

Predictable Dual-view Hashing

MOHAMMAD RASTEGARI

JONGHYUN CHOI

SHOBEIR FAKHRAEI*

HAL DAUMÉ III

LARRY S. DAVIS

University of Maryland, College Park, MD, USA



Integrate information from different modalities

- Text - Image



- Sound - Text

Dual-View Hashing

Textual Space

Human riding a horse

Person laughing
at restaurant

Cow standing in a farm

Visual Space

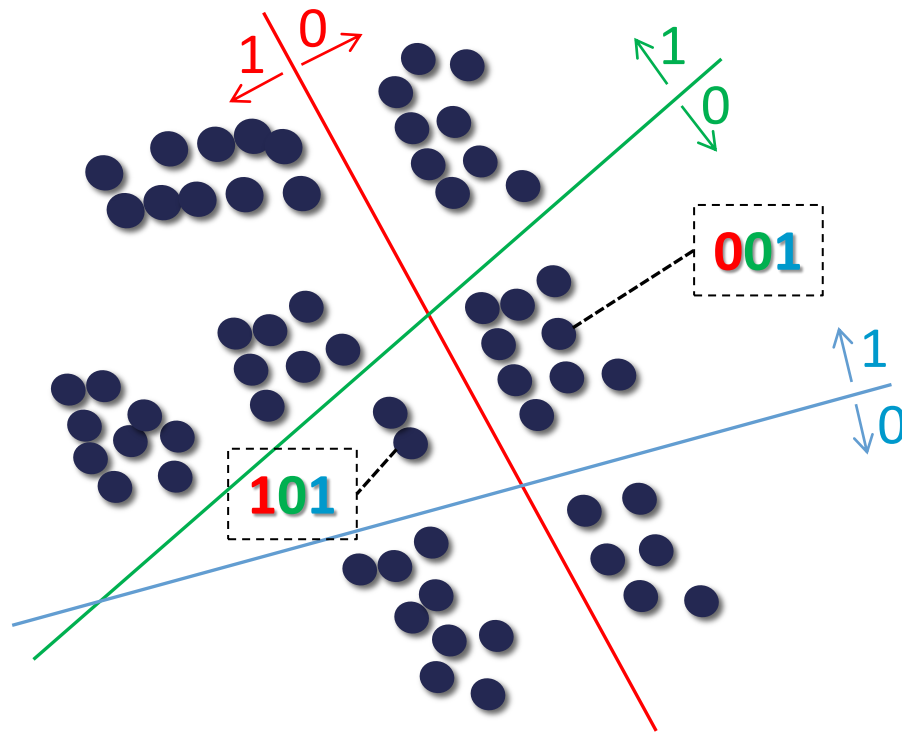


Hamming Space

010111010
101010101
110010101

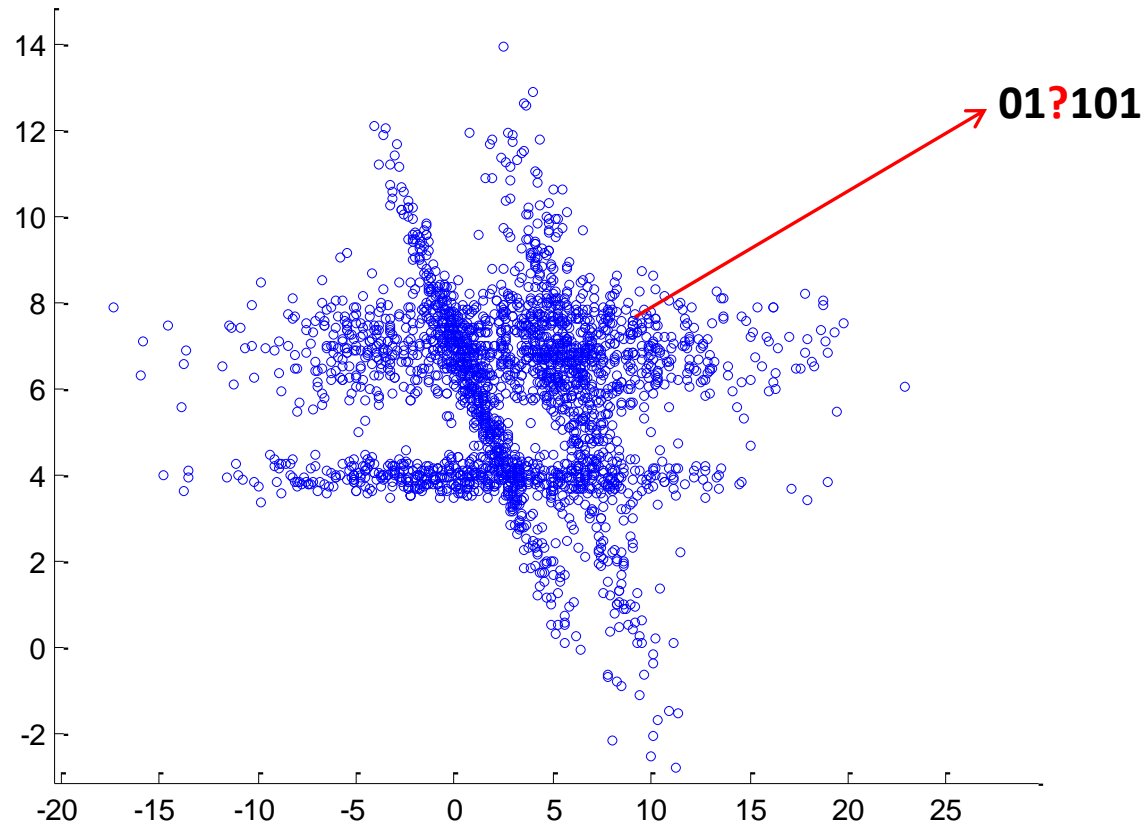


Binary code assignments with splits

