LEHIGH
UNIVERSITY

CASHEIRS: Cloud Assisted Scalable Hierarchical Encrypted Based Image Retrieval System

Xin Li, Qinghan Xue, Mooi Choo Chuah

INFOCOM, 2017

OUTLINE

- Motivation and challenges
- Selected image retrieval schemes
- Our solution
- Evaluations
- Future work
- References

OUTLINE

- Motivation and challenges
- Selected image retrieval schemes
- Our solution
- Evaluations
- Future work
- References

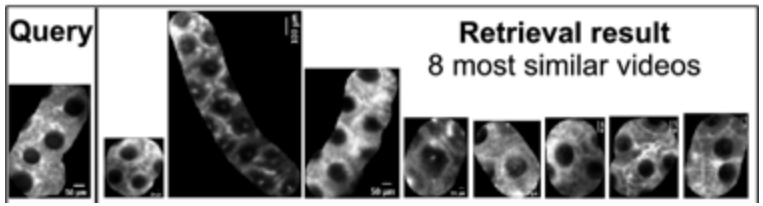
Motivation

1. Image retrieval techniques are needed in different applications.

Normal search

Online shopping

Biological and medical research

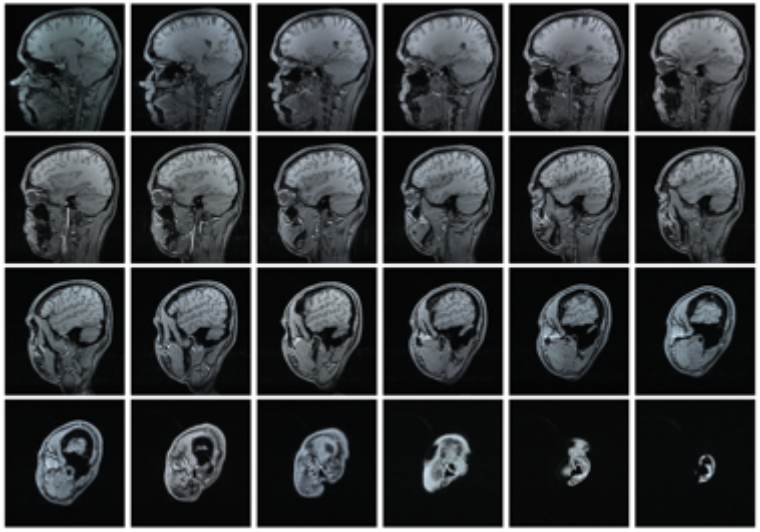


Social Network

Motivation

2. Sensitive Images

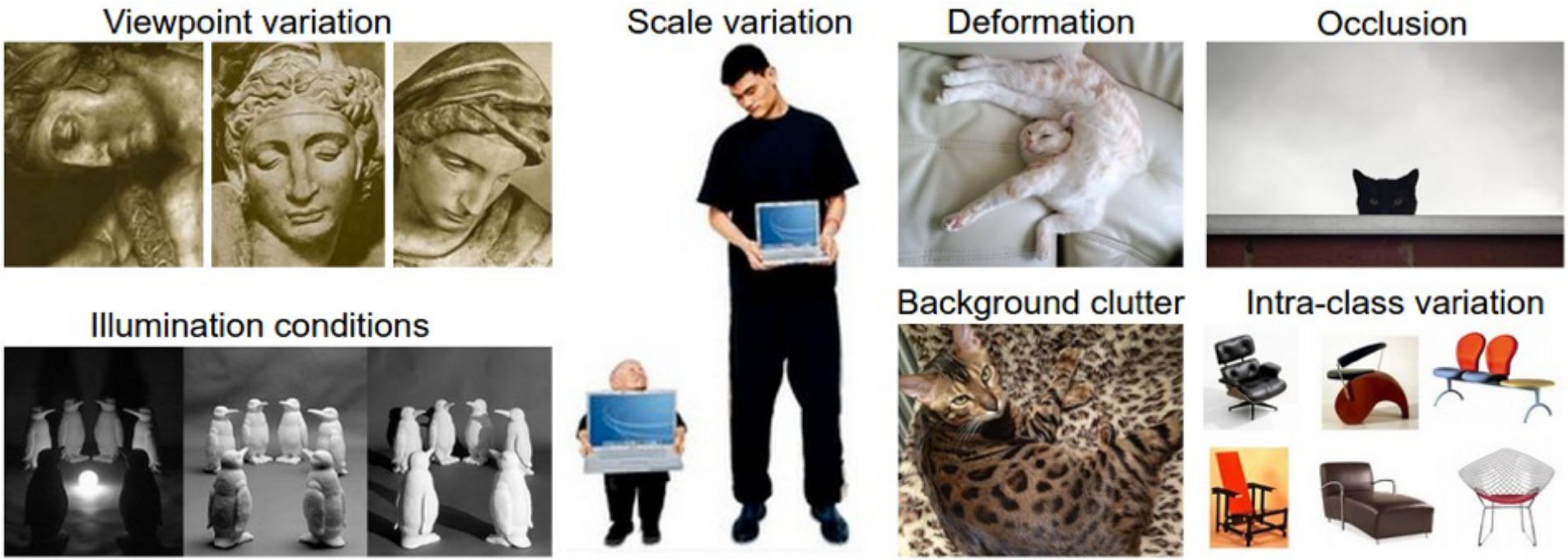
Magnetic resonance images



Famous people

Challenges

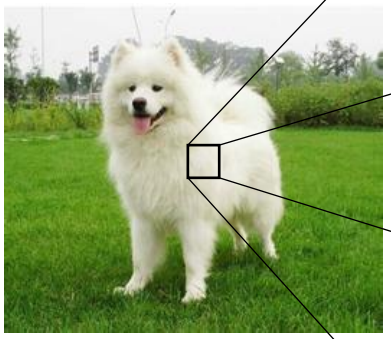
- 1. Variation of images:



(Image source: http://cs231n.github.io/classification/)

Challenges

- 2. Semantic gap:



13	97	70	10	17	32	86	58	43	23	86	25	21	85	56
26	46	74	93	80	12	57	07	11	33	87	85	01	42	04
73	19	34	75	02	75	14	68	30	69	05	01	96	14	84
70	88	50	94	66	73	84	45	79	95	92	57	94	26	12
74	99	24	18	42	31	60	24	56	34	81	11	96	45	03
53	89	14	90	00	12	08	08	09	75	11	17	09	22	17
66	10	20	91	70	56	53	06	74	60	34	19	25	41	82
73	69	69	32	15	85	79	52	86	62	75	60	30	02	62
31	54	09	61	83	33	57	47	67	85	100	42	28	20	45
82	24	81	35	63	34	85	94	11	63	24	81	04	24	45
43	28	65	36	01	77	60	13	90	34	44	71	10	82	55
08	69	57	70	86	55	59	94	47	31	96	59	98	59	26
24	24	42	45	49	85	00	33	04	49	56	10	15	74	53
77	44	95	07	90	93	39	97	57	70	49	40	81	25	40
81	30	19	60	48	01	65	82	01	78	02	89	54	23	75

