Label Consistent Matrix Factorization based Hashing for Cross-Modal Retrieval

Devraj Mandal & Soma Biswas
{devraj89,soma.biswas}@ee.iisc.ernet.in
Indian Institute of Science

CROSS-MODAL HASHING

A matrix factorization based hashing technique for cross-modal data

- ❑ Single label or Multi label
- ❑ Supervised algorithm
- ❑ Capability of handling large amounts of data
- ❑ Online settings

`shadow' `man' `ground' `swing' `black' `night'

`minicooper' `car' `automobile' `hotdog' `street'

`sky' `reflection' `building' `home' `windows'

PROBLEM STATEMENT

Data $X_t \in R^{d_t \times N}$ Labels $X_L \in R^{d_c \times N}$

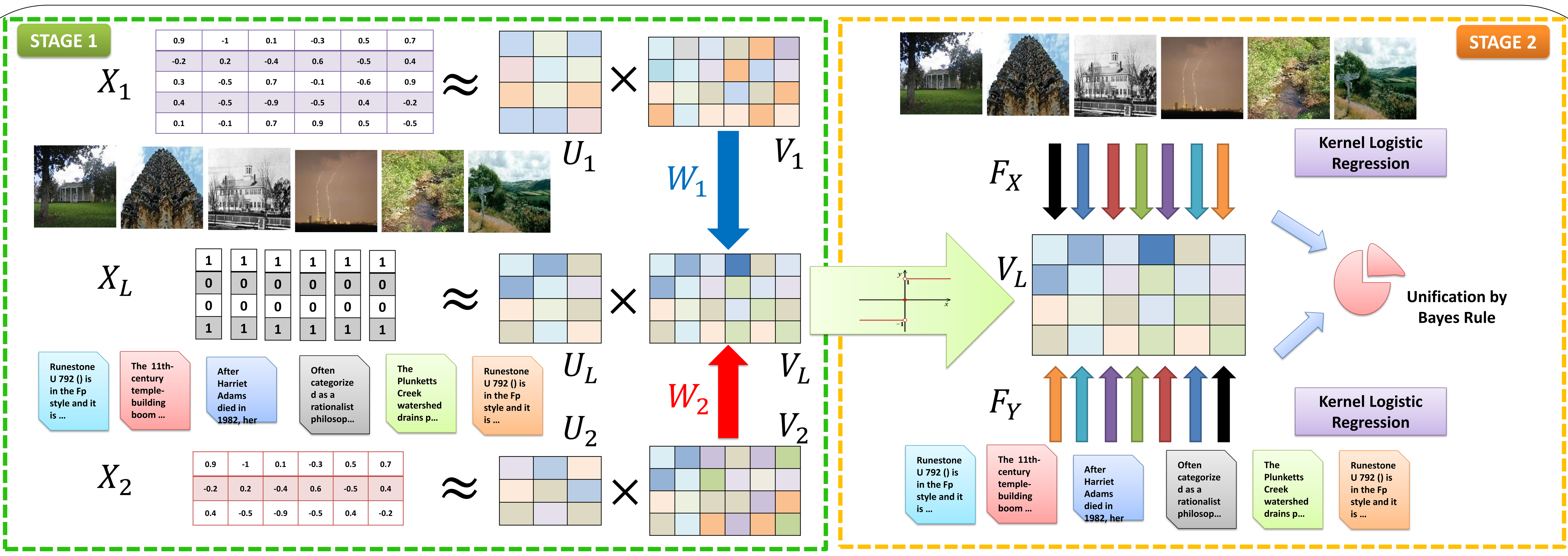
$$F = \sum_{t=1}^{2} \lambda_t \|X_t - U_t V_t\|_F^2 + \lambda_L \|X_L - U_L V_L\|_F^2 + \sum_{t=1}^{2} \alpha_t \|V_L - W_t V_t\|_F^2$$

Algorithm: For each iteration

- Fix V, W Update U using $\nabla_U F$
- Fix V, U Update W using $\nabla_W F$
- Fix U, W Update V using $\nabla_V F$