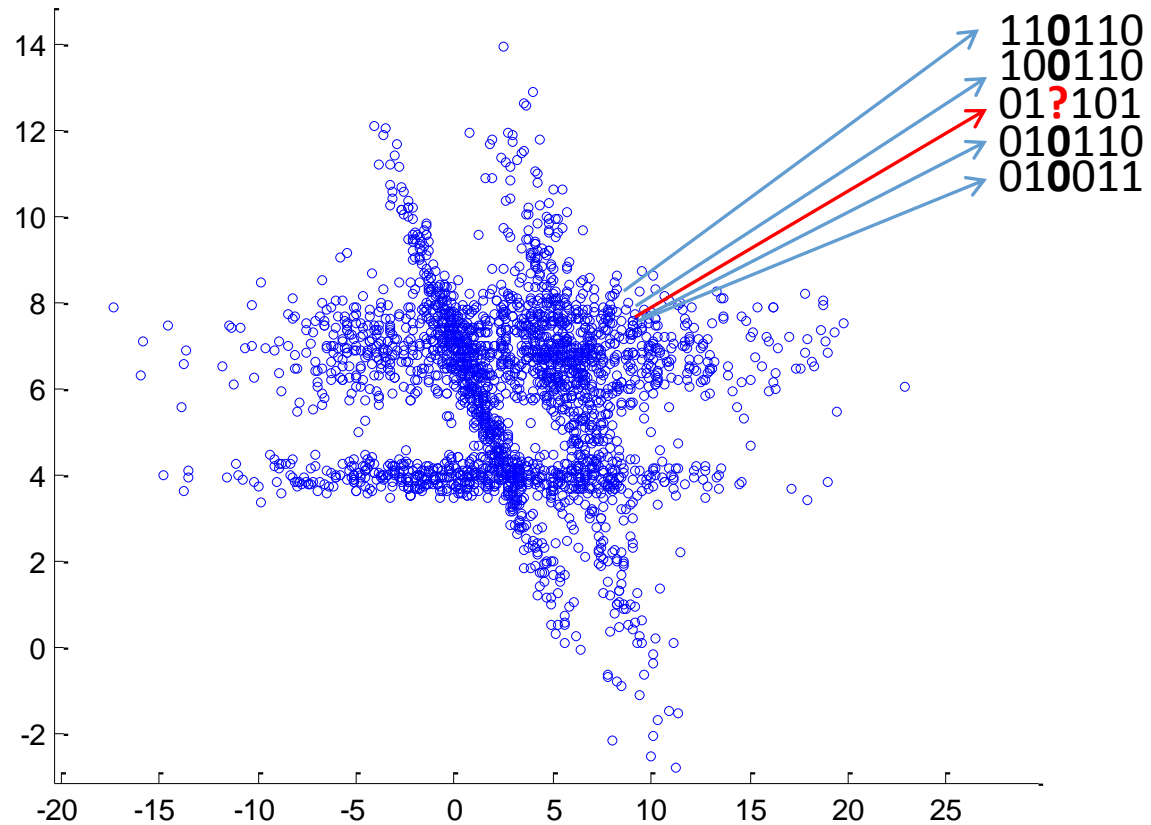


$$B = \text{sgn}(W^T X)$$

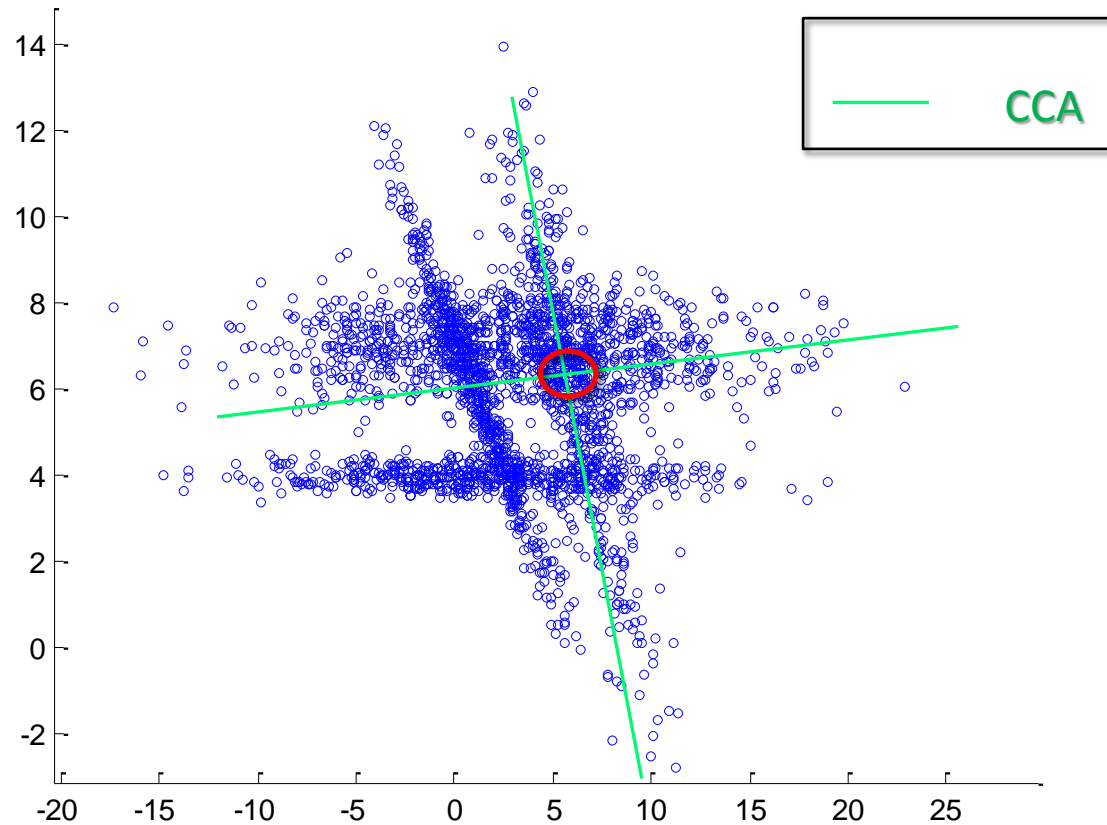
Predictability



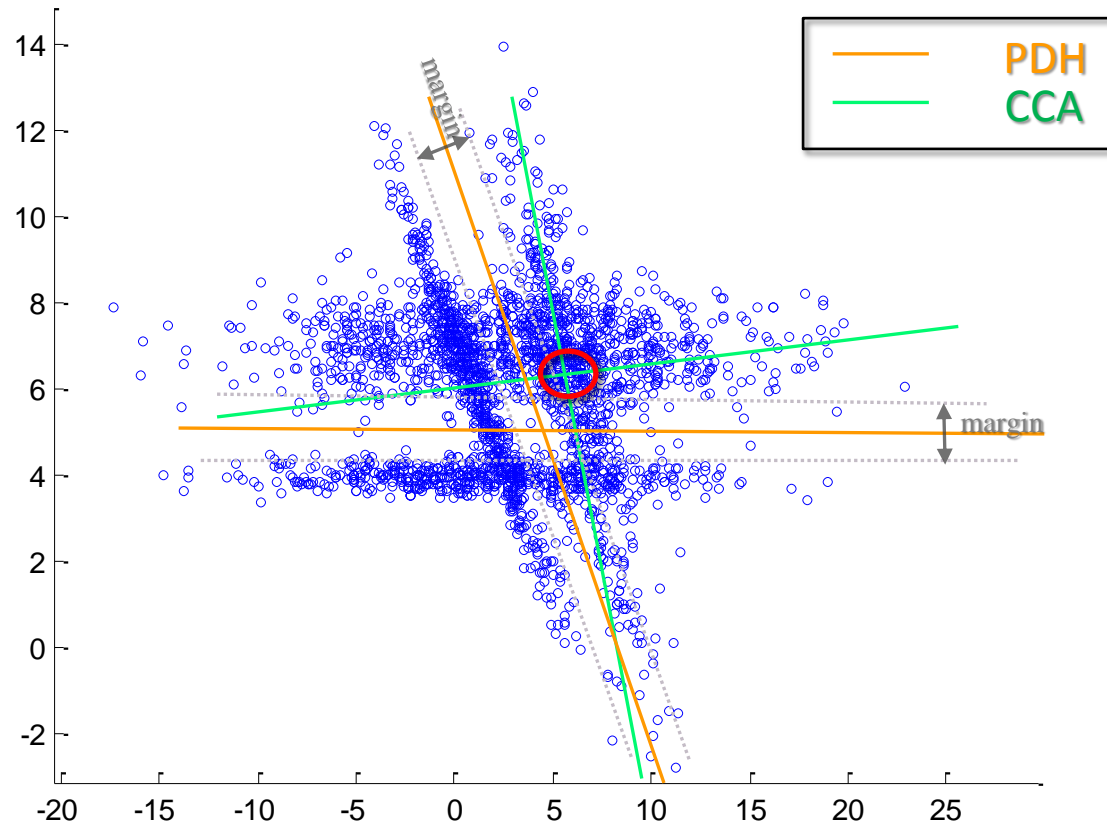
Predictability



Predictability



Predictability



Optimization

$$\min_{W_{\mathcal{V}}, W_{\mathcal{T}}, \xi_{\mathcal{V}}, \xi_{\mathcal{T}}} \|B_{\mathcal{T}} B_{\mathcal{T}}^T - I\|_2^2 + \|B_{\mathcal{V}} B_{\mathcal{V}}^T - I\|_2^2 \\ + \sum \|w_{\mathcal{V}i}\| + \sum \|w_{\mathcal{T}i}\| + C_1 \sum \xi_{\mathcal{V}} + C_2 \sum \xi_{\mathcal{T}}$$

s.t.

