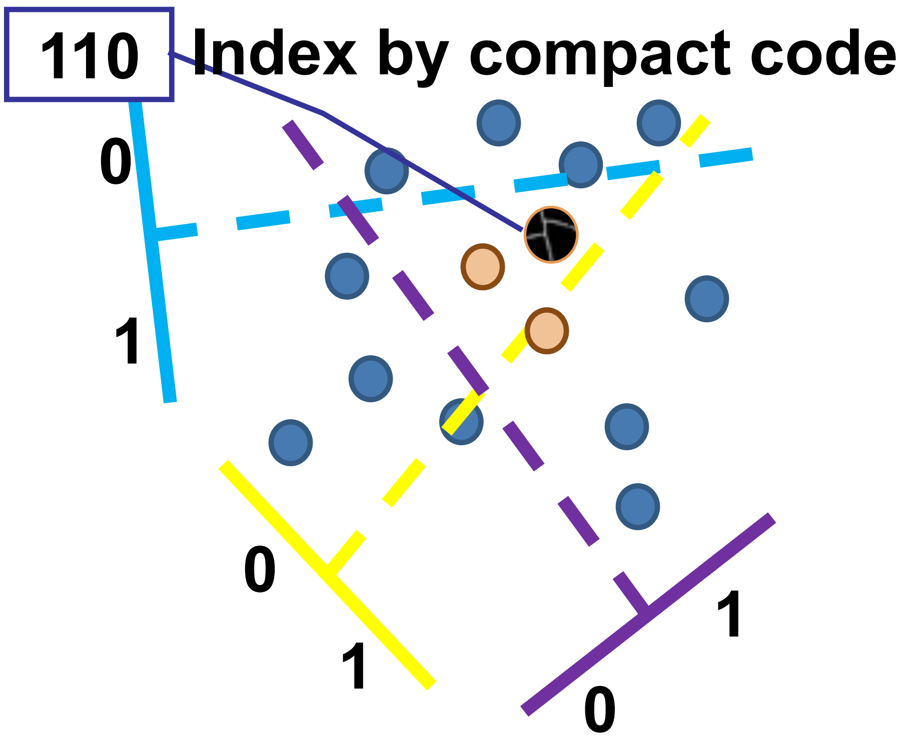
x —data point w_k —projection vector b_k —threshold



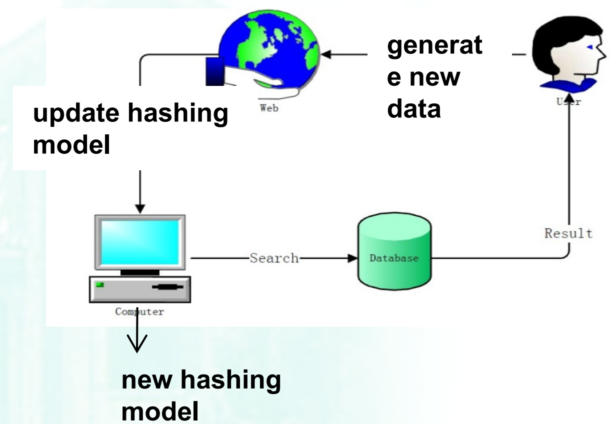
Different choices of W and f(.) lead to different hashing approaches.

About Hashing based Search

$$h_k(x) = sgn(w^T x + b \quad)$$



Learning Hashing Codes Online



Long-Kai Huang, Qiang Yang, Wei-Shi Zheng*. Online Hashing. IEEE Transactions on Neural Networks and Learning Systems, 2017 (DOI: 10.1109/TNNLS.2017.2689242)

Online Hashing

- Update a hashing model online to process sequential pairwise data

- We assume a pairwise data comes at a time
  - It can be easily generalised

- Label: similar or non-similar

$$\mathbb{x}^t = [\mathrm{x}_\mathrm{i}^\mathrm{t}, \mathrm{x}_\mathrm{j}^\mathrm{t}]$$

$$s = \begin{cases} 1, & if\ \mathbf{x}_\mathrm{i} \text{ and } \mathbf{x}_\mathrm{j} \text{ are similar} \\ -1, & if\ \mathbf{x}_\mathrm{i} \text{ and } \mathbf{x}_\mathrm{j} \text{ are not similar} \end{cases}$$

Background

- **The Hash Function:**

$$h_k(\mathbf{x}) = sgn(\mathbf{w}_k^T\mathbf{x} + b_k) = \begin{cases} 1, & if \quad \mathbf{w}_k^T\mathbf{x} + b_k \geq 0, \\ -1, & otherwise, \end{cases}$$

- **Online Learning History for Hashing**
  - Online Kernel Hashing (IJCAI 2013)
  - Sketching Hashing (CVPR 2015)
  - LEGO-LSH (NIPS, 2008)

Online Hashing (OH)

- The hashing model we address

$$\mathbf{h}(\mathbf{x}) = sgn(\mathbf{W}^T\mathbf{x})$$

$$\Downarrow$$

$$\mathbf{h} = \arg\max_{\mathbf{f}\in\{-1,1\}^r} \mathbf{f}^T\mathbf{W}^T\mathbf{x}$$

Goodness: we can directly learn W more easily by getting rid of sgn

Quantifying the Loss

Hamming distance

$$R(\mathbb{h}^t, s^t) = \begin{cases} max\{0, \mathcal{D}_h(\mathbf{h_i}^t, \mathbf{h_j}^t) - \alpha\}, & if\ \ s^t = 1, \\ max\{0, \beta r - \mathcal{D}_h(\mathbf{h_i}^t, \mathbf{h_j}^t)\}, & if\ \ s^t = -1, \end{cases}$$

$$s^t = \begin{cases} 1, & if\ \ \mathbf{x_i}^t \text{ and } \mathbf{x_j}^t \text{ are similar,} \\ -1, & if\ \ \mathbf{x_i}^t \text{ and } \mathbf{x_j}^t \text{ are not similar.} \end{cases}$$

The Processing

$$\mathbb{x}^t = [\mathbf{x_i}^t, \mathbf{x_j}^t]$$

$R(\mathbb{h}^t, s^t) = 0$ → **No update**

$R(\mathbb{h}^t, s^t) > 0$ ↓

Update the previous model

Find zero- loss hash code pair $\mathbb{g}^t = [\mathbf{g_i}^t, \mathbf{g_j}^t]$

$$R(\mathbb{g}^t, s^t) = 0$$

Obtain an updated hash projection matrix $\mathbf{W}^{t+1}$

$$\ell^t(\mathbf{W}^{t+1}) = \max_{\mathbb{f} \in \{-1,1\}^{r\times 2}} H^t(\underbrace{\mathbf{W}^{t+1}, \mathbb{f}) - G^t(\mathbf{W}^{t+1}}_{\text{predicts similar hash code pair towards the zero-loss hash code pair}}) + \sqrt{R(\mathbb{h}^t, s^t)}.$$

Online Kernel Hashing (OKH)

- Criterion to optimise

$$\mathbf{W}^{t+1} = \arg\min_{\mathbf{W}} \frac{1}{2}||\mathbf{W} - \mathbf{W}^t||_F^2 + C\xi$$

$$s.t. \quad \ell^t(\mathbf{W}) \leq \xi \quad \text{and} \quad \xi \geq 0,$$

$$\ell^t(\mathbf{W}^{t+1}) = \max_{\mathbb{f} \in \{-1,1\}^{r \times 2}} H^t(\mathbf{W}^{t+1}, \mathbb{f}) - G^t(\mathbf{W}^{t+1}) + \sqrt{R(\mathbb{h}^t, s^t)}.$$

$$H^t(\mathbf{W}, \mathbb{f}) = \mathbf{f_i^t}^T \mathbf{W}^T \mathbf{x_i}^t + \mathbf{f_j^t}^T \mathbf{W}^T \mathbf{x_j}^t$$

$$G^t(\mathbf{W}) = \mathbf{g_i}^{tT} \mathbf{W}^T \mathbf{x_i}^t + \mathbf{g_j}^{tT} \mathbf{W}^T \mathbf{x_j}^t$$

Bounds

Lemma 1. *Let* $(\mathbb{x}^1, s^1), \cdots, (\mathbb{x}^t, s^t)$ *be a sequence of pairwise examples, each with a similarity label* $s^t \in \{1, -1\}$. *The data pair* $\mathbb{x}^t \in \mathbb{R}^{d\times 2}$ *is mapped to a* r*-bit hash code pair* $\mathbb{h}^t \in \mathbb{R}^{r\times 2}$ *through the hash projection matrix* $\mathbf{W}^t \in \mathbb{R}^{d\times r}$. *Let* $\mathbf{U}$ *be an arbitrary matrix in* $\mathbb{R}^{d\times r}$. *If* τ^t *is defined as that in Eq. (10), we then have*

$$\sum_{t=1}^{\infty} \tau^t (2\ell^t(\mathbf{W}^t) - \tau^t ||\mathbb{x}^t(\mathbb{g}^t - \mathbb{h}^t)^T||_F^2 - 2\ell_U^t) \leq ||\mathbf{U} - \mathbf{W}^1||_F^2,$$

where $\mathbf{W}^1$ *is the initialized hash projection matrix that consists of non-zero vectors.*

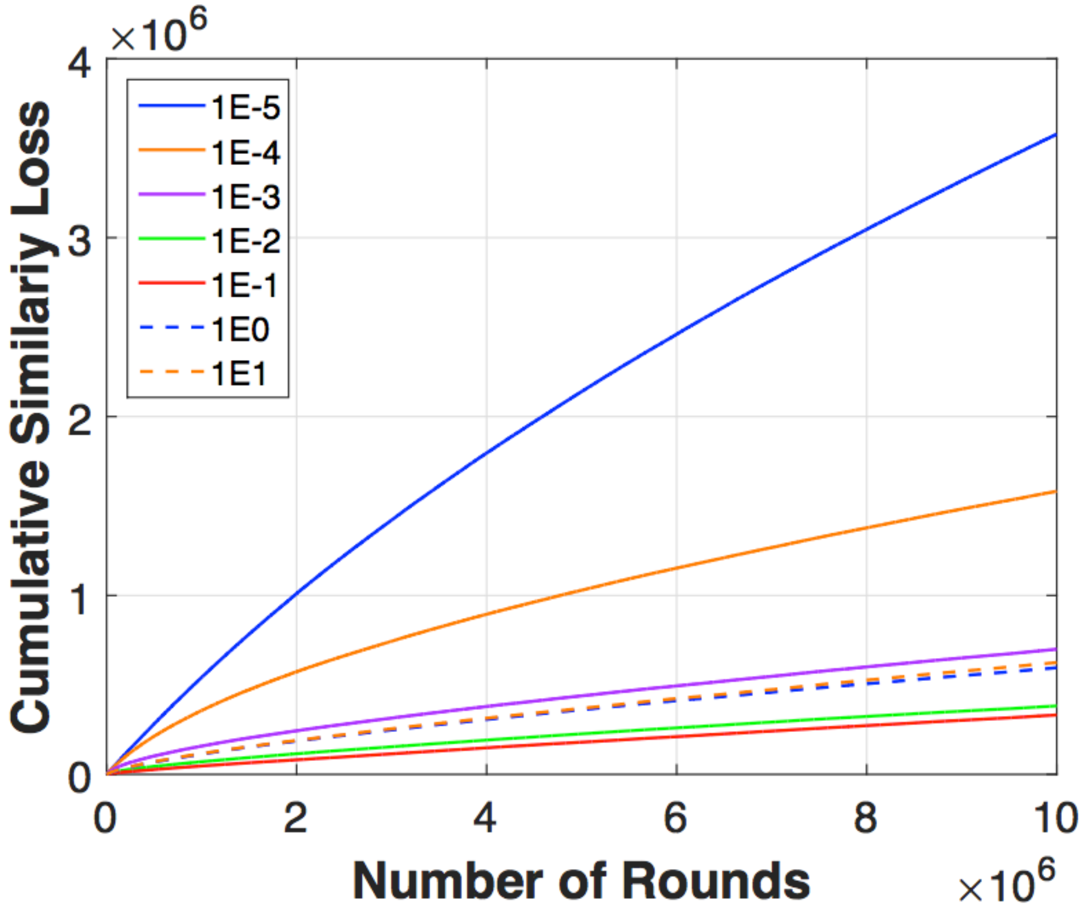
Bounds

Theorem 2. *Let* $(\mathbb{x}^1, s^1), \cdots, (\mathbb{x}^t, s^t)$ *be a sequence of pair-wise examples, each with a similarity label* $s^t \in \{1, -1\}$ *for all* t*. The data pair* $\mathbb{x}^t \in \mathbb{R}^{d\times 2}$ *is mapped to a* r*-bit hash code pair* $\mathbb{h}^t \in \mathbb{R}^{r\times 2}$ *through the hash projection matrix* $\mathbf{W}^t \in \mathbb{R}^{d\times r}$*. If* $||\mathbb{x}^t(\mathbb{g}^t - \mathbb{h}^t)^T||_F^2$ *is upper bounded by* F^2 *and the margin parameter* C *is set as the upper bound of* $\frac{\sqrt{R(\mathbb{h}^t, s^t)}}{F^2}$*, then the cumulative similarity loss (Eq. (3)) is bounded for any matrix* $\mathbf{U} \in \mathbb{R}^{d\times r}$*, i.e.*

$$\sum_{t=1}^{\infty} R(\mathbb{h}^t, s^t) \leq F^2(||\mathbf{U} - \mathbf{W}^1||_F^2 + 2C\sum_{t=1}^{\infty} \ell_U^t),$$

where C *is the margin parameter defined in Criterion (7).*

Bounds



(a) Photo Tourism

Fig. 5. Cumulative similarity loss of OH with C ranging from 0.00001 to 10. (Best viewed in color.)