What the computer sees

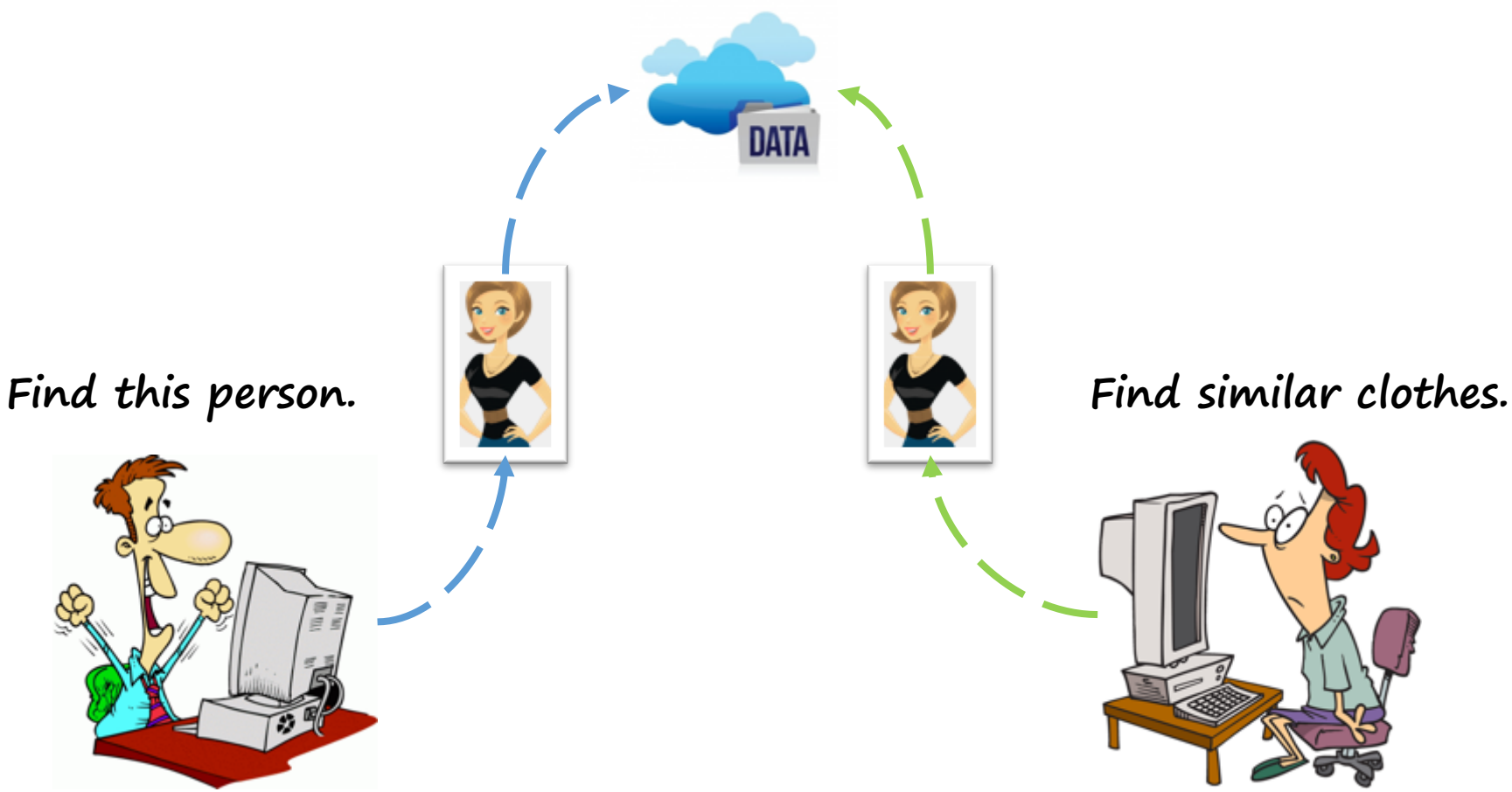
?

Semantic Meaning
Cat
Dog
Hat
…
Boat

What humans see

Challenges

- 3. Evaluation Criterion:



OUTLINE

- Motivation and challenges
- Selected image retrieval schemes
- Our solution
- Evaluations
- Future work
- References

Selected image retrieval schemes

B-Hie: Hierarchical Semantic Indexing for Large Scale Image Retrieval [1]

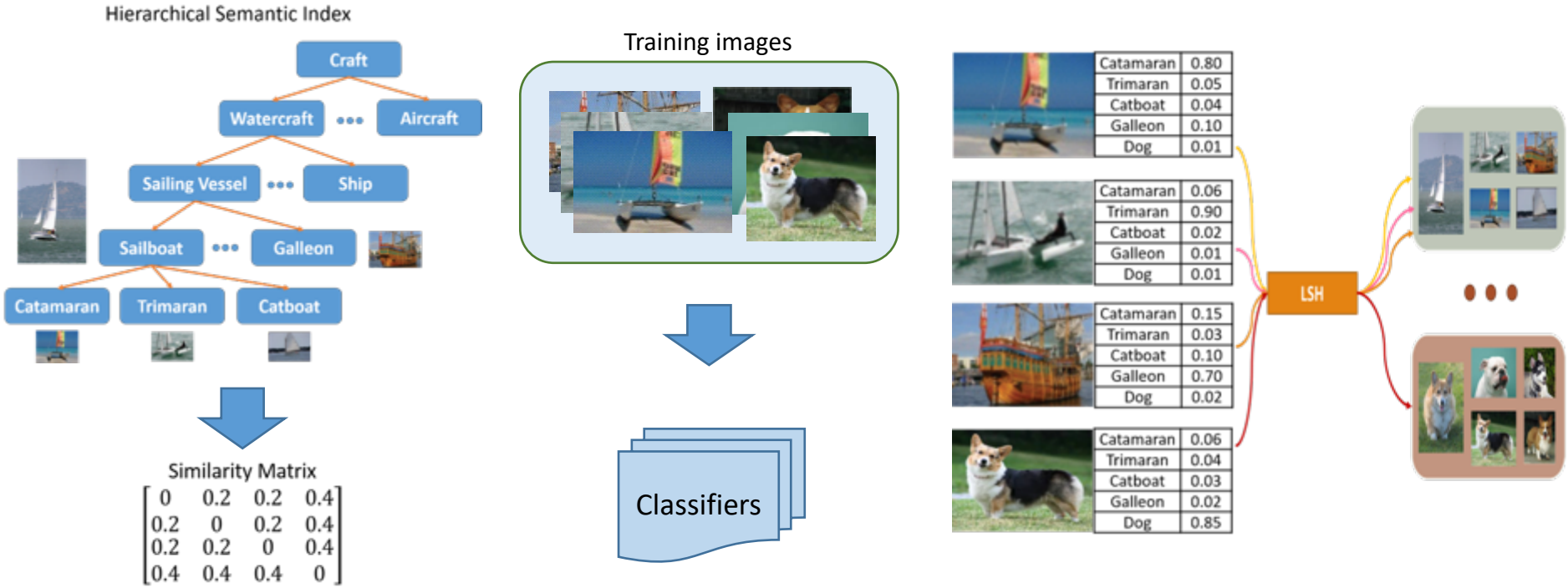
- Jia Deng, Alexander C. Berg, Li Fei-Fei
- CVPR, 2011

- Goal:
  - This paper aims to address the problem of similar image retrieval, especially in the setting of large-scale datasets with millions to billions of images.

System Flow

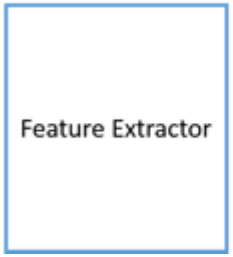
Four steps to build the system:

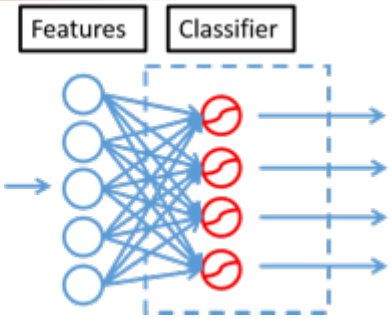
1. Calculate a similarity matrix based on a given hierarchical semantic tree.
2. Train a 1-vs-all classifier for each category.
3. Generate an attribute vector for each training image using trained classifiers.
4. Hash all images into different bins based on attribute vectors.



Query process

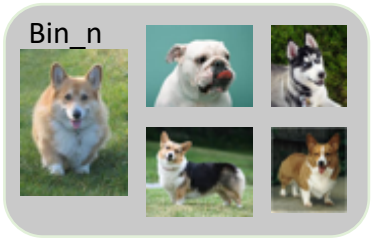
Query Image





Category	Attribute
Catamaran	0.10
Trimaran	0.20
Catboat	0.10
Galleon	0.60

LSH



For each candidate image

Category	Attribute
Catamaran	0.05
Trimaran	0.80
Catboat	0.05
Galleon	0.10

Calculate Similarity Score:

$$Sim(a, b) = f(a)^T S f(b)$$

$$= [0.1\ 0.2\ 0.1\ 0.6] \begin{bmatrix} 1 & 0.6 & 0.6 & 0.4 \\ 0.6 & 1 & 0.6 & 0.4 \\ 0.6 & 0.6 & 1 & 0.4 \\ 0.4 & 0.4 & 0.4 & 1 \end{bmatrix} \begin{bmatrix} 0.05 \\ 0.80 \\ 0.05 \\ 0.10 \end{bmatrix}$$

$$= 0.576$$

Sort by Similarity Score

Limitation

- 1. The dimension of attribute vector may be very large.
- 2. Use all classifiers for each query image.
  - => Slow query response time.
- 3. Need to re-compute the attribute vectors for all images even if we only add one new category.
- 4. No privacy-aware.