Closed Form updates

PROPOSED APPROACH



TRAINING WITH LARGE AMOUNTS OF DATA

A trick to handle large amounts of data

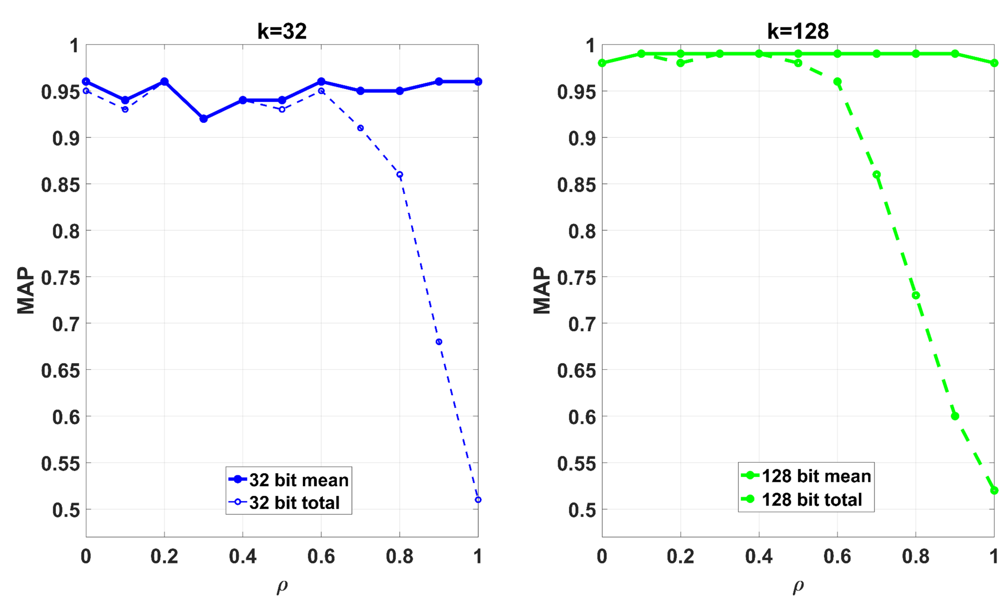
- ❑ No need for re-generation if training set = retrieval set
- ❑ An iterative scheme to handle data in mini-batches

For each mini-batch $i+1$

- Get $X_t^{(i+1)}$ and $X_L^{(i+1)}$ ← the new batch
- Get V^i, U^i, W^i ← from the ith batch
- Update $U^{(i+1)} = (1-\rho)U^i + \rho \nabla_U F^{(i+1)}$
- Update $W^{(i+1)} = (1-\rho)W^i + \rho \nabla_W F^{(i+1)}$
- Finally Compute $V^{(i+1)}$

ANALYSIS OF THE ALGORITHM

Effect of ρ on the hash code learning framework for the NUS-WIDE dataset



Complexity of Algorithm : $O(d_t N k)$

EXPERIMENTAL RESULTS

MAP@50 on the MIRFLICKR dataset

Method	I-T		T-I	
	K=32	K=64	K=32	K=64
CMSSH	0.66	0.66	0.64	0.63
CVH	0.63	0.62	0.65	0.64
MLBE	0.58	0.58	0.61	0.65
QCH	0.57	0.56	0.60	0.57
LSSH	0.64	0.64	0.67	0.69
CMFH	0.60	0.63	0.60	0.64
CMCQ	0.67	0.67	0.73	0.73
SePH	0.73	0.75	0.73	0.74
Ours_1	0.68	0.69	0.74	0.73
Ours_2	0.68	0.85	0.86	0.86

MAP@50 on the NUS-WIDE dataset

Method	I-T		T-I	
	K=32	K=64	K=32	K=64
CMSSH	0.52	0.52	0.42	0.41
CVH	0.52	0.50	0.54	0.51
MLBE	0.45	0.47	0.48	0.50
QCH	0.52	0.52	0.51	0.50
LSSH	0.55	0.56	0.66	0.68
CMFH	0.48	0.51	0.56	0.58
CMCQ	0.59	0.59	0.70	0.71
SePH	0.58	0.60	0.72	0.74
Ours_1	0.59	0.59	0.75	0.76
Ours_2	0.74	0.76	0.85	0.84

FUTURE WORK

Possible directions of future work

- ❑ Make the latent factors same i.e., $V_1 = V_2 = V$ and $W_1 = W_2 = W$
- ❑ Make the dictionary variables similar i.e., $U_1 = U_2 = U$ (to handle dimensionality mismatch, use CCA, GMA, random projections, etc.)
- ❑ Learning hash functions using non-linearity (end-to-end)
- ❑ A Online adaptation of the hash functions to the incoming data.

Supported by IEEE Signal Processing Society Travel Grant