Bounds

- When retaining more than one model as candidates?

$m_0 = \arg\min_m R_m(\mathbb{h}_m^t, s^t)$

Theorem 3. *Let $(\mathbb{x}^1, s^1), \cdots, (\mathbb{x}^t, s^t)$ be a sequence of pairwise examples, each with a similarity label $s^t \in \{1, -1\}$ for all t. The data pair $\mathbb{x}^t \in \mathbb{R}^{d\times 2}$ is mapped to a r-bit hash code pair $\mathbb{h}^t \in \mathbb{R}^{r\times 2}$ through the hash projection matrix $\mathbf{W}^t \in \mathbb{R}^{d\times r}$. Suppose $||\mathbb{x}^t(\mathbb{g}_m^t - \mathbb{h}_m^t)^T||_F^2$ is upper bounded by F^2, and the margin parameter C is set as the upper bound of $\frac{\sqrt{R_m^*(\mathbb{h}_m^t, s^t)}}{F^2}$ for all m, where $R_m^*(\mathbb{h}_m^t, s^t)$ is an auxiliary function defined as:*

$$R_m^*(\mathbb{h}_m^t, s^t) = \begin{cases} R_m(\mathbb{h}_m^t, s^t), & \textit{if the } m^{th} \textit{ model is selected for update at step } t, \\ 0, & \textit{otherwise.} \end{cases} \tag{28}$$

Then for any matrix $\mathbf{U} \in \mathbb{R}^{d\times r}$, the cumulative similarity loss (Eq. (3)) is bounded, i.e.,

$$\sum_{t=1}^{\infty} \sum_{m=1}^{T} R_m^*(\mathbb{h}_m^t, s^t) \le TF^2(||\mathbf{U} - \mathbf{W}^1||_F^2 + 2C \sum_{t=1}^{\infty} \ell_U^t),$$

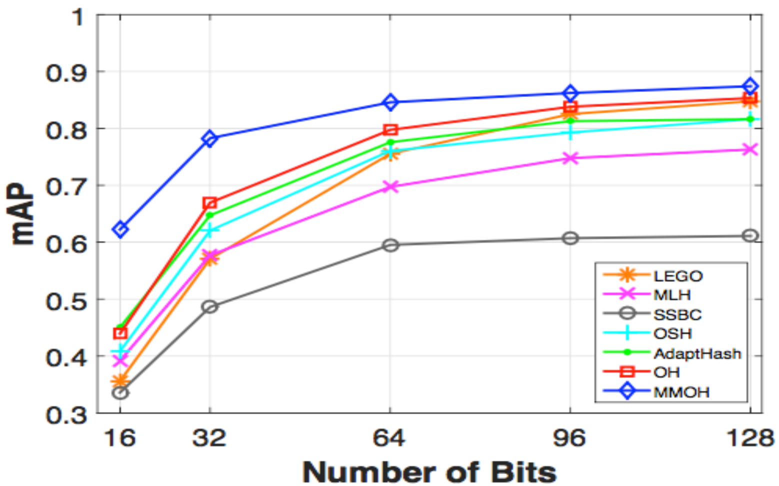
where C is the margin parameter defined in Criterion (7).

Evaluation

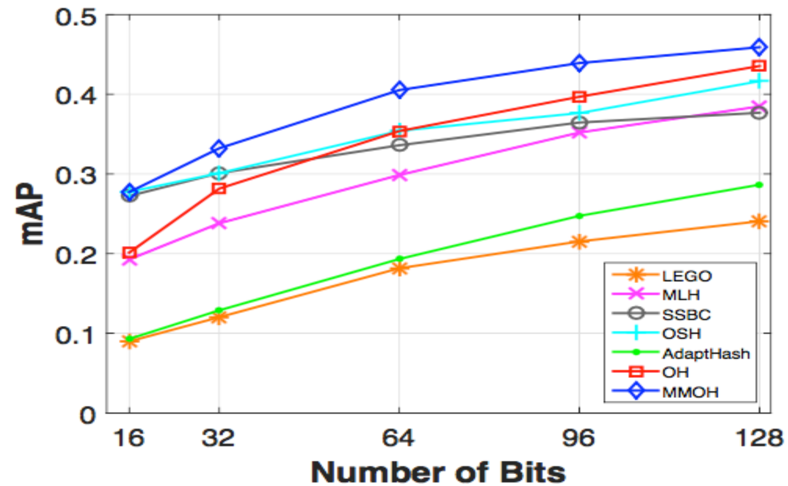
- **Datasets**
  - Photo Tourism: 100K, 512D Gist Features
  - LabelMe: 22K, 512D Gist Features
  - Gist1M: 1M, 960-D Gist Features
  - CIFAR-10 and Tiny Image 80M: 80M, 2048-D deep features
- **Compared Methods**
  - ITQ, KLSH, SDH
  - LEGO-LSH, MLH, SSBC, OSH and AdaptHash.
- **Measurement**
  - MAP: Mean Average Precision
  - Time cost

Evaluation

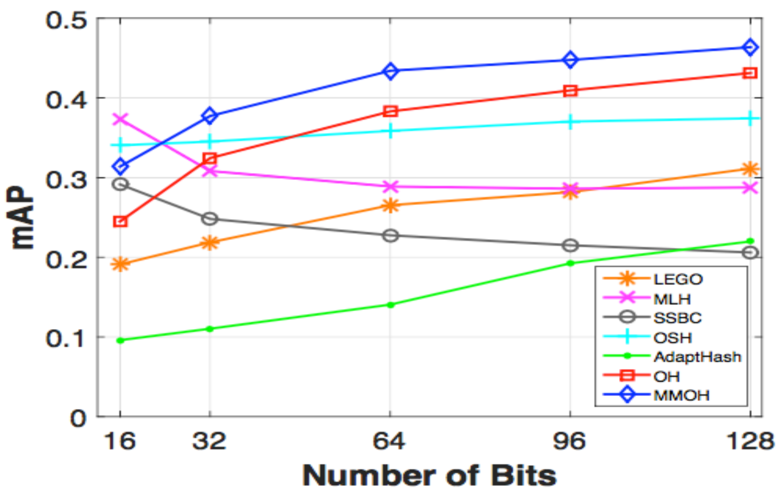
- Comparison with related online methods



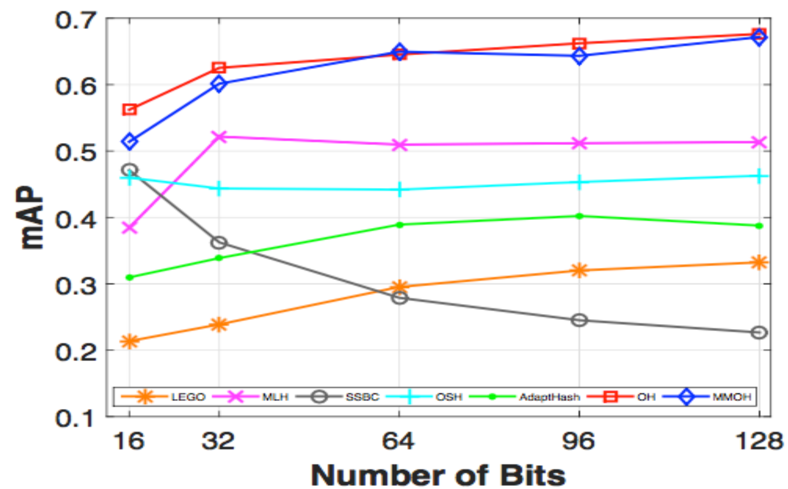
(a) Photo Tourism



(b) 22K LabelMe



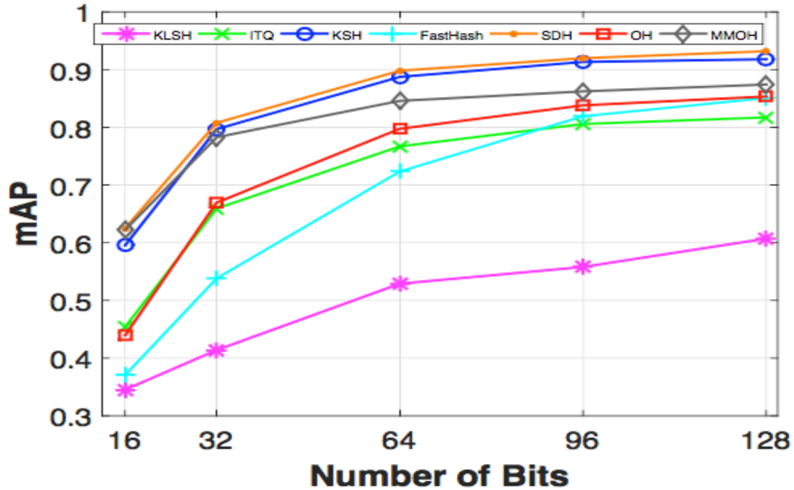
(c) GIST 1M



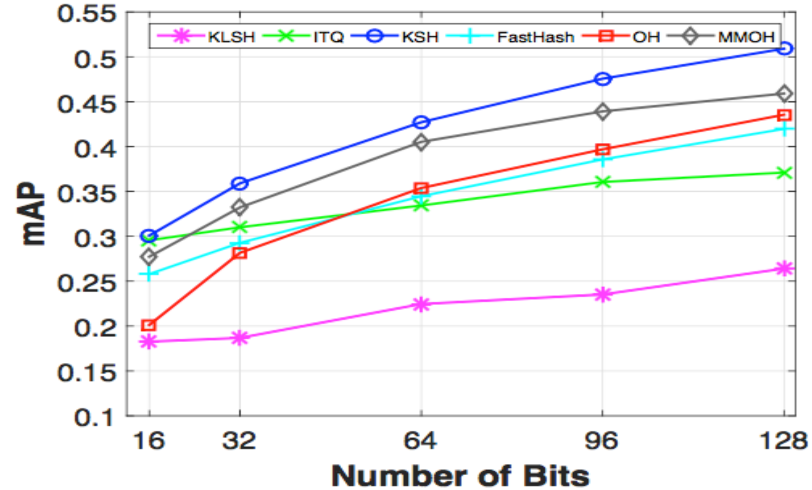
(d) CIFAR-10

Evaluation

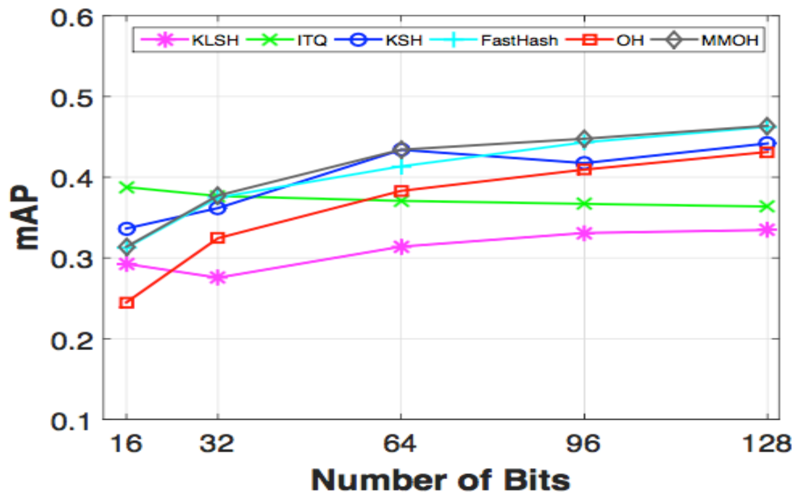
- Comparison with related offline models