Selected image retrieval schemes

SEISA: Secure and Efficient Encrypted Image Search With Access Control [2]

- Jiawei Yuan, Shucheng Yu, Linke Guo
- INFOCOM, 2015

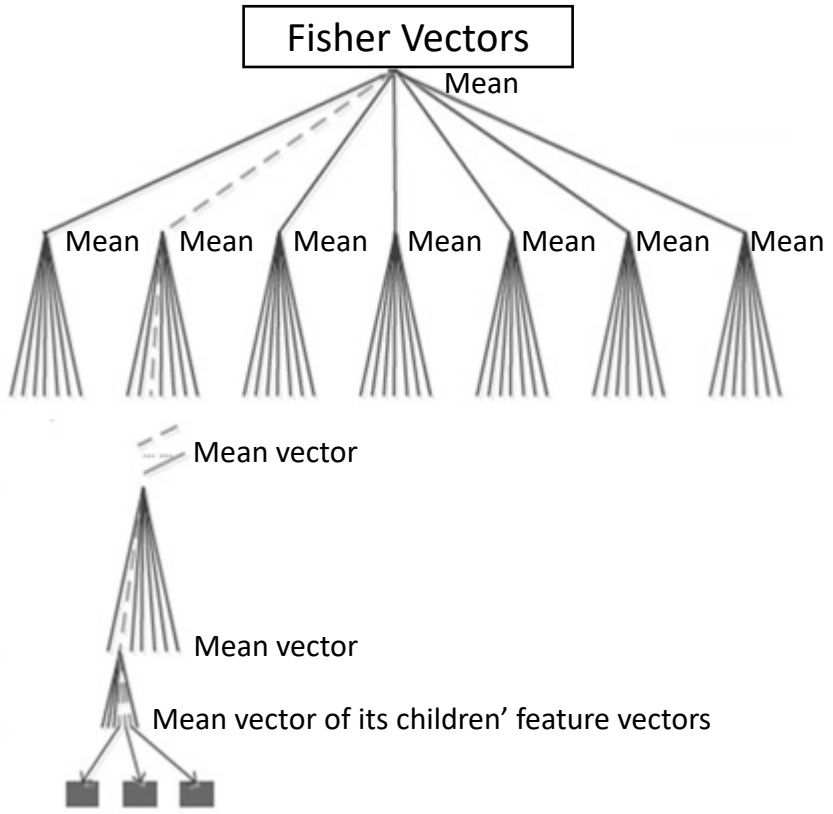
- Goal:
  - This paper aims to search encrypted images in a secure and efficient way.

1. Build an index tree

a. Extract Fisher vector from each image.

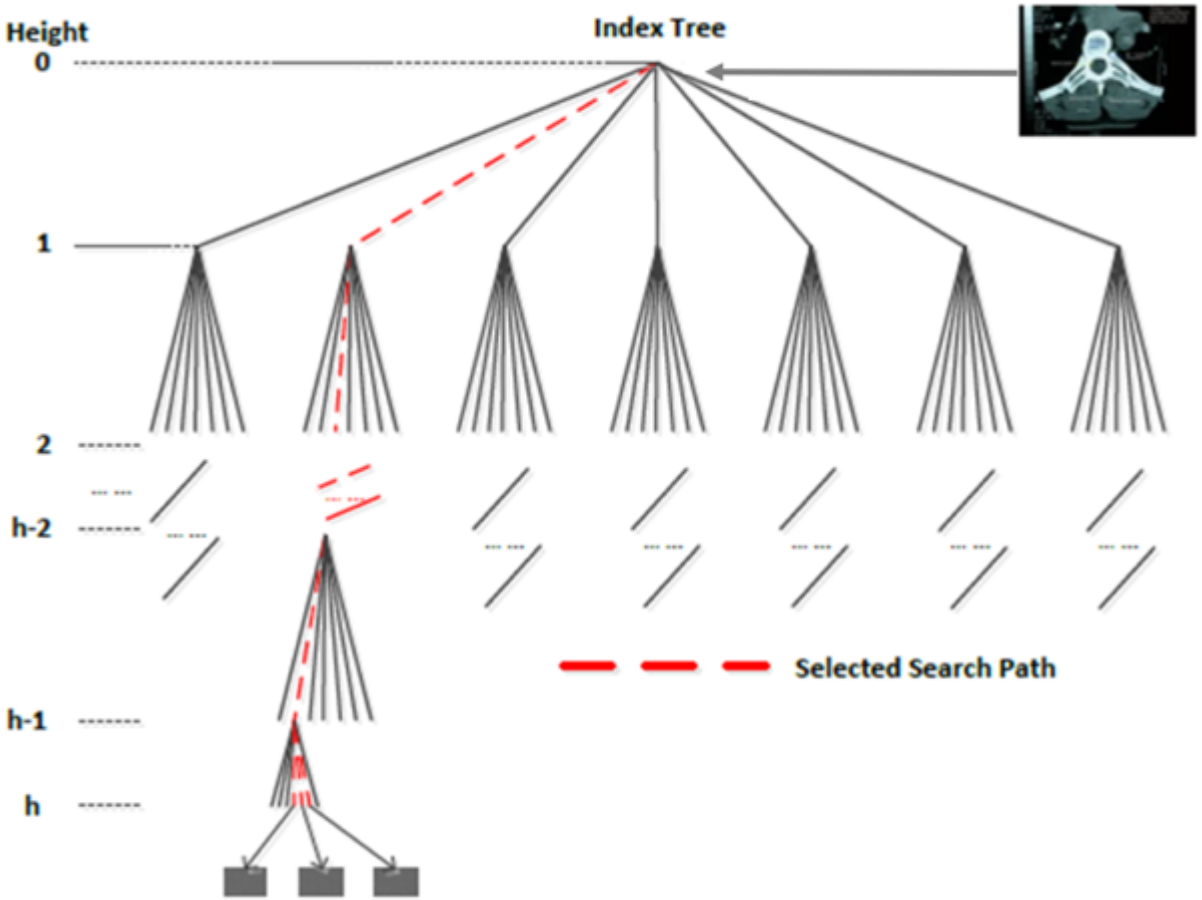
b. Employ the K-Means to generate the tree.

c. Assigns a mean vector to each intermediate node.



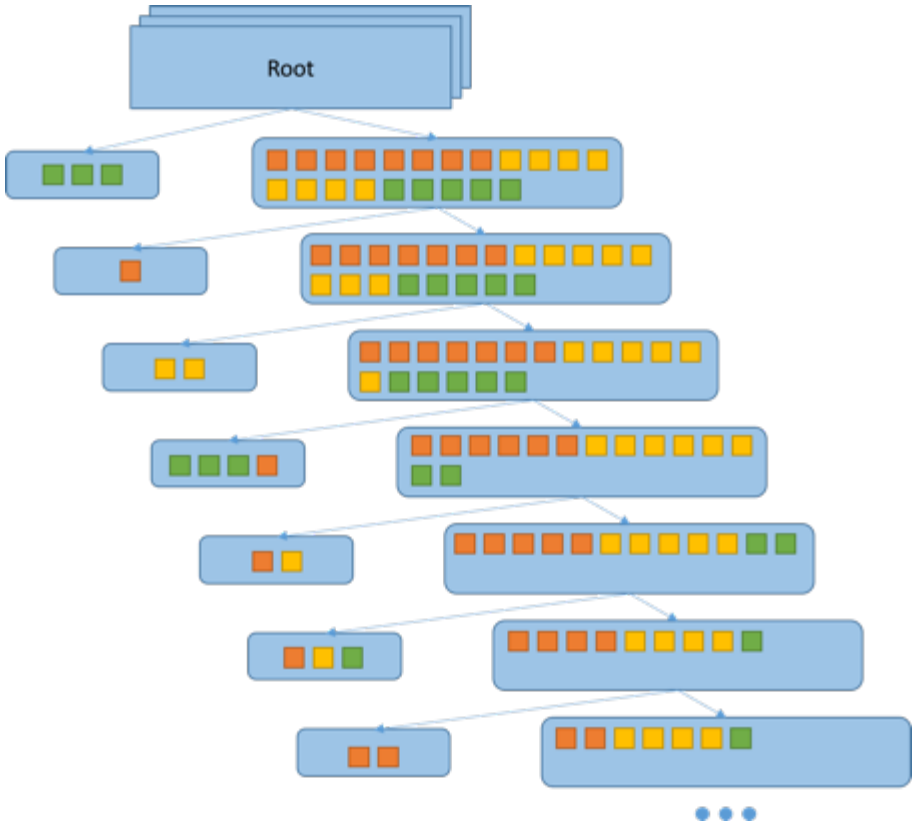
2. Image retrieval

Search the most similar nodes through the index tree.



Limitation

- It is possible that the depth of the index tree can become very deep.
- => Slow query response time and low retrieval accuracy.