$$B_{\mathcal{T}} = \text{sgn}(W_{\mathcal{T}}^T X_{\mathcal{T}}),$$

$$B_{\mathcal{V}} = \text{sgn}(W_{\mathcal{V}}^T X_{\mathcal{V}}),$$

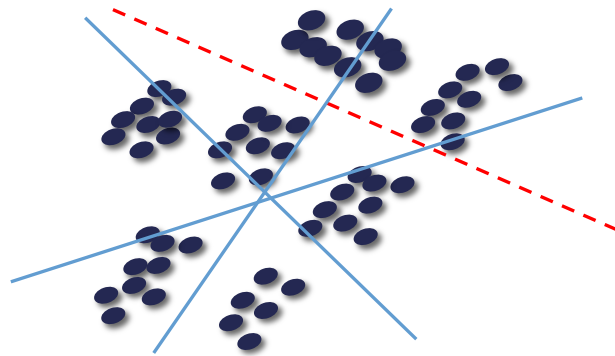
$$B_{\mathcal{T}}^{ij}(w_{\mathcal{V}i}^T X_{\mathcal{V}}^j) \geq 1 - \xi_{\mathcal{V}}^{ij} \quad \forall i, j,$$

$$B_{\mathcal{V}}^{ij}(w_{\mathcal{T}i}^T X_{\mathcal{T}}^j) \geq 1 - \xi_{\mathcal{T}}^{ij} \quad \forall i, j.$$

Solved by a block coordinate descent algorithm

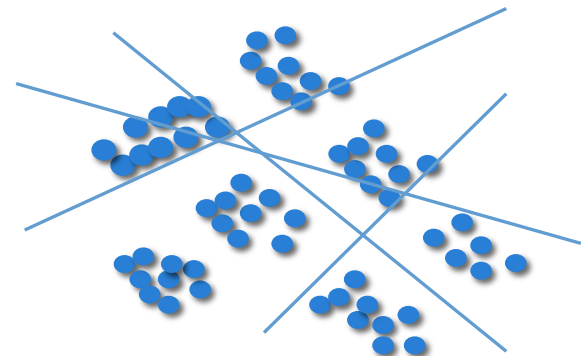
Predictable dual-view hashing

Modality 1



1101
0010
1010
0101
1110

Modality 2

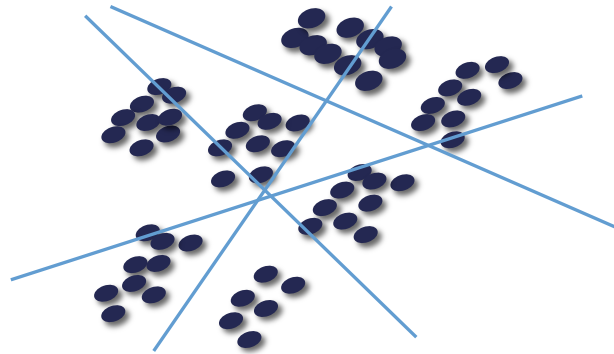


label →

0101
1110
1010
0101
1001

Predictable dual-view hashing

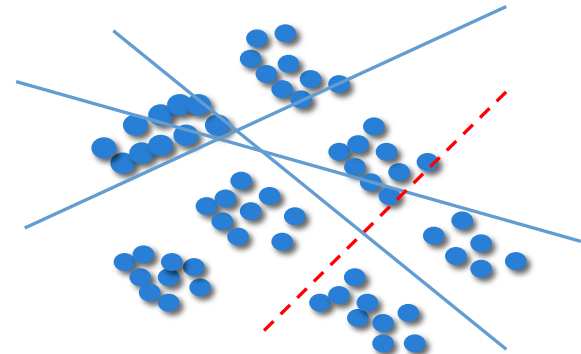
Modality 1



1101
0010
1010
0101
1110

← label

Modality 2



0101
1110
1010
0101
1001

Image → image (with 32 bits)

Query



Retrieval Set



Text → Image (with 32 bits)