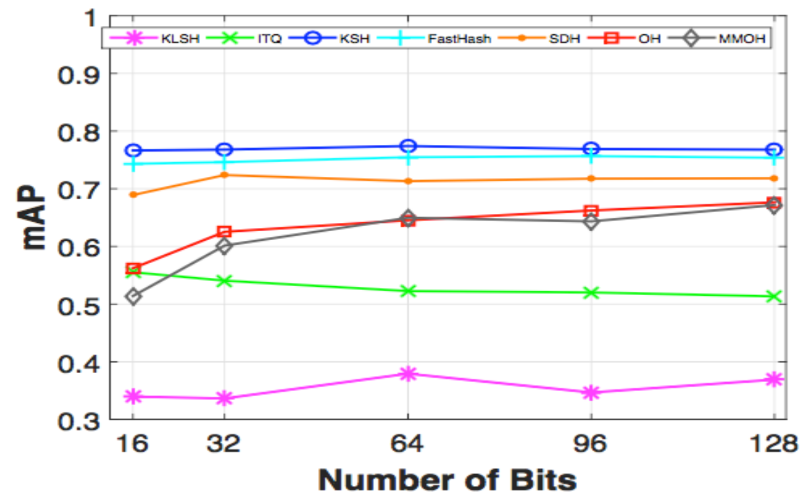
(a) Photo Tourism



(b) 22K LabelMe



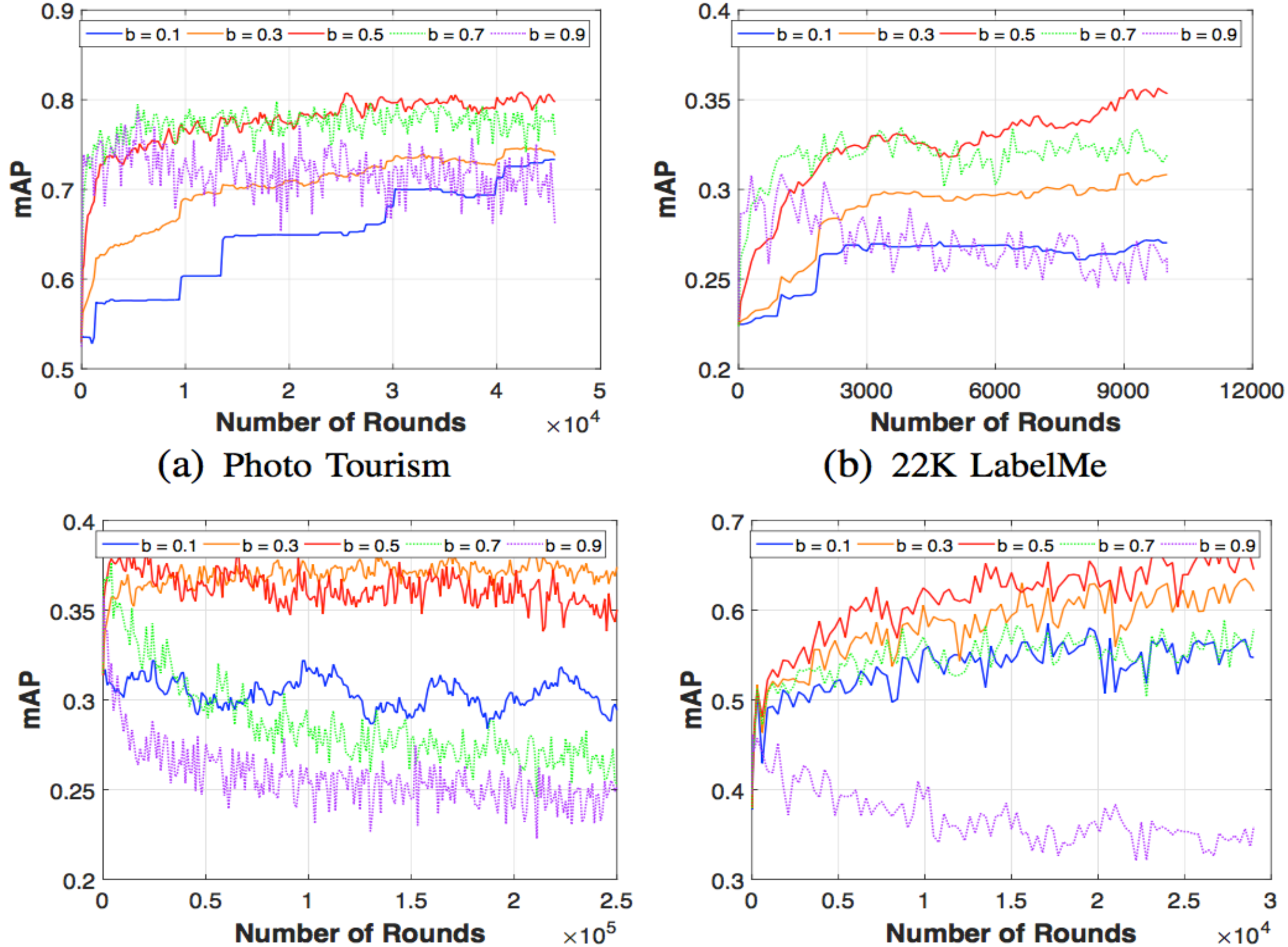
(c) GIST 1M



(d) CIFAR-10

Evaluation

$$R(\mathbb{h}^t, s^t) = \begin{cases} max\{0, \mathcal{D}_h(\mathbf{h_i}^t, \mathbf{h_j}^t) - \alpha\}, & if\ s^t = 1, \\ max\{0, \beta r - \mathcal{D}_h(\mathbf{h_i}^t, \mathbf{h_j}^t)\}, & if\ s^t = -1, \end{cases}$$