OUTLINE

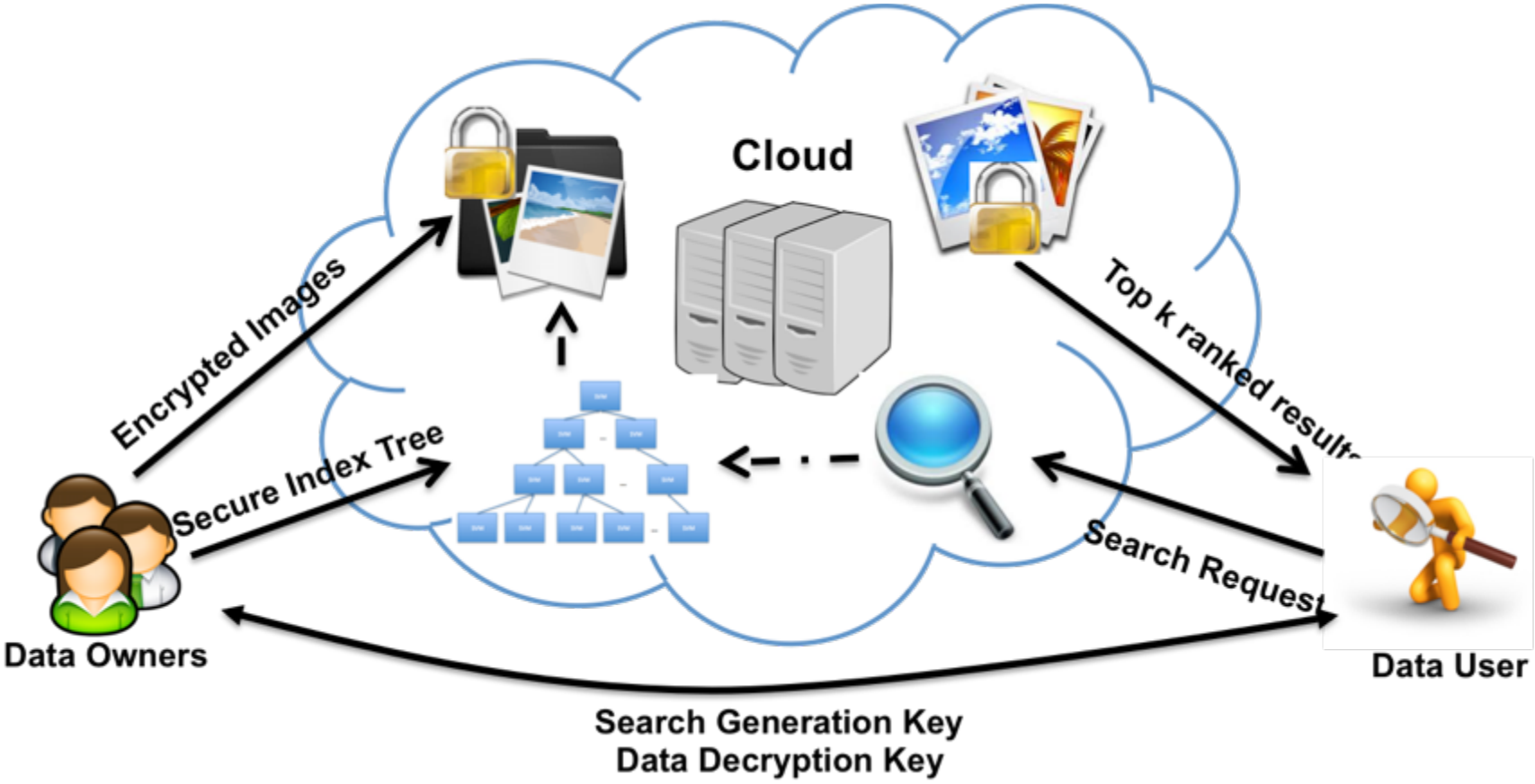
- Motivation and challenges
- Selected image retrieval schemes
- Our solution
- Evaluations
- Future work
- References

Our solution

Build an efficient image retrieval system

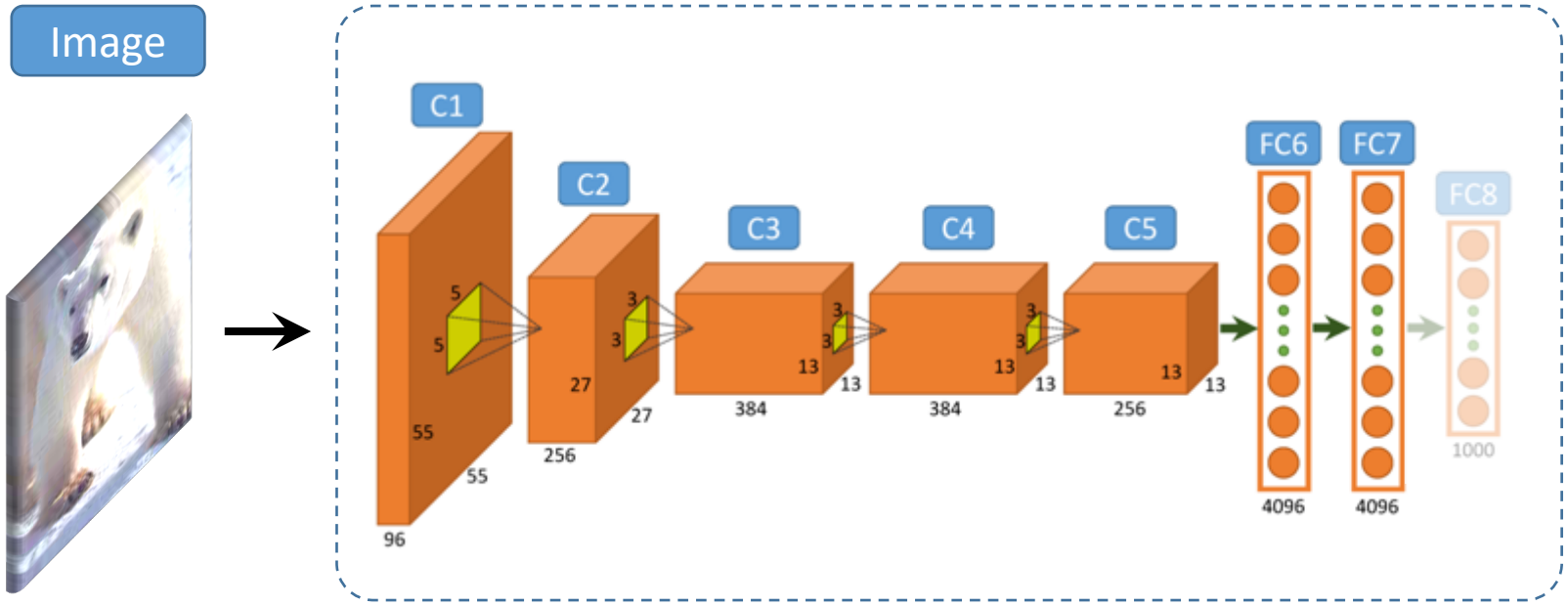
- Scalable
  - Deal with large-scale datasets.
- Hierarchical
  - Speed up the query process by using hierarchical structure to quickly identify a small subset of candidate images.
- Secure
  - Prevent the sensitive information of images from being leaked.

System Model



Visual feature representation

Image feature:



Fischer, Philipp, Alexey Dosovitskiy, and Thomas Brox. "Descriptor matching with convolutional neural networks: a comparison to sift." *arXiv preprint arXiv:1405.5769* (2014).