Bike riding in a field



Cows standing in a village



Laptop placed on the table



Persons standing in a room



Plane flying on the air

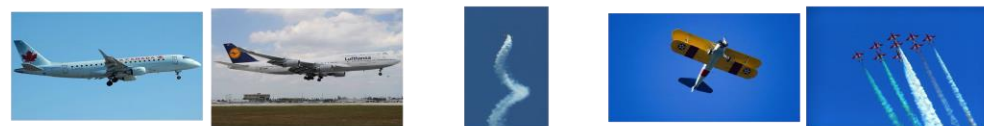
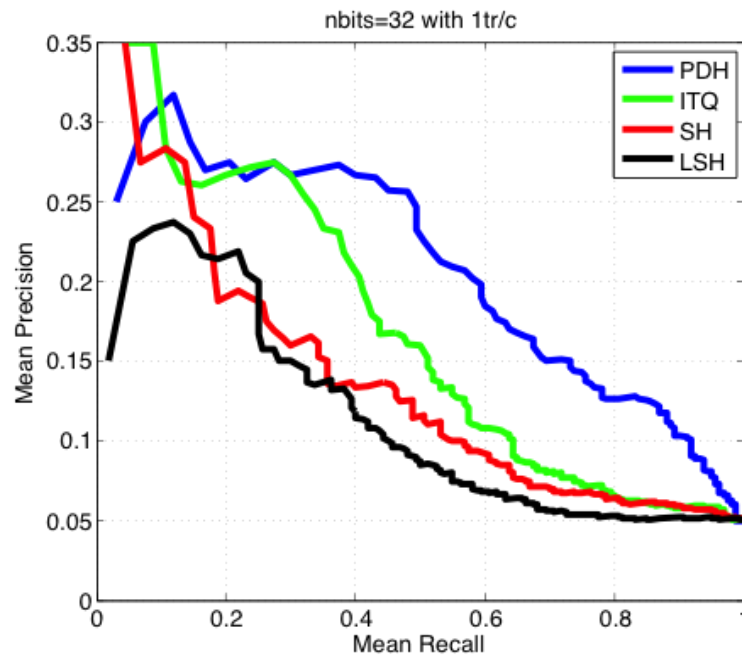
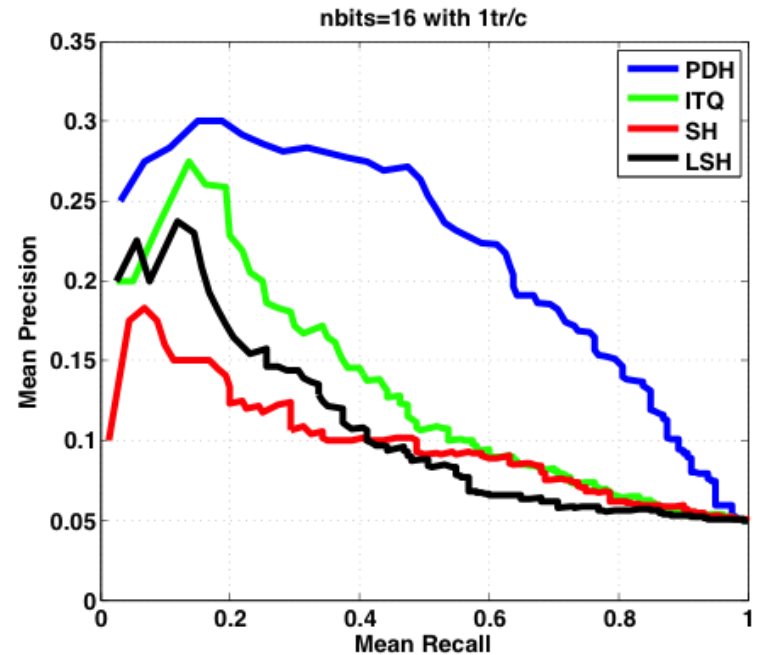


Image Retrieval



Pascal Sentence [Farhadi et al]



Sun Attribute [Peterson et al]

Thank you

Predictable Dual-view Hashing

Mohammad Rastegari, Jonghyun Choi, Shobeir Fakhraei*,
Hal Daumé III and Larry S. Davis

University of Maryland, College Park, MD, USA



UMIACS