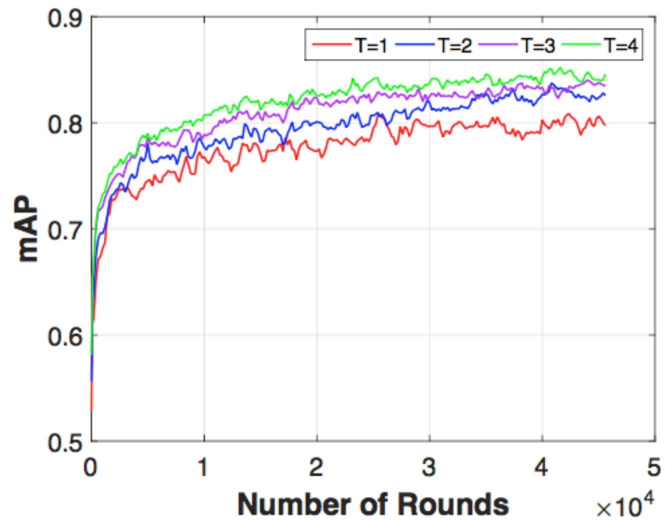


(a) Photo Tourism (b) 22K LabelMe

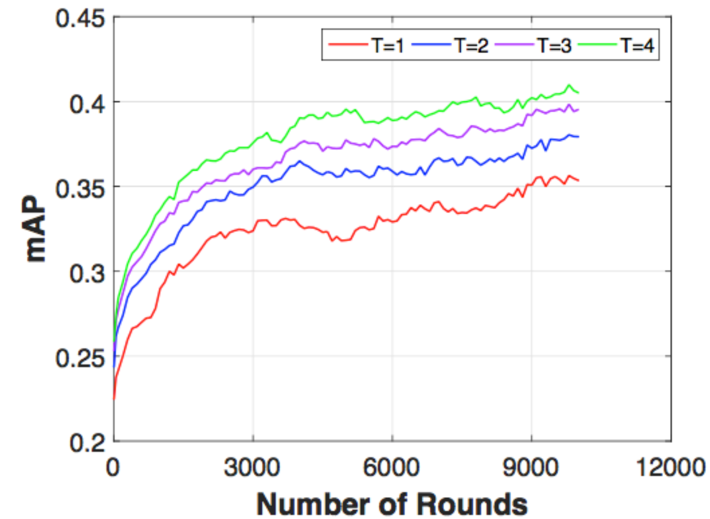
(c) GIST1M (d) CIFAR-10

Evaluation

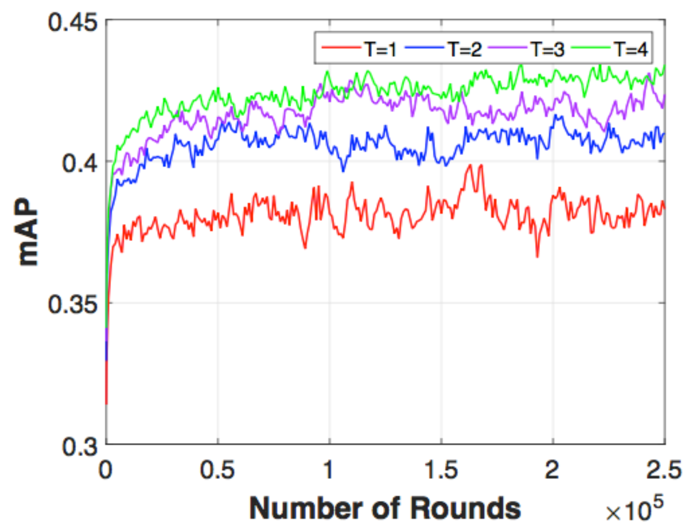
- Different numbers of models



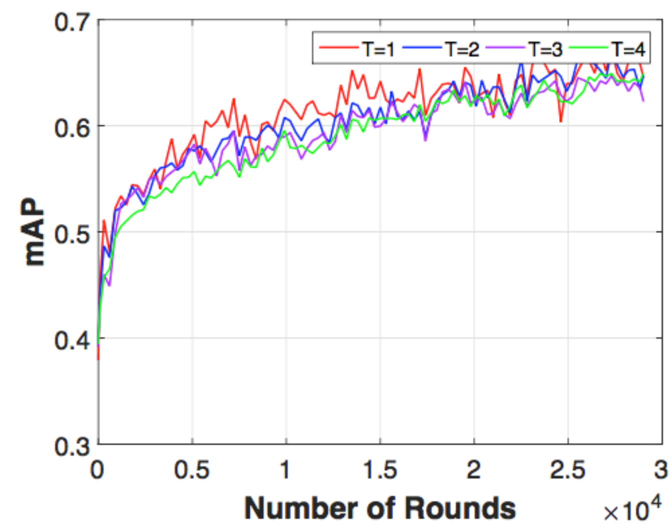
(a) Photo Touris



(b) 22K LabelMe



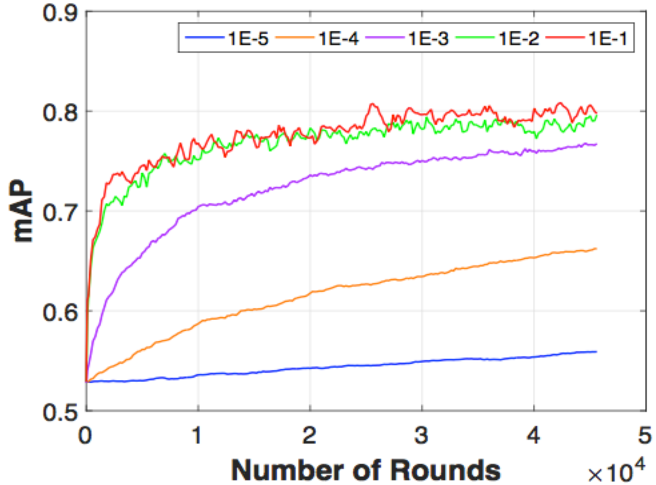
(c) GIST1M



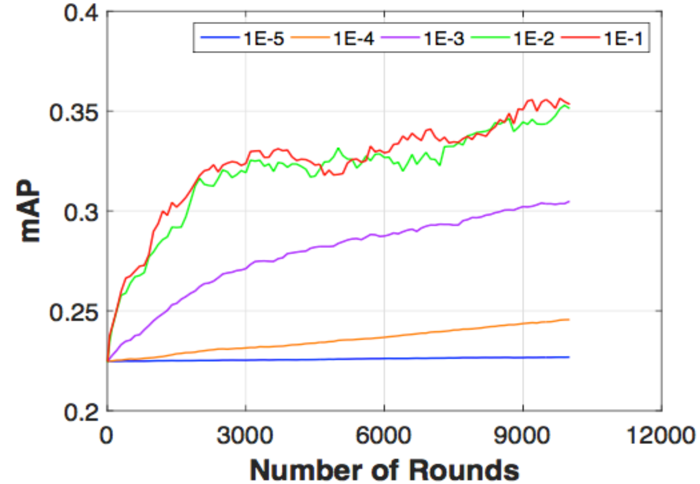
(d) CIFAR-10

Evaluation

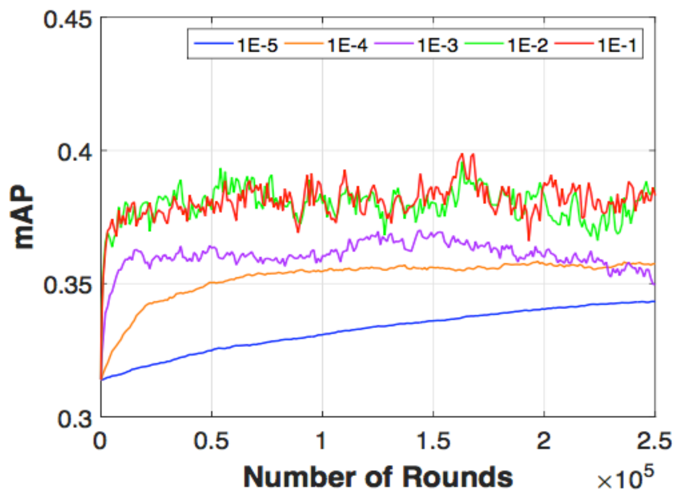
$$\sum_{t=1}^{\infty} R(\mathbb{h}^t, s^t) \le F^2(\|\mathbf{U}-\mathbf{W}^1\|_F^2 + 2C\sum_{t=1}^{\infty}\ell_U^t)$$



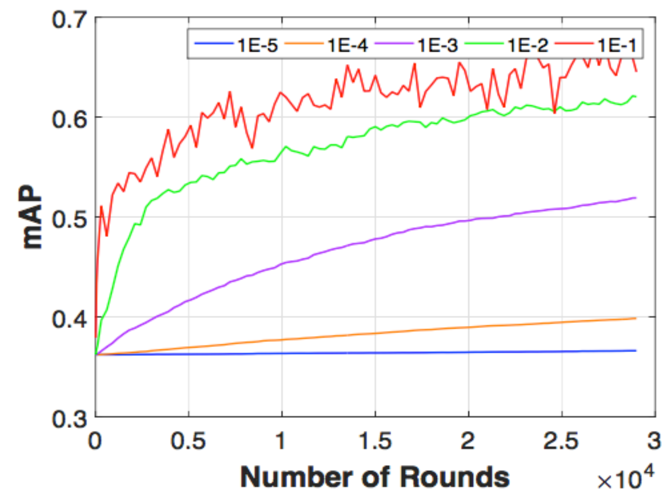
(a) Photo Tourism



(b) 22K LabelMe



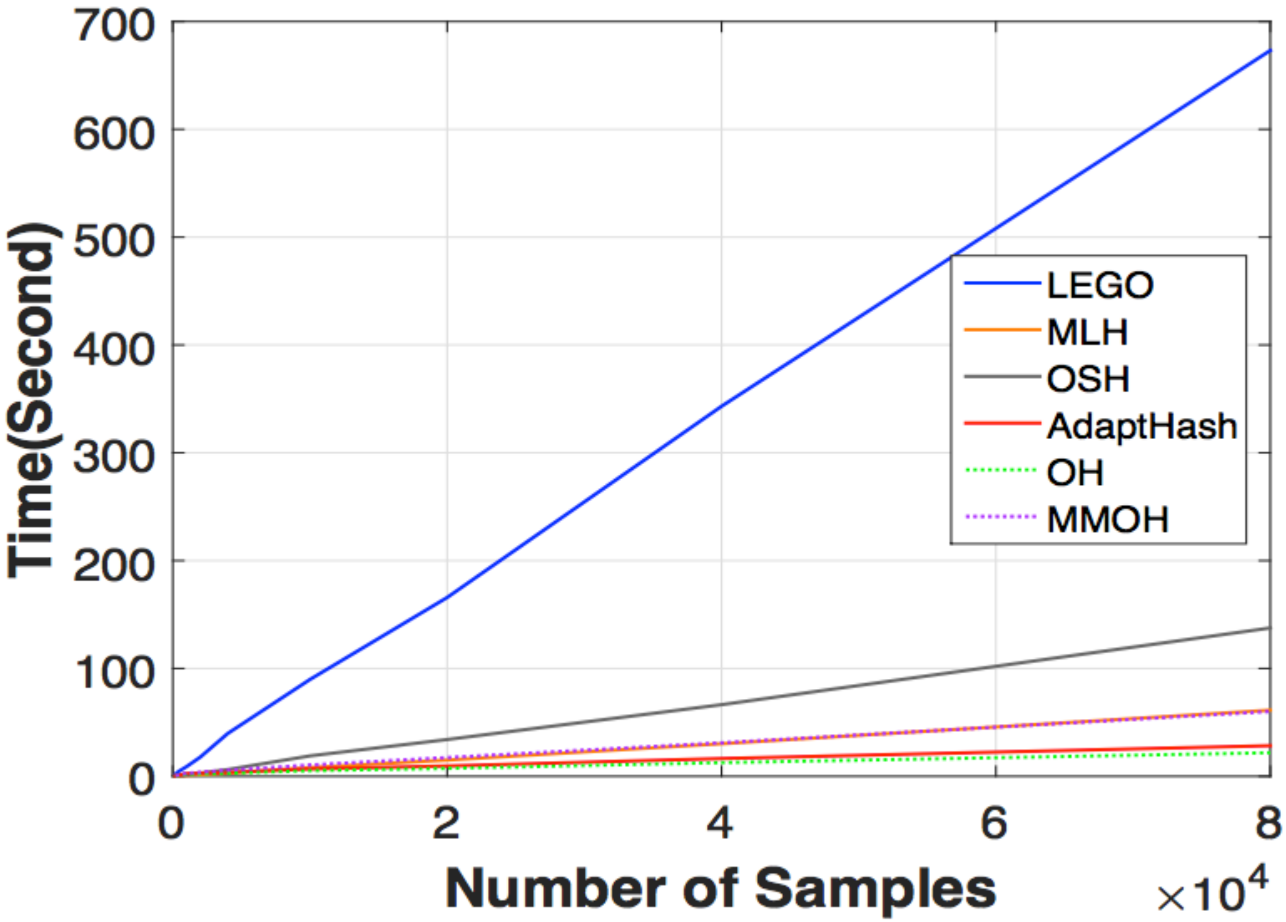
(c) GIST1M



(d) CIFAR-10

Evaluation

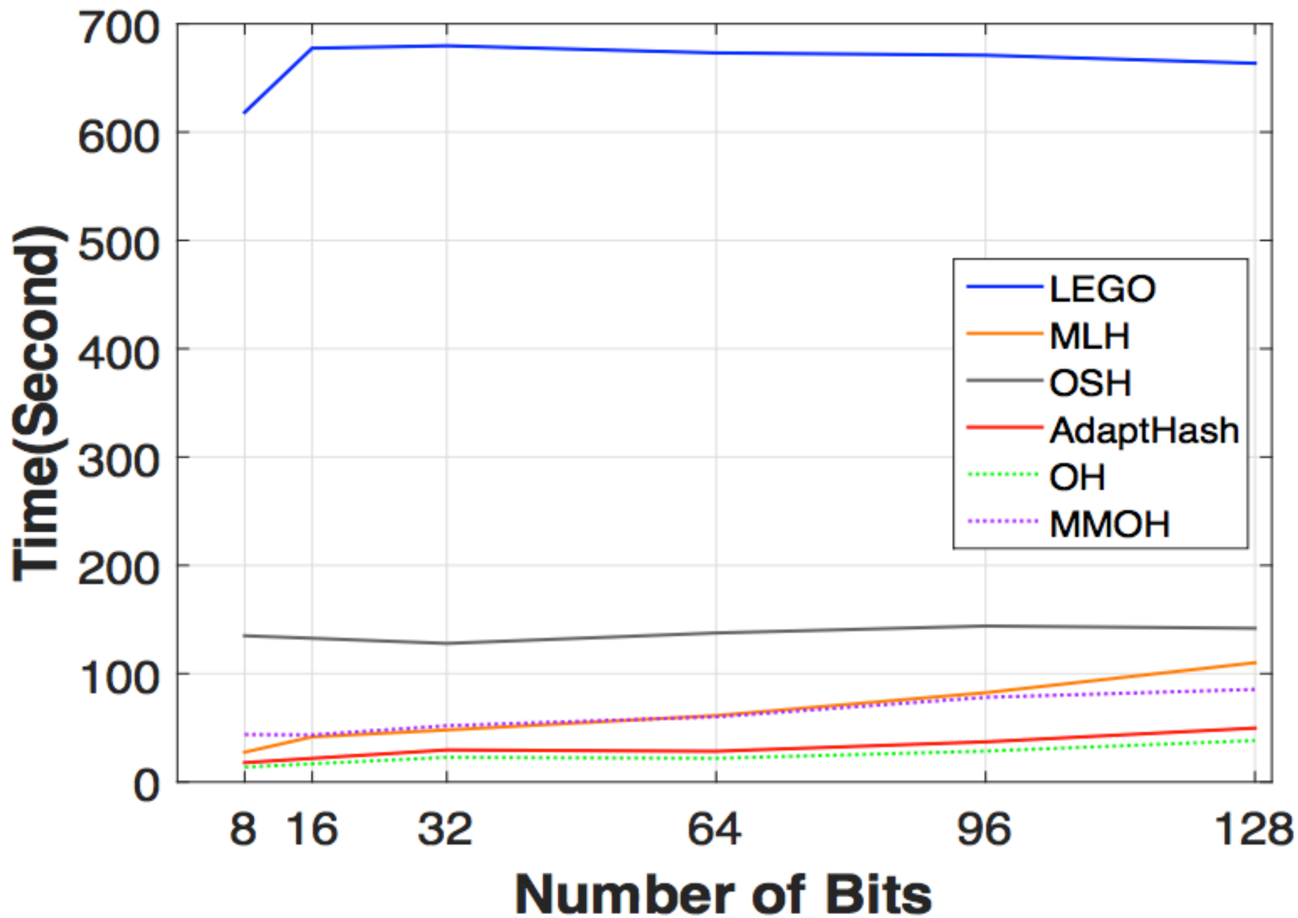
- Training time



(a) Training time comparison among different algorithms when the number of samples increases.

Evaluation

- Training time



(b) Training time comparison among different algorithms with different code lengths.

Learning Hashing with Multiple Features

Chenghao Zhang, and Wei-Shi Zheng. Semi-supervised Multi-view Discrete Hashing for Fast Image Search. IEEE Transactions on Image Processing, vol. 26, no. 6, pp. 2604-2617, 2017.