Visual feature representation

Image feature --> binary code:

High-dimensional descriptor vectors

Binary encoding

1011001011…0101
1110101101…0010
0001111000…1010
1111111001…0001
1010101010…1001
0001111110…1010
0101101001…1111
1001111000…1010
0001001001…0010
1001110010…1010
1101111000…0010
1101111001…0001
0001111000…0010
1010011011…1111

Visual feature representation

Image feature --> binary code:

Similar images

$$||x - y|| \approx 0$$

Binary encoding

1011001011…0101
1110101101…0010
0001111000…1010
1111111001…0001
1010101010…1001
0001111110…1010
0101101001…1111
1001111000…1010
0001001001…0010
1001110010…1010
1101111000…0010
1101111001…0001
0001111000…0010
1010011011…1111

Visual feature representation

Image feature --> binary code:

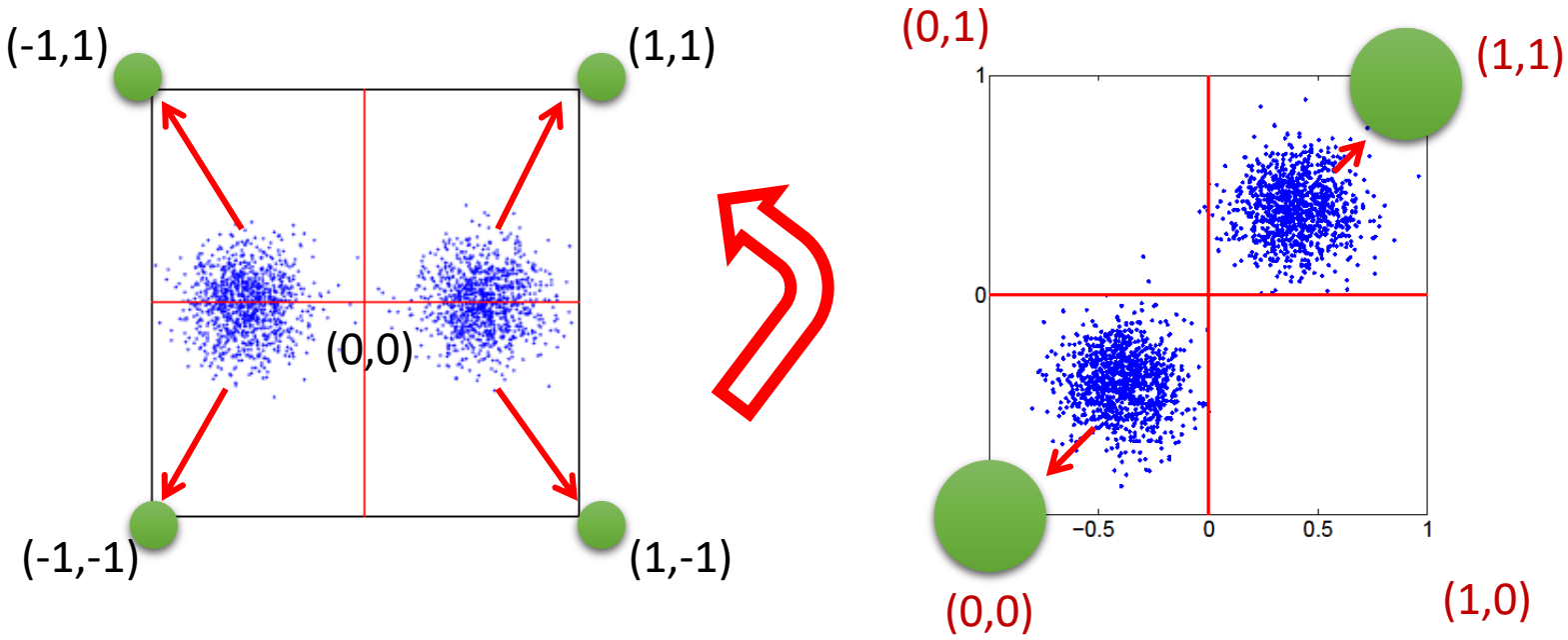
Dissimilar images
$||x - y||$ is large

Binary encoding

1011001011…0101
1110101101…0010
0001111000…1010
1111111001…0001
1010101010…1001
0001111110…1010
0101101001…1111
1001111000…1010
0001001001…0010
1001110010…1010
1101111000…0010
1101111001…0001
0001111000…0010
0010011011…1111

Visual feature representation

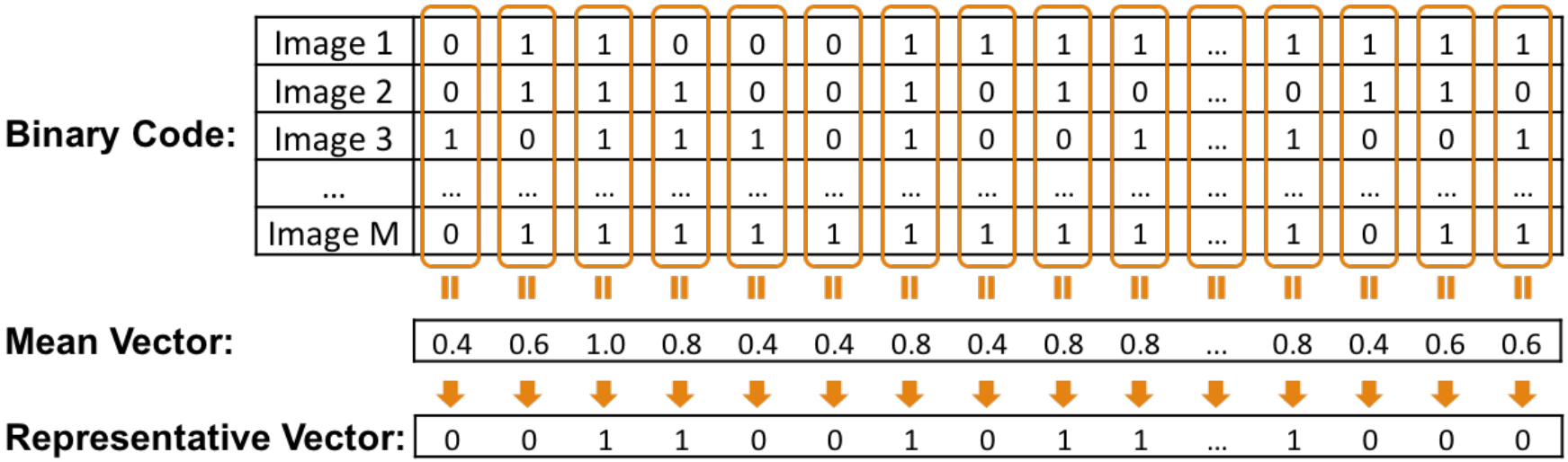
Image feature --> binary code (ITQ):



Gong, Yunchao, et al. "Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.12 (2013): 2916-2929.

Visual feature representation

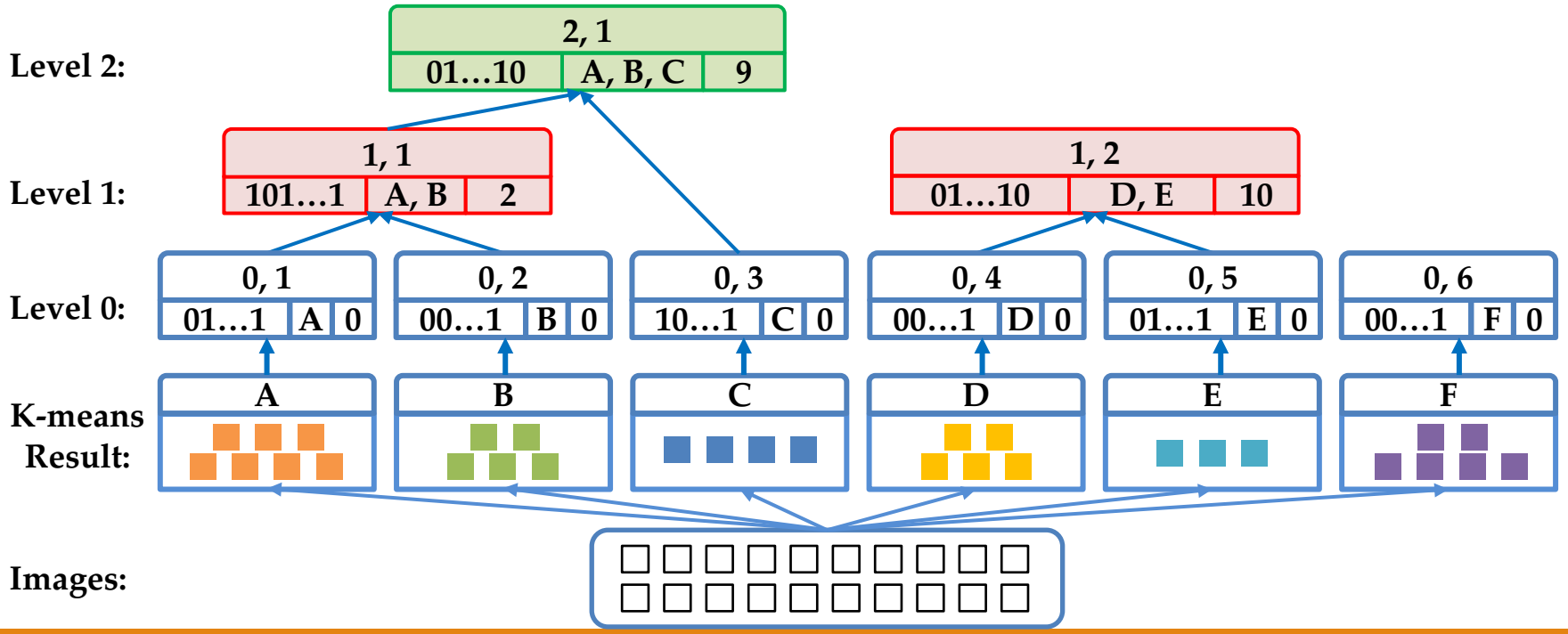
Representative Vector (RV):