Semi-supervised Multi-Modality Hashing

$$\min_{\mathbf{B},\mathbf{C},\{\mathbf{W}_i\}_{i=1}^{K}} L = \frac{1}{n_\ell}||\mathbf{Y} - \mathbf{C}^T\mathbf{B}_\ell||^2$$

$$+ \frac{v}{n}||\mathbf{B} - \sum_{i=1}^{K}\mathbf{W}_i^T\phi_i(\mathbf{X}^{(i)})||^2$$

$$+ \frac{\alpha}{2}\sum_{i=1}^{K}\sum_{j\neq i}||\mathbf{W}_i^T Cov_{ij}\mathbf{W}_j||^2$$

$$+ \beta \cdot trace(\sum_{i=1}^{K}\sum_{j=1}^{K}\mathbf{W}_i^T\mathbf{L}_{ij}\mathbf{W}_j) + \theta||\mathbf{C}||^2$$

$$s.t. \quad \tilde{\mathbf{Y}} = \mathbf{C}^T\mathbf{B}_\ell, \ \mathbf{B} \in \{-1,1\}^{L*n}.$$

Modeling

- Multi-view Discrete Modeling

$$\mathbf{b} = sgn(\sum_{i=1}^{K} \mathbf{W}_i^T \phi_i(\mathbf{x}^{(i)})).$$

View-specific transformation

$$\min_{\mathbf{B},\{\mathbf{W}_i\}_{i=1}^{K}} L = ||\mathbf{B} - \sum_{i=1}^{K} \mathbf{W}_i^T \phi_i(\mathbf{X}^{(i)})||^2$$

Modeling

- Regression on Class Label Vectors

$$\tilde{\mathbf{y}} = G(\mathbf{b}) = \left[\mathbf{C}_1^T\mathbf{b}, \ldots, \mathbf{C}_L^T\mathbf{b}\right]^T$$

$$\min_{\mathbf{B}_\ell, \mathbf{C}} \frac{1}{n_\ell} ||\mathbf{Y} - \mathbf{C}^T\mathbf{B}_\ell||^2 + \theta||\mathbf{C}||^2$$

Modeling

- Extracting Statistically Uncorrelated View-specific Features

$$\min_{\{\mathbf{W}_i\}_{i=1}^{K}} \sum_{i=1}^{K} \sum_{j \neq i}^{K} ||\mathbf{W}_i^T Cov_{ij} \mathbf{W}_j||^2$$

$$\begin{aligned}
&\mathbb{E}_{\{(\mathbf{x}^{(i)},\mathbf{x}^{(j)}) \mid \mathbf{W}_i,\ \mathbf{W}_j\}} \left\{ \mathbf{W}_i^T (\mathbf{x}^{(i)} - \mathbb{E}_{\mathbf{x}^{(i)}})(\mathbf{x}^{(j)} - \mathbb{E}_{\mathbf{x}^{(j)}})^T \mathbf{W}_j \right\} \\
\approx & \frac{1}{n} \sum_{r=1}^{n} \mathbf{W}_i^T (\mathbf{x}_r^{(i)} - \mathbf{u}^{(i)})(\mathbf{x}_r^{(j)} - \mathbf{u}^{(j)})^T \mathbf{W}_j \\
= & \mathbf{W}_i^T \left\{ \frac{1}{n} \sum_{r=1}^{n} (\mathbf{x}_r^{(i)} - \mathbf{u}^{(i)})(\mathbf{x}_r^{(j)} - \mathbf{u}^{(j)})^T \right\} \mathbf{W}_j \\
= & \mathbf{W}_i^T Cov_{ij} \mathbf{W}_j,
\end{aligned}$$

Modeling

- Composite local data variation

$$\min_{\{\mathbf{W}_i\}_{i=1}^K} \quad trace(\sum_{i=1}^{K}\sum_{j=1}^{K}\mathbf{W}_i^T\mathbf{L}_{ij}\mathbf{W}_j).$$

$$\mathbf{L} = \frac{1}{n^2}\sum_{r=1}^{n}\sum_{t=1}^{n}S(r,t)(\sum_{i=1}^{K}\mathbf{W}_i^T(\phi_i(\mathbf{x}_r^{(i)}) - \phi_i(\mathbf{x}_t^{(i)}))) \times (\sum_{i=1}^{K}\mathbf{W}_i^T(\phi_i(\mathbf{x}_r^{(i)}) - \phi_i(\mathbf{x}_t^{(i)})))^T \qquad = \sum_{i=1}^{K}\sum_{j=1}^{K}\mathbf{W}_i^T\mathbf{L}_{ij}\mathbf{W}_j$$

$$\mathbf{L}_{ij} = \frac{2}{n^2}\phi_i(\mathbf{X}^{(i)})(\mathbf{D}-\mathbf{S})\phi_j(\mathbf{X}^{(j)})^T$$

$$S(r,t) = \frac{1}{K}\sum_{i=1}^{K}exp(-||\mathbf{x}_r^{(i)} - \mathbf{x}_t^{(i)}||^2/2\sigma_i^2)$$

Evaluation

- **Datasets**
  - WIKI: image (SIFT), text (LDA)
  - CIFAR-10: Gist, HOG features
  - NUS-WDIE: image (SIFT), text (tag vector)
  - ILSVRC-150K: 17 &18 layer features of VGG net
- **Compared Methods**
  - Unsupervised Multi-view Hash Models
  - Semi-supervised Hash Models
- **Measurement**
  - MAP: Mean Average Precision

Evaluation

Multiple Modality vs. Single

TABLE II
EVALUATION ON PROPOSED SSMDH: MAP ON CIFAR-10

	SSMDH	GIST	HOG	Naive Fusion
16 bits	**0.3919**	0.3632	0.3543	0.3575
32 bits	**0.4102**	0.3801	0.3892	0.3847
48 bits	**0.4349**	0.3905	0.3902	0.3921
64 bits	**0.4479**	0.3909	0.3924	0.3930
80 bits	**0.4574**	0.3980	0.3965	0.3992

TABLE IV
EVALUATION ON PROPOSED SSMDH: MAP ON NUS-WIDE

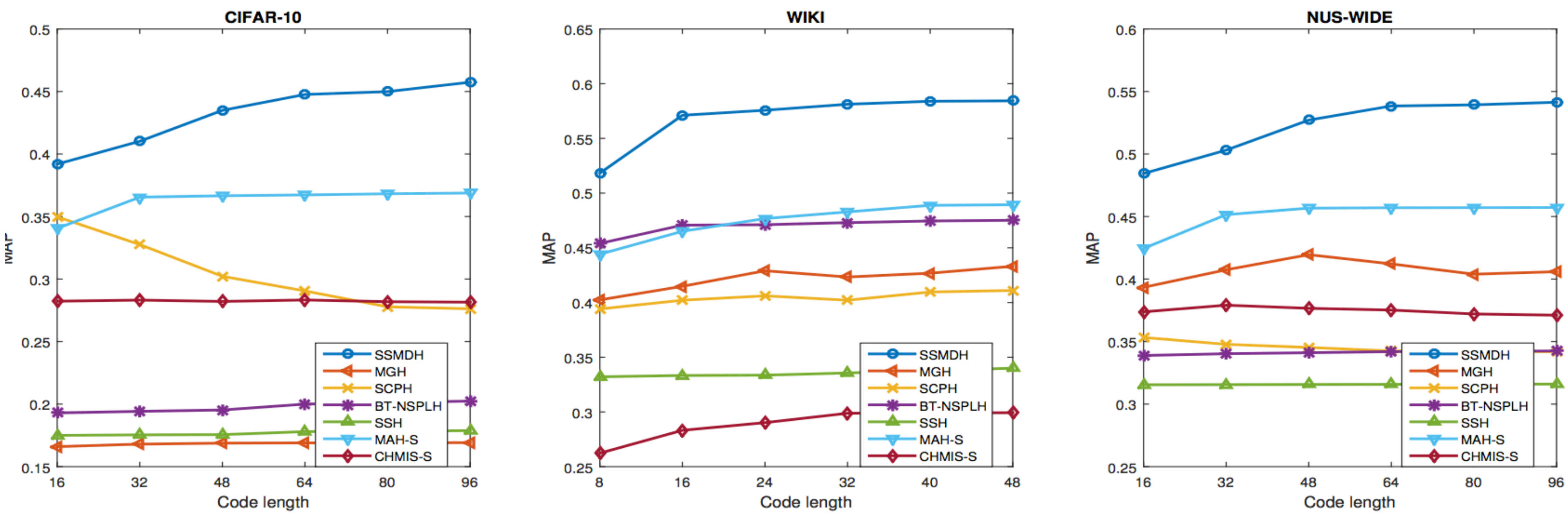
	SSMDH	IMG(SIFT)	TXT	Naive Fusion
16 bits	**0.4844**	0.4334	0.4726	0.4586
32 bits	**0.5031**	0.4547	0.4619	0.4644
48 bits	**0.5271**	0.4691	0.4664	0.4749
64 bits	**0.5384**	0.4752	0.4649	0.4723
80 bits	**0.5393**	0.4710	0.4683	0.4772

TABLE III
EVALUATION ON PROPOSED SSMDH: MAP ON WIKI

	SSMDH	IMG(SIFT)	TXT	Naive Fusion
8 bits	**0.5182**	0.2524	0.4273	0.3075
16 bits	**0.5711**	0.2623	0.5054	0.4495
24 bits	**0.5757**	0.2618	0.5281	0.4405
32 bits	**0.5812**	0.2621	0.5317	0.4503
40 bits	**0.5839**	0.2629	0.5319	0.4543

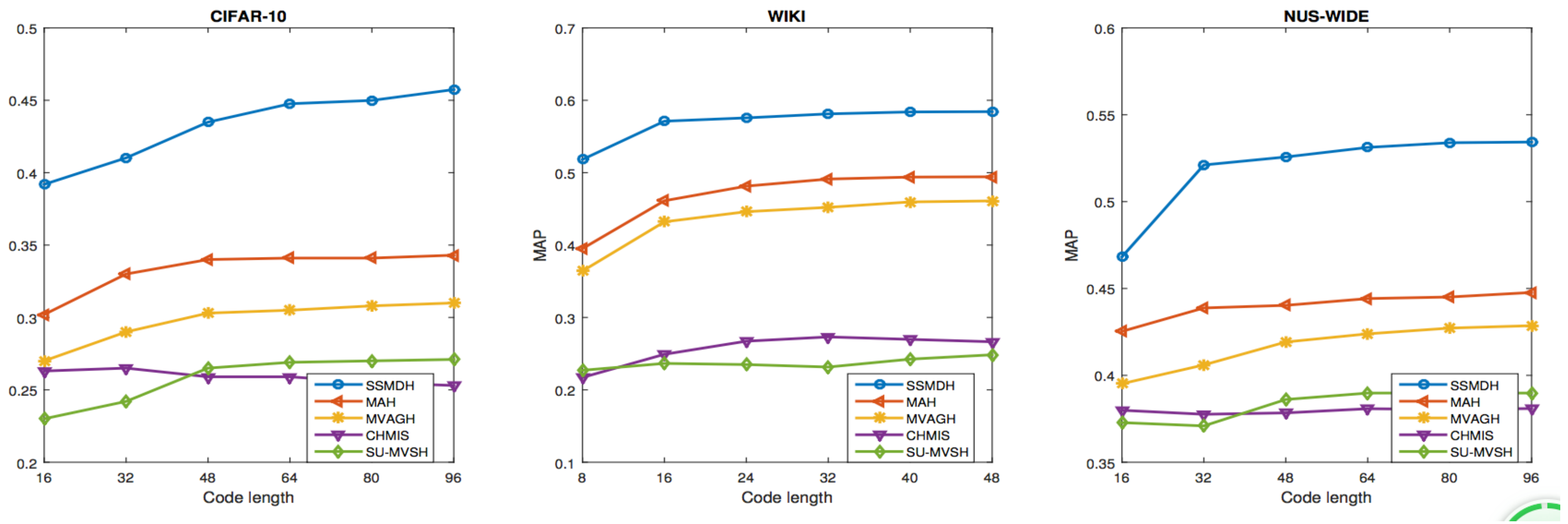
Evaluation

Comparison to Semi-supervised Models



Evaluation

- Comparison to Unsupervised Models



Evaluation

- When deep features are used?

TABLE V
MAP Evaluation on CIFAR-10(CNN+HOG+GIST)

Code Length	16	32	48	64	80	96
SSMDH	**0.5918**	**0.6040**	**0.6351**	**0.6461**	**0.6471**	**0.6483**
MAH	0.5303	0.5467	0.5820	0.5924	0.5943	0.5961
MVAGH	0.3421	0.3915	0.3991	0.4038	0.4091	0.4131
CHMIS	0.3044	0.3623	0.3870	0.3896	0.3904	0.3910
SU-MVSH	0.2758	0.2754	0.2743	0.2731	0.2720	0.2727
MGH	0.1744	0.1805	0.2141	0.2219	0.2240	0.2255
SCPH	0.3952	0.3849	0.3753	0.3431	0.2716	0.2701
BT-NSPLH	0.2083	0.2157	0.2320	0.2405	0.2423	0.2431
SSH	0.1939	0.2048	0.2257	0.2333	0.2361	0.2380
MAH-S	0.3520	0.4097	0.4148	0.4281	0.4342	0.4388
CHMIS-S	0.2838	0.2873	0.2893	0.2901	0.2909	0.2915

TABLE VI
MAP Evaluation on NUS-WIDE(CNN+SIFT+TXT)

Code Length	16	32	48	64	80	96
SSMDH	**0.6383**	**0.6671**	**0.6895**	**0.6987**	**0.7073**	**0.7129**
MAH	0.5845	0.6149	0.6382	0.6502	0.6614	0.6687
MVAGH	0.5649	0.5893	0.5990	0.6032	0.6067	0.6091
CHMIS	0.4458	0.4531	0.4562	0.4580	0.4588	0.4592
SU-MVSH	0.4281	0.4366	0.4398	0.4403	0.4389	0.4371
MGH	0.5431	0.5586	0.5548	0.5560	0.5563	0.5567
SCPH	0.5459	0.5340	0.5295	0.5248	0.5224	0.5209
BT-NSPLH	0.5146	0.5263	0.5279	0.5281	[illegible]	[illegible]
SSH	0.4739	0.4752	0.4756	0.4760	[illegible]	[illegible]
MAH-S	0.5714	0.5925	0.6003	0.6041	0.6068	0.6072
CHMIS-S	0.4618	0.4780	0.4825	0.4841	0.4862	0.4867

Evaluation

- On more diverse dataset

TABLE VII
MAP Evaluation on ILSVRC-150k

Code Length	32	64	96	128	160
SSMDH	**0.1120**	**0.2027**	**0.2418**	**0.2460**	**0.2571**
MAH	0.0897	0.1082	0.1124	0.1206	0.1258
MVAGH	0.0734	0.0859	0.0980	0.1064	0.1113
CHMIS	0.0342	0.0358	0.0366	0.0378	0.0375
SU-MVSH	0.0298	0.0303	0.0312	0.0316	0.0321
MGH	0.0699	0.0733	0.0758	0.0772	0.0778
SCPH	0.0645	0.0691	0.0739	0.0765	0.0773
BT-NSPLH	0.0378	0.0407	0.0454	0.0467	0.0479
SSH	0.0143	0.0159	0.0167	0.0171	0.0173
MAH-S	0.0972	0.1211	0.1259	0.1283	0.1301
CHMIS-S	0.0349	0.0368	0.0387	0.0392	0.0397

Evaluation

- When different classes are imbalanced?

TABLE IX
MAP EVALUATION ON IMBALANCED ILSVRC-150K

Code Length	32	64	96	128	160
SSMDH	**0.1017**	**0.1916**	**0.2439**	**0.2655**	**0.2710**
MAH	0.0901	0.1357	0.1422	0.1460	0.1473
MVAGH	0.0741	0.0949	0.1018	0.1040	0.1199
CHMIS	0.0287	0.0301	0.0325	0.0379	0.0404
SU-MVSH	0.0254	0.0331	0.0380	0.0392	0.0408
MGH	0.0534	0.0668	0.0722	0.0760	0.0752
SCPH	0.0514	0.0659	0.0723	0.0737	0.0745
BT-NSPLH	0.0328	0.0377	0.0391	0.0403	0.0410
SSH	0.0124	0.0129	0.0134	0.0138	0.0136
MAH-S	0.0954	0.1236	0.1259	0.1301	0.1306
CHMIS-S	0.0299	0.0307	0.0316	0.0331	0.0338

Evaluation

- Are the models transferable?

TABLE XII
MAP EVALUATION ON TRANSFER CASE

Database	CIFAR-10			ILSVRC-150K		
Code Length	32	64	96	32	64	96
SSMDH	**0.7200**	**0.7214**	**0.7227**	**0.4065**	**0.4144**	**0.4225**
MAH	0.6684	0.6732	0.6741	0.3253	0.3382	0.3402
MVAGH	0.6706	0.6683	0.6620	0.3447	0.3581	0.3639
CHMIS	0.4526	0.4569	0.4611	0.1841	0.1878	0.1890
SU-MVSH	0.4024	0.4110	0.4123	0.1394	0.1386	0.1381
MGH	0.4418	0.4450	0.4456	0.3062	0.3104	0.3149
SCPH	0.5315	0.5247	0.5174	0.2859	0.2900	0.2923
BT-NSPLH	0.4727	0.4739	0.4746	0.1455	0.1527	0.1540
SSH	0.4312	0.4427	0.4475	0.0942	0.0955	0.0959
MAH-S	0.6869	0.6904	0.6918	0.3438	0.3511	0.3537
CHMIS-S	0.4635	0.4658	0.4666	0.2003	0.2015	0.2019

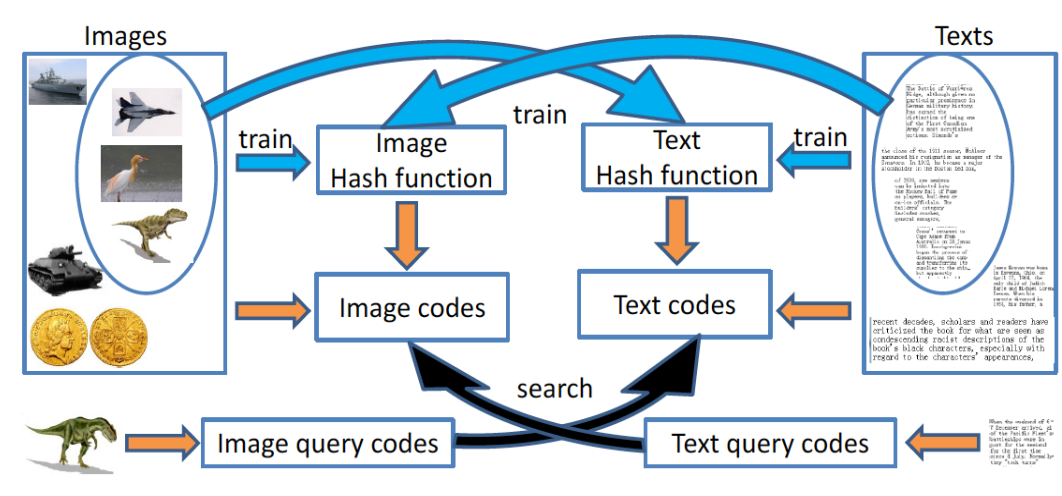
Evaluation

- Why not use classifiers to generate binary codes?

TABLE XI
MAP EVALUATION ON CLASSIFIER-BASED HASHING

Database	Method	Code Length	MAP
CIFAR-10	Classifier+ONE-HOT	4	0.4235
	Classifier+LSH	64	0.4542
	SSMDH	64	**0.6461**
WIKI	Classifier+ONE-HOT	4	**0.5938**
	Classifier+LSH	16	0.5816
	SSMDH	16	0.5711
NUS-WIDE	Classifier+ONE-HOT	4	0.6570
	Classifier+LSH	64	0.6394
	SSMDH	64	**0.6987**
ILSVRC-150K	Classifier+ONE-HOT	10	0.0536
	Classifier+LSH	64	0.0482
	SSMDH	64	**0.2027**

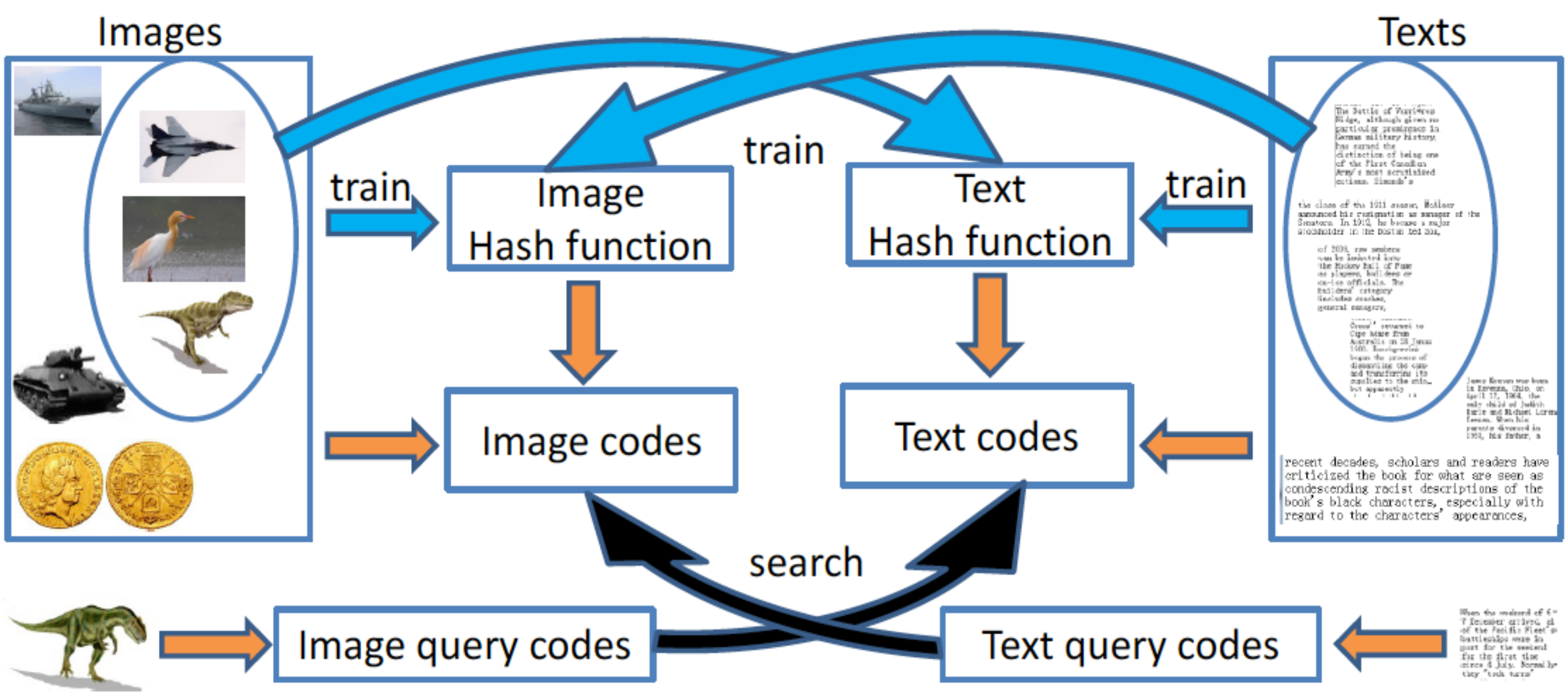
Hashing for Cross-Modalities



Botong Wu, Qiang Yang, Wei-Shi Zheng*, Yizhou Wang, and Jingdong Wang. Quantized Correlation Hashing for Fast Cross-modal Search. In International Joint Conference on Artificial Intelligence (IJCAI), 2015.