Hierarchical index tree building

Cluster Index: Lev h , Ind i		
RV	E	IND

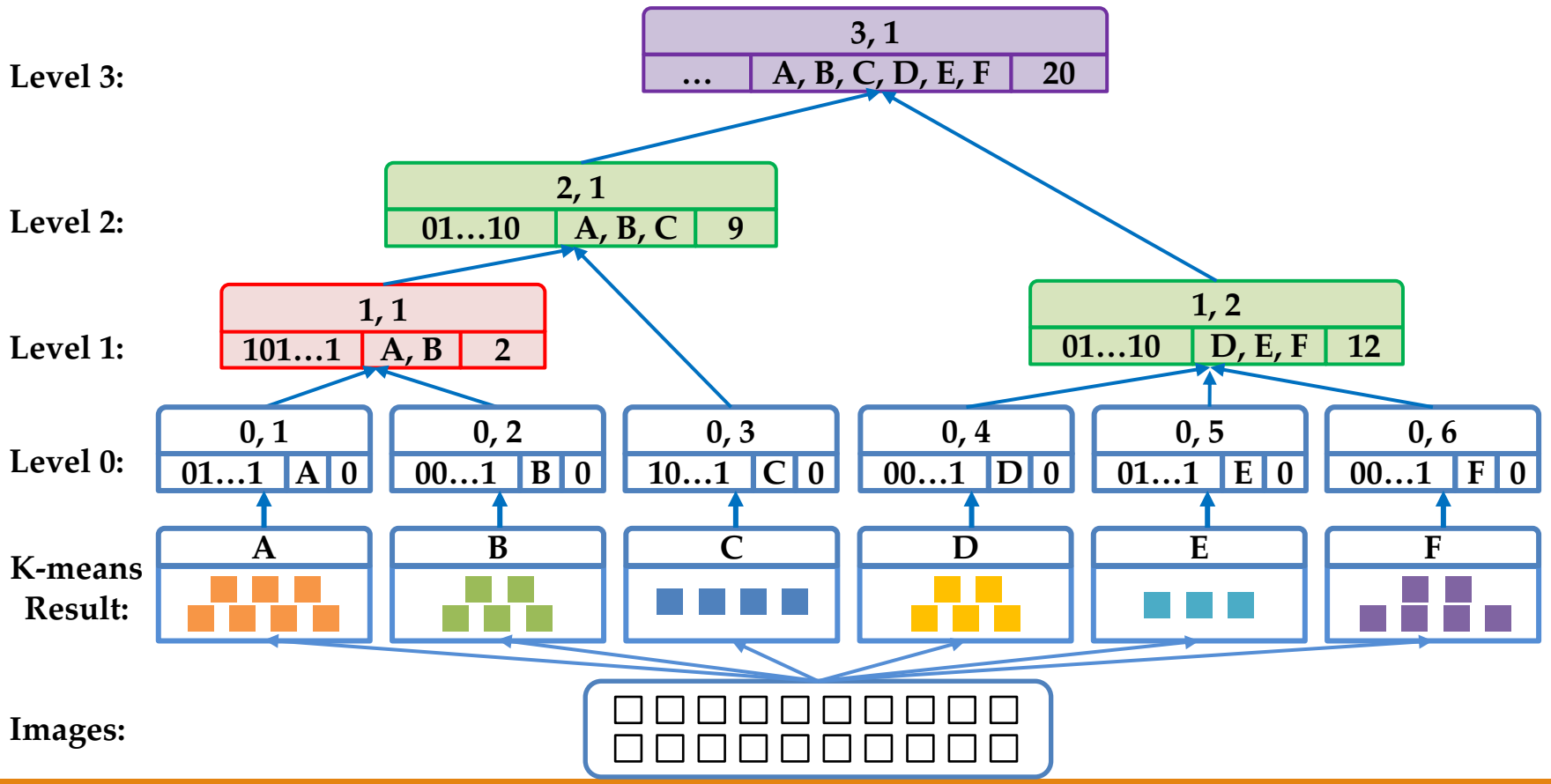
$$\frac{|D(a,b)-\max(IND_a,IND_b)|}{\max(IND_a,IND_b)} > 0.5 \text{ (threshold)}$$



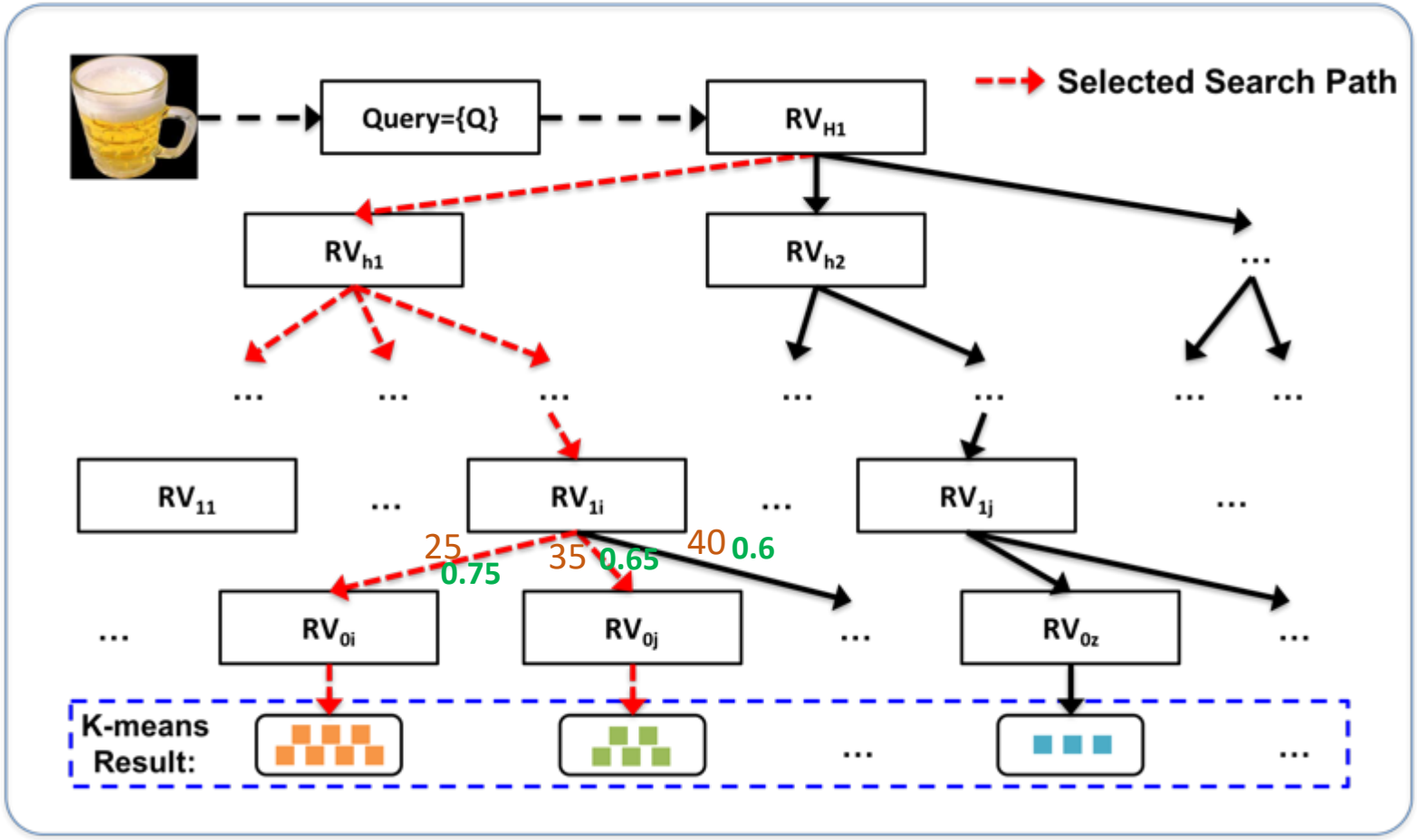
Hierarchical index tree building

Cluster Index: Lev h , Ind j		
RV	E	IND

$$\frac{|D(a,b)-\max(IND_a,IND_b)|}{\max(IND_a,IND_b)} > 0.5 \text{ (threshold)}$$



Query process



$$w_{hj} = 1 - \frac{D(Q, RV_{hj})}{\sum_{i=0}^{z} D(Q, RV_{hi})}$$

Sum_Chosen $= \sum_{c=0}^{k} w_c$ > 0.8 (threshold)

Privacy

	RV_1 (Cloud)	RV_2 (User)
Binary Code:	0 0 1 1 0 0 1	0 1 1 1 0 0 0

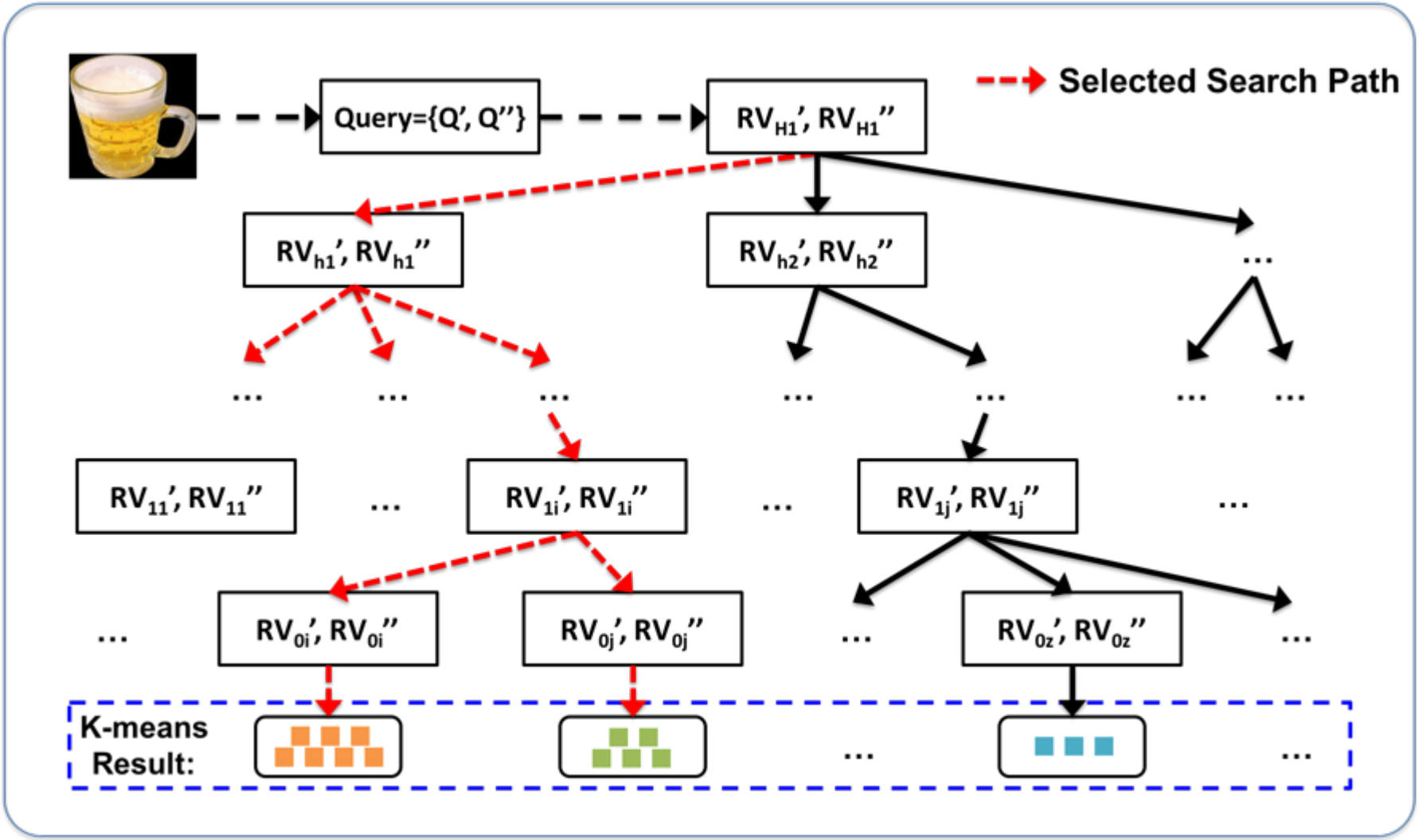
Step 1:	RV_1'	-1 -1 1 1 -1 -1 1	RV_2' -1 1 1 1 -1 -1 -1
	S_1	1 0 0 1 0 1 1	S_2 0 0 1 1 0 1 0
Step 2:	RV_{11}'	-1 a_1 a_2 1 a_4 -1 1	RV_{21}' -1 1 a_2 a_3 -1 a_5 -1
	RV_{12}'	-1 b_1 b_2 1 b_4 -1 1	RV_{22}' -1 1 b_2 b_3 -1 b_5 -1

Step 3: Generate invertible random matrices M_1 and M_2

Step 4: $Enc(RV_1') = \{M_1^T RV_{11}' , \quad M_2^T RV_{12}'\}$ $\quad Enc(RV_2') = \{M_1^{-1} RV_{21}' , \quad M_2^{-1} RV_{22}'\}$

Distance: $$D(RV_1, RV_2) = \frac{n - RV_1' \cdot RV_2'}{2} = \frac{n - Enc(RV_1') \cdot Enc(RV_2')}{2}$$

Query process (with encryption)



OUTLINE

- Motivation and challenges
- Selected image retrieval schemes
- Our solution
- Evaluations
- Future work
- References

Evaluations

- Metrics:
  - Precision at top k (P@k)

$$P@k = \frac{num\_correct}{k}$$

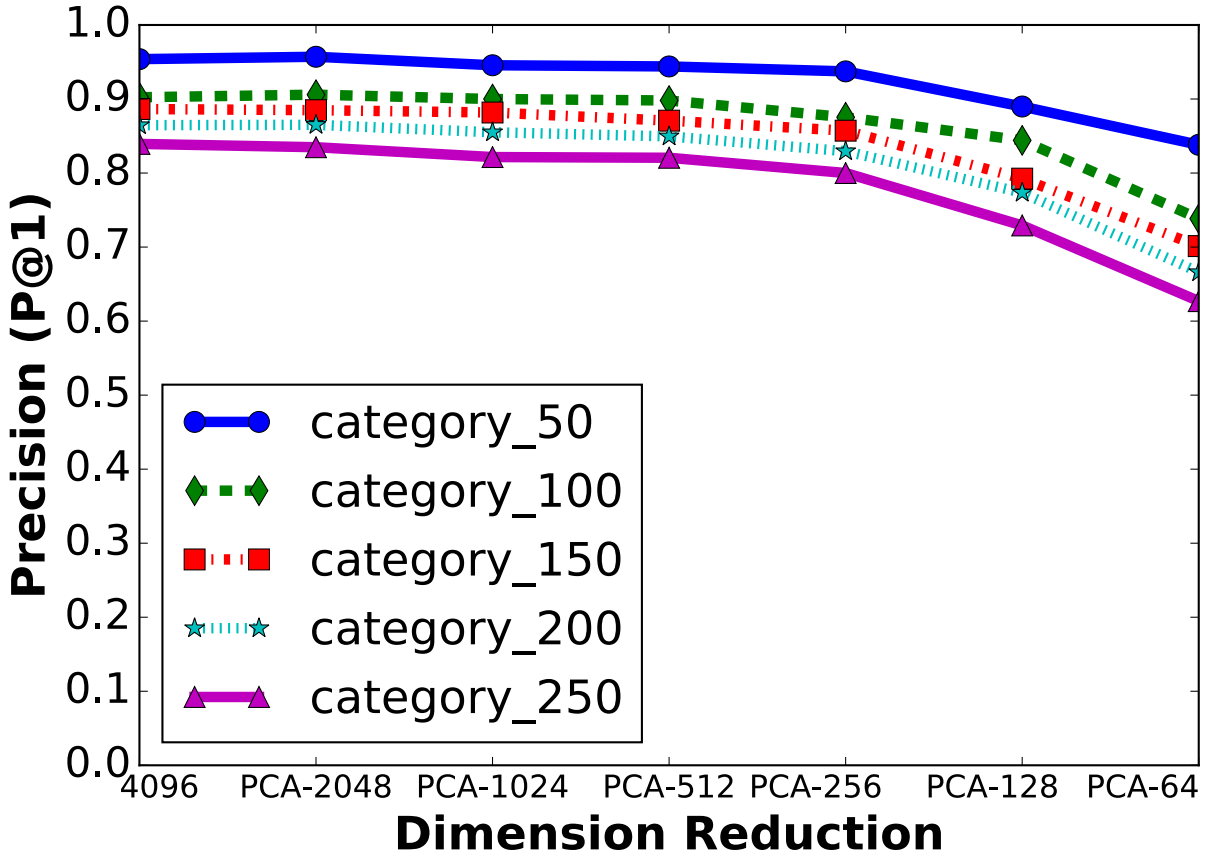
  - Mean average precision (mAP)

$$AP(q) = \frac{\sum_{k=1}^{n}(P@k * rel(k))}{N}$$

- Dataset:
  - Caltech256
    - 30608 images, 256 object categories
  - INRIA Holiday
    - 1491 images, 500 image groups
- Evaluate on:
  - A laptop running OS X
  - 2.5GHz Intel Core i7 CPU
  - 16GB Memory