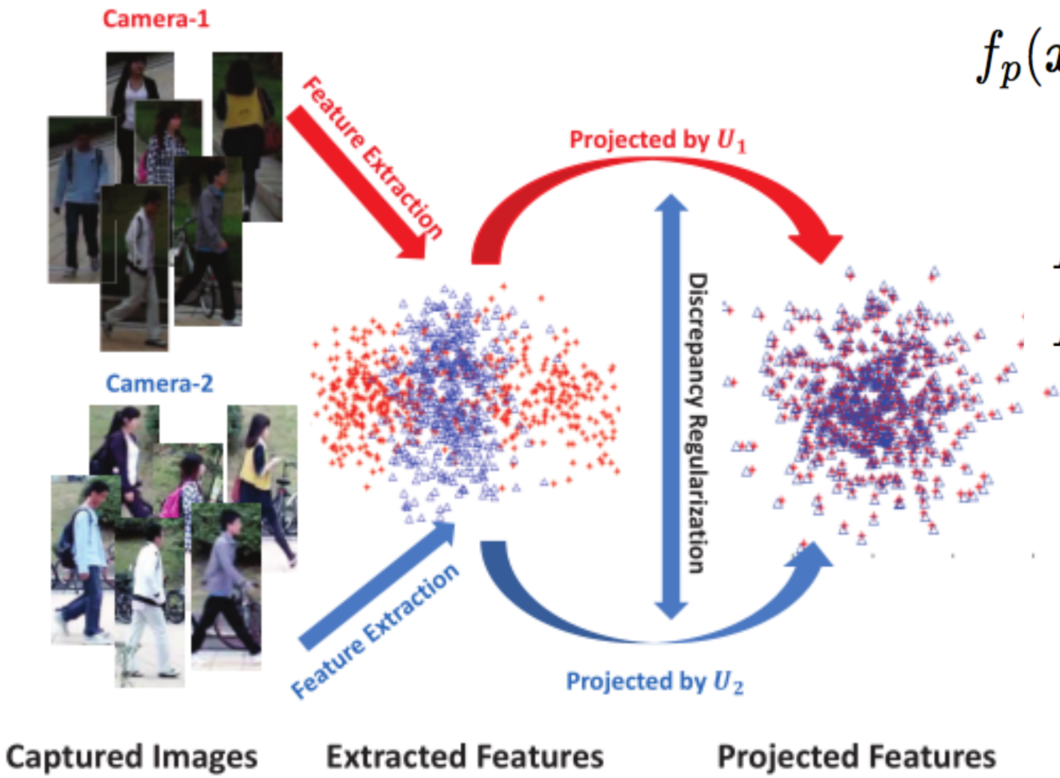
Xiatian Zhu, Botong Wu, Dongcheng Huang, Wei-Shi Zheng*. Fast Open-World Person Re-Identification. Submitted to IEEE Transactions on Image Processing. (Minor Revision)

Cross-modality Hashing



Cross-modality Hashing

- How is it related to person re-id



$$f_p(\boldsymbol{x}_i^p) = \boldsymbol{x}_i^p \boldsymbol{W}_p, \qquad f_g(\boldsymbol{x}_j^g) = \boldsymbol{x}_j^g \boldsymbol{W}_g$$

$$\boldsymbol{B}_p = \text{sign}(\boldsymbol{X}_p \boldsymbol{W}_p) \in \{-1, 1\}^{n_p \times c},$$
$$\boldsymbol{B}_g = \text{sign}(\boldsymbol{X}_g \boldsymbol{W}_g) \in \{-1, 1\}^{n_g \times c},$$

Evaluation

- **Datasets**
  - CUHK03
  - SYSU
  - Market

- **Setting: open-world re-id**
  - Re-id with large scale of imposters

Wei-Shi Zheng, Shaogang Gong, and Tao Xiang. Towards Open-World Person Re-Identification by One-Shot Group-based Verification. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 38, no. 3, pp. 591-606, 2016.

Evaluation

Comparison to other related Hashing functions

TABLE III: Comparing state-of-the-art hashing methods. (Metrics: TTR (%) at varying FTRs (%), and mAP (%).)

Dataset	CUHK03 [39]						SYSU [36]						Market-1501 [40]					
Metric	Individual Verification					mAP	Individual Verification					mAP	Individual Verification					mAP
	1%	5%	10%	20%	30%	(%)	1%	5%	10%	20%	30%	(%)	1%	5%	10%	20%	30%	(%)
LSH [25]	15.03	34.87	48.12	64.66	75.20	1.91	21.21	43.73	57.42	72.17	81.17	5.48	37.00	61.56	73.43	85.12	90.92	8.28
SH [27]	11.97	27.8	39.99	54.73	65.35	1.49	17.60	34.21	45.63	60.14	70.27	4.18	38.54	58.69	69.47	80.71	86.75	9.29
SGH [65]	16.95	37.37	50.71	66.45	76.76	2.36	27.18	49.21	61.74	75.34	82.92	8.03	37.75	63.16	75.05	86.13	91.76	8.69
ITQ [66]	17.31	39.29	53.06	69.12	80.24	2.70	26.51	49.55	63.18	77.04	84.87	7.47	40.72	67.17	78.67	88.51	93.41	10.84
CCA+ITQ [66]	28.11	51.15	65.05	78.95	86.37	4.28	50.50	73.75	83.16	91.08	94.67	18.12	57.75	80.30	87.53	93.47	96.17	15.02
KSH [30]	32.29	57.54	69.78	81.73	88.96	5.49	53.23	77.28	85.88	92.62	95.62	22.29	59.03	81.83	89.01	94.26	96.41	17.34
FH [68]	20.01	40.07	52.32	67.35	77.39	1.03	29.48	50.56	62.42	75.58	83.65	8.07	28.24	48.88	60.59	73.92	81.61	5.07
SDH [31]	38.80	66.82	78.83	88.15	93.03	7.31	46.09	72.34	82.76	90.75	94.51	17.99	58.03	81.26	88.22	94.16	96.29	15.57
COSDISH [69]	13.19	29.18	40.33	56.88	68.23	1.55	38.04	61.43	72.73	83.95	89.38	11.51	39.29	62.26	73.49	83.68	89.04	8.44
CMSSH [48]	10.46	32.46	49.80	68.67	80.45	1.25	11.06	33.76	50.51	70.55	82.29	3.18	8.88	29.25	46.49	67.02	79.56	1.55
CVH [47]	2.83	10.09	17.81	31.05	42.51	0.39	5.76	19.67	31.77	49.33	62.21	1.30	3.51	13.38	22.62	37.31	50.55	0.53
CMFH [71]	11.85	31.23	46.40	64.63	75.56	1.27	25.73	54.24	68.73	82.09	89.14	6.32	24.96	52.96	67.52	81.62	89.11	4.07
SCM [51]	5.43	17.84	28.77	44.72	58.70	0.59	14.83	32.93	45.22	60.35	70.95	3.92	13.41	31.49	43.44	58.75	69.09	2.04
SePH [70]	26.98	52.69	65.88	79.29	86.24	4.18	37.15	64.01	75.75	86.09	91.56	13.56	41.88	70.39	80.72	88.89	93.09	8.80
X-ICE(hinge)	49.67	**79.60**	**89.50**	**96.09**	**98.48**	**11.66**	61.86	84.10	91.47	**96.35**	**98.26**	**29.93**	**66.52**	**88.03**	**93.66**	**97.15**	98.55	**21.47**
X-ICE(reg)	**49.96**	78.18	88.96	95.88	97.98	11.23	**63.13**	**84.86**	**91.52**	96.17	98.08	29.44	64.18	86.98	92.91	97.09	**98.59**	20.68
Metric	Set Verification					mAP	Set Verification					mAP	Set Verification					mAP
LSH [25]	4.81	13.97	21.83	35.01	45.88	1.91	7.25	18.13	27.35	41.18	52.31	5.48	17.17	33.22	43.85	57.75	67.36	8.28
SH [27]	3.91	12.00	19.93	31.72	43.41	1.49	7.22	16.36	24.39	36.62	46.96	4.18	21.52	36.77	46.41	58.33	68.00	9.29
SGH [65]	5.42	14.34	22.87	36.84	48.89	2.36	10.20	22.06	31.91	46.08	56.88	8.03	16.59	34.04	45.33	59.12	69.60	8.69
ITQ [66]	5.45	14.9	23.96	37.95	49.51	2.70	8.81	21.47	31.42	45.82	56.73	7.47	17.39	35.66	47.25	60.90	70.46	10.84
CCA+ITQ [66]	10.52	23.86	34.05	49.95	59.58	4.28	21.08	42.10	53.63	67.05	75.76	18.12	18.30	45.05	60.16	73.75	81.38	15.02
KSH [30]	11.65	27.43	38.19	52.68	63.62	5.49	22.07	43.13	55.44	69.26	77.25	22.29	21.74	47.36	61.28	74.78	82.22	17.34
FH [68]	7.66	17.92	26.82	40.15	52.08	1.03	12.82	26.19	35.86	48.86	59.28	8.07	13.32	27.17	36.78	50.13	60.84	5.07
SDH [31]	13.59	31.37	43.59	60.08	70.34	7.31	15.46	36.48	49.00	63.30	72.79	17.99	20.95	47.01	61.17	75.19	82.16	15.57
COSDISH [69]	5.07	13.37	21.71	34.27	43.74	1.55	16.86	33.10	43.36	57.20	66.38	11.51	15.80	35.00	46.81	60.49	69.66	8.44
CMSSH [48]	2.32	10.32	19.42	32.44	45.06	1.25	2.66	11.13	19.29	33.87	46.11	3.18	2.20	8.93	17.42	31.57	44.29	1.55
CVH [47]	1.31	5.62	12.06	23.29	32.45	0.39	1.75	7.65	14.44	27.18	38.12	1.30	1.57	6.83	12.83	24.39	35.74	0.53
CMFH [71]	3.49	11.91	20.06	34.46	45.35	1.27	7.25	21.42	32.99	48.97	60.98	6.32	7.95	22.60	33.55	48.81	60.49	4.07
SCM [51]	1.91	7.64	14.74	27.22	37.74	0.59	5.75	15.55	23.82	37.04	48.23	3.92	5.29	15.23	23.20	36.53	47.93	2.04
SePH [70]	9.64	23.22	33.51	49.79	60.94	4.18	12.40	30.40	42.85	57.81	67.73	13.56	11.78	32.98	46.91	63.28	73.64	8.80
X-ICE(hinge)	**16.41**	**37.50**	**50.14**	**66.56**	**77.30**	**11.66**	23.32	46.84	60.48	74.20	82.37	**29.93**	**26.81**	**52.73**	**66.47**	**79.66**	**86.16**	**21.47**
X-ICE(reg)	16.37	37.36	49.71	65.49	76.03	11.23	**25.94**	**49.94**	**62.59**	**75.91**	**83.30**	29.44	22.27	48.12	62.87	77.13	84.55	20.68

Evaluation

- Comparison to re-id methods

TABLE V: Comparing state-of-the-art non-hashing person re-id methods. (Metrics: TTR (%) at varying FTRs (%), and mAP (%); ST: *Search Time* (smaller is better), with unit set as the search time of X-ICE.)