Evaluations – Effectiveness of RV

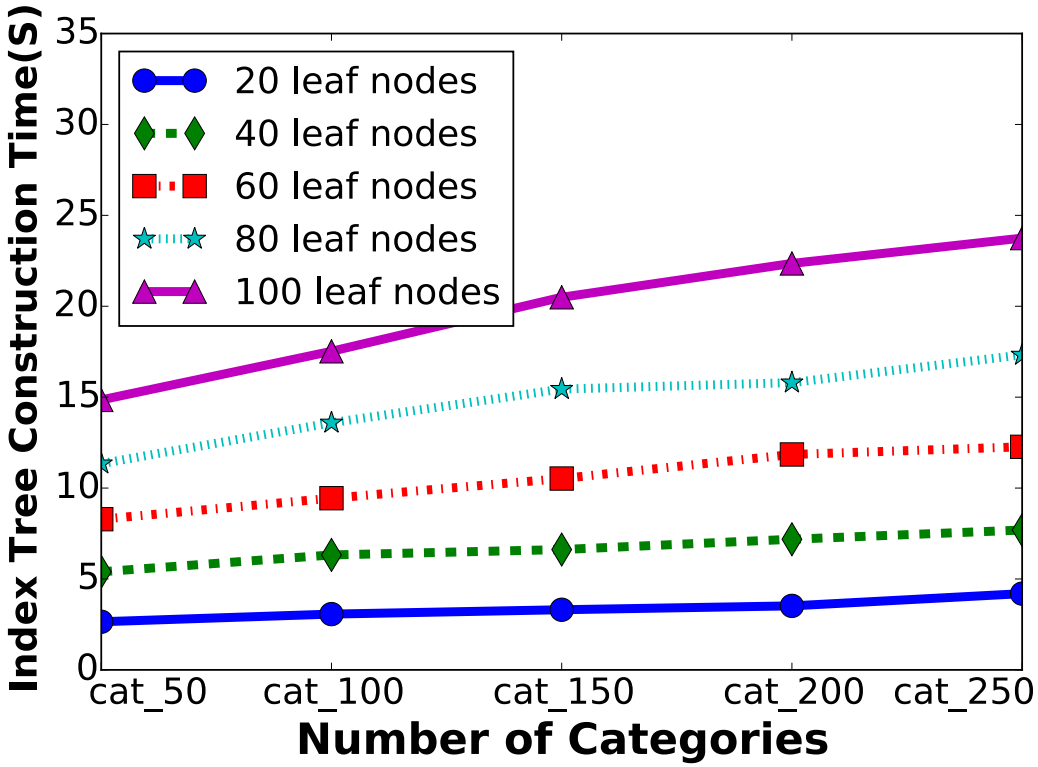
- 50 images per category are used
- Using image category label to build the index tree instead of K-Means.



Evaluations – System Construction

Computation time of building a hierarchical index tree.

- 50 images per category are used
- binary code: PCA-128



Evaluations – Storage Cost

CASHEIRS: (binary code: PCA-128)

- CNN model: 244.7 MB [Data owner | Data user]
- PCA matrix: 3.3 MB [Data owner | Data user]
- Rotation matrix: 105 KB [Data owner | Data user]
- Encrypted key: 1.7 KB [Data owner | Data user]

- Encrypted images: 14.28 KB/image (300 x 200 pixels) [Data owner | Cloud]
- Encrypted index tree: 440 KB (cat_50) [Data owner | Cloud]
- Image features: 5.3 MB (cat_50) [Data owner | Cloud]

Cloud Storage Cost (MB)	Cat_50	Cat_100	Cat_150	Cat_200	Cat_250
CASHEIRS	40.6	81.0	121.5	161.4	201.7
B-Hie	59.8	120.4	181.0	241.7	302.4
OASIS	109.3	197.4	285.5	373.6	461.7
SEISA (PCA-128)	40.3	80.8	122.3	167.1	211.8
SEISA (PCA-512)	55.7	113.9	175.9	239.6	303.3

Evaluations – Search Evaluation

TABLE II: Comparison Results

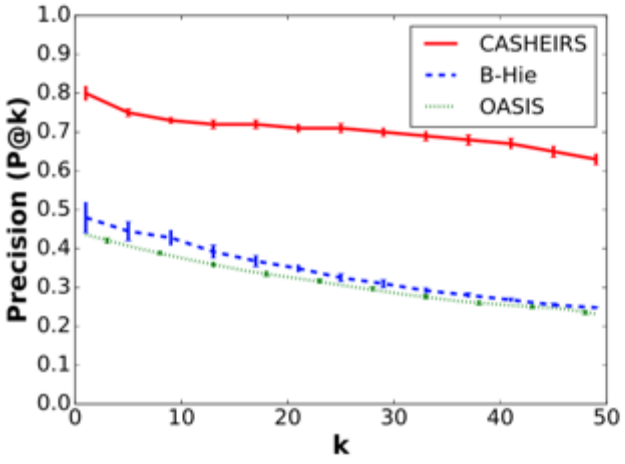
Caltech256			
Schemes	**Search Time (ms)**		
	Category-10	Category-20	Category-50
CASHEIRS	4.1	9.6	12.9
B-Hie	13.3	23.2	52.2
OASIS	131.9	132.6	133.9

INRIA Holiday (10 million images)		
Scheme	Search Time (ms)	mAP
CASHEIRS	95.2	0.64
SEISA	87.5	0.55

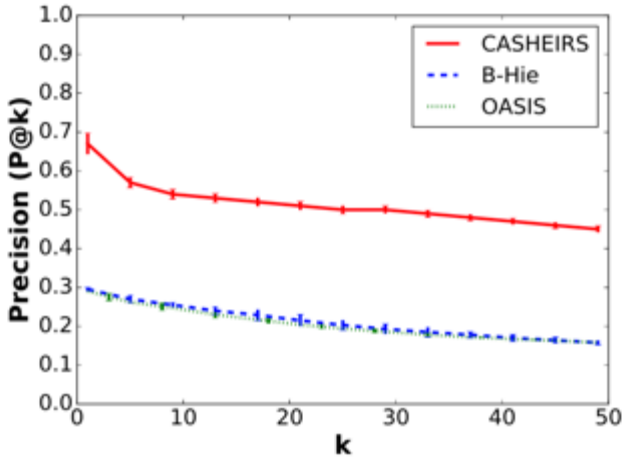
We randomly generate 10 Million image features (128 dimension), and then mix them with the features extracted from INRIA Holiday dataset. (Same set-up as SEISA)

Accuracy on Caltech256 dataset

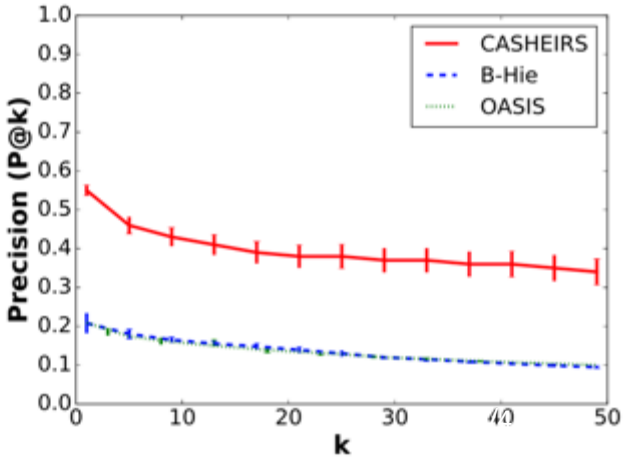
10 classes



20 classes



50 classes



Evaluations – Search Evaluation

Communication Cost:

Data owner to the cloud:

11.2 MB encrypted index tree [one time]

14.28 KB/encrypted image (300 x 200 pixels) [one time]

A user:

[one time]

receive 248.1 MB CNN model and encryption keys from the Data owner

[Querying