	Dataset	CUHK03 [39]							SYSU [36]							Market-1501 [40]						
	Metric	Individual Verification					mAP	ST	Individual Verification					mAP	ST	Individual Verification					mAP	ST
		1%	5%	10%	20%	30%	(%)	-	1%	5%	10%	20%	30%	(%)	-	1%	5%	10%	20%	30%	(%)	-
	X-ICE(hinge)	49.67	**79.60**	**89.50**	**96.09**	**98.48**	**11.66**	**1**	61.86	84.10	91.47	**96.35**	**98.26**	**29.93**	**1**	**66.52**	**88.03**	**93.66**	**97.15**	98.55	**21.47**	**1**
	X-ICE(reg)	**49.96**	78.18	88.96	95.88	97.98	11.23	**1**	**63.13**	**84.86**	**91.52**	96.17	98.08	29.44	**1**	64.18	86.98	92.91	97.09	**98.59**	20.68	**1**
non-hash	KISSME [34]	33.66	61.67	74.69	86.14	91.29	7.97	954	33.67	58.30	70.79	83.57	90.08	19.31	2056	59.40	81.55	89.19	94.87	96.84	21.64	1447
non-hash	CVDCA [36]	41.73	68.65	80.63	89.91	94.52	8.50	367	58.39	81.89	89.75	94.94	97.09	23.88	486	32.87	52.90	63.15	74.81	81.45	11.25	456
non-hash	XQDA [6]	56.96	82.67	91.71	97.04	98.29	15.93	1775	**59.93**	**83.31**	**91.04**	**96.32**	**98.09**	**34.28**	4311	71.40	90.00	94.70	97.98	99.04	28.95	3185
non-hash	MLAPG [11]	53.97	83.05	**92.35**	**97.56**	**99.02**	12.79	**128**	55.61	79.55	88.17	94.77	97.54	29.13	**159**	66.71	87.93	94.04	97.73	98.95	22.30	**154**
non-hash	DNS [12]	**59.68**	**84.68**	92.25	97.05	98.50	**17.52**	316	59.62	82.12	89.52	95.10	97.37	29.29	164	**75.33**	**91.55**	**95.83**	**98.26**	**99.13**	**31.48**	262
	Metric	Set Verification					mAP	ST	Set Verification					mAP	ST	Set Verification					mAP	ST
	X-ICE(hinge)	**16.41**	**37.50**	**50.14**	**66.56**	**77.30**	**11.66**	**1**	23.32	46.84	60.48	74.20	82.37	**29.93**	**1**	**26.81**	**52.73**	**66.47**	**79.66**	**86.16**	**21.47**	**1**
	X-ICE(reg)	16.37	37.36	49.71	65.49	76.03	11.23	**1**	**25.94**	**49.94**	**62.59**	**75.91**	**83.30**	29.44	**1**	22.27	48.12	62.87	77.13	84.55	20.68	**1**
non-hash	KISSME [34]	9.00	23.77	34.78	50.64	62.29	7.97	954	12.35	25.26	35.66	50.95	62.16	19.31	2056	31.42	50.39	61.30	73.88	81.37	21.64	1447
non-hash	CVDCA [36]	15.27	33.04	44.94	60.61	70.89	8.50	367	21.55	**46.22**	**59.37**	**73.29**	81.31	23.88	486	15.55	31.37	41.00	53.67	63.02	11.25	456
non-hash	XQDA [6]	15.43	38.83	54.55	71.38	80.87	15.93	1775	**23.33**	45.07	58.14	73.04	**82.23**	**34.28**	4311	34.18	58.71	70.53	82.21	88.50	28.95	3185
non-hash	MLAPG [11]	13.86	38.19	54.15	71.54	82.00	12.79	**128**	22.00	43.00	56.29	71.31	80.06	29.13	**159**	30.79	55.14	68.69	81.78	88.62	22.30	**154**
non-hash	DNS [12]	**18.91**	**41.09**	**54.86**	**71.88**	**81.65**	**17.52**	316	20.42	39.01	50.95	65.82	75.57	29.29	164	**41.34**	**62.74**	**73.12**	**83.63**	**89.42**	**31.48**	262

Evaluation

- When using more powerful features?

TABLE IX: Evaluating the effect of different visual features. (Metrics: TTR (%) at FTR = 1%, and mAP (%). IV: Individual Verification, SV: Set Verification.)

Method	Feature	CUHK03 [39]			SYSU [36]			Market-1501 [40]		
		IV	SV	mAP	IV	SV	mAP	IV	SV	mAP
DCNN [74]	Deep	47.87	14.38	13.62	58.77	19.64	29.69	78.37	31.58	33.65
KSH [30]	LOMO	32.29	11.65	5.49	53.23	22.07	22.29	59.03	21.74	17.34
	Deep	51.08	**19.00**	14.77	60.38	21.55	31.54	79.50	34.39	34.96
SDH [31]	LOMO	38.80	13.59	7.31	46.09	15.46	17.99	58.03	20.95	15.57
	Deep	36.88	15.93	9.11	52.93	16.60	22.23	71.00	35.61	26.73
X-ICE(hinge)	LOMO	49.67	16.41	11.66	61.86	23.32	29.93	66.52	26.81	21.47
	Deep	51.79	18.29	**15.33**	63.08	23.35	32.95	**80.90**	**41.52**	**37.34**
X-ICE(reg)	LOMO	49.96	16.37	11.23	63.13	25.94	29.44	64.18	22.27	20.68
	Deep	**52.62**	17.63	15.21	**64.05**	**24.57**	**33.07**	80.37	41.01	37.23

Evaluation

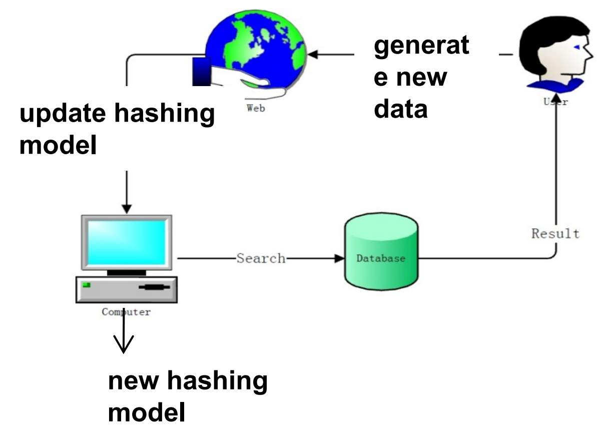
■ Visual Comparison

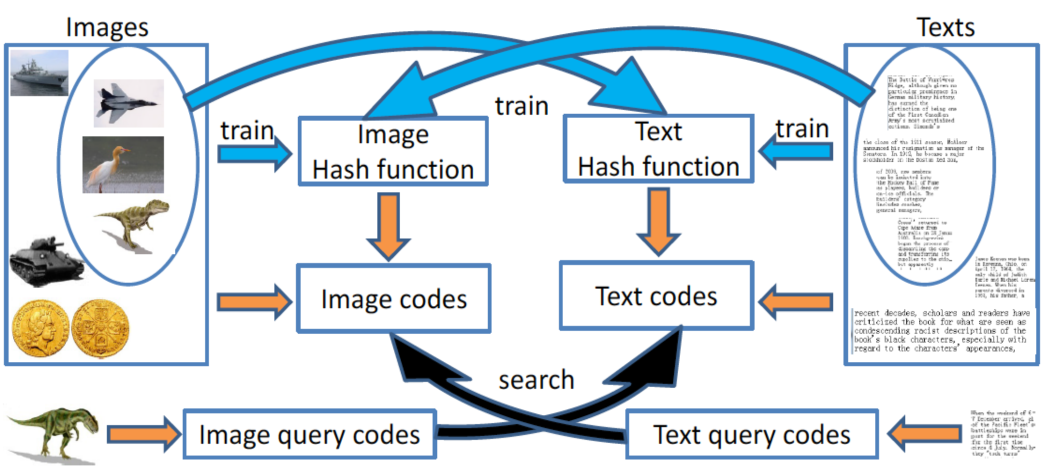
(a) CUHK03 [39] (b) SYSU [36] (c) Market-1501 [40]

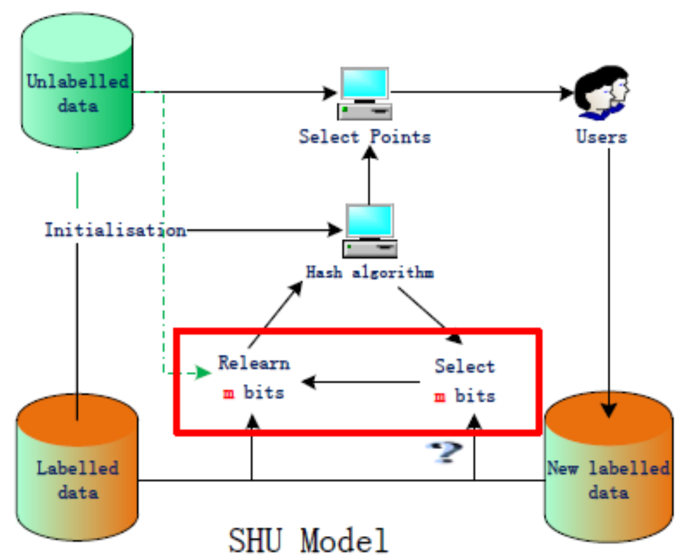
Fig. 5: Visualising person re-id performance by four top methods X-ICE (1st row), KSH (2nd row), CCA+ITQ (3rd row) and SDH (4th row). For each dataset, the left-most image is the probe person image, followed by top 10 most matched gallery images by respective methods with red boxes indicating true matches.

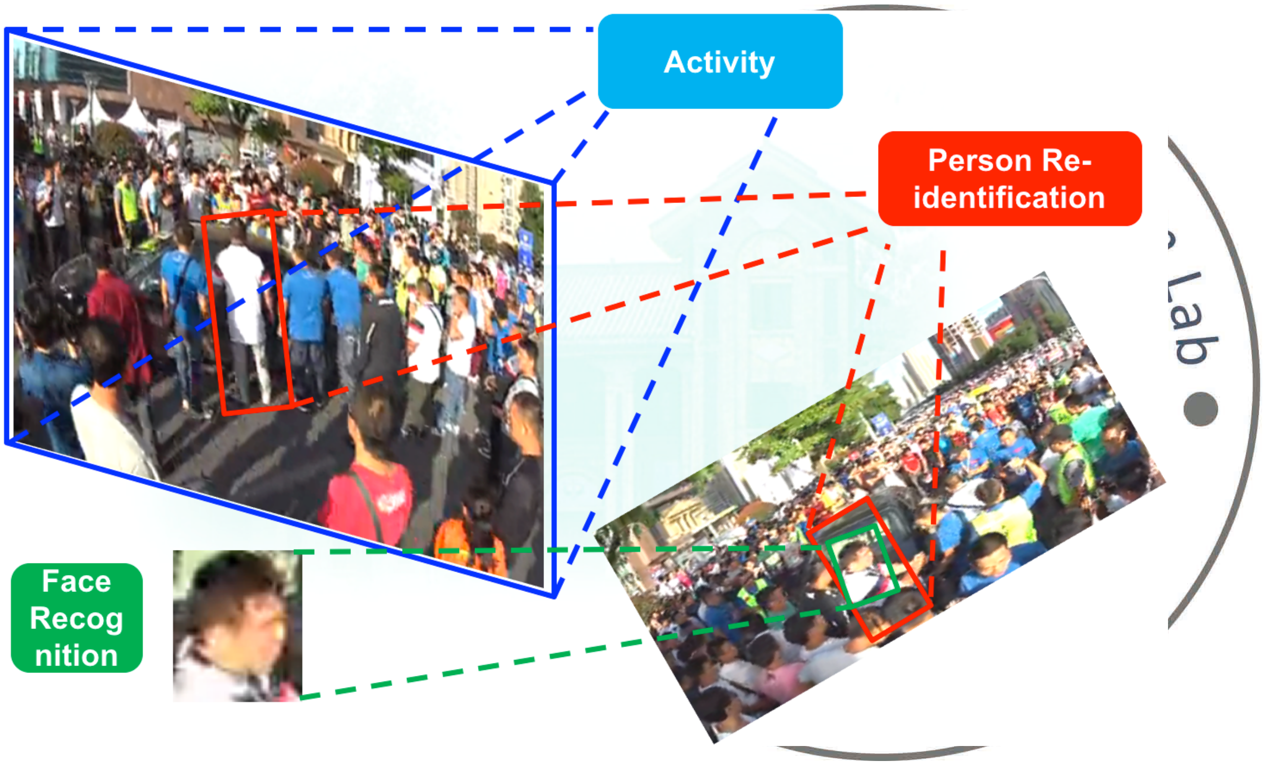
Large-scale Search

- **Fast Search based on Hashing**
  - Online Hashing
    (IEEE TNNLS 2017, IJCAI 2013)
  - Smart Model for Active Hashing
    (IJCAI 2013)
  - Multi/Cross Modality Search
    (IEEE TIP 2017; IJCAI 2015)









感谢各位老师和同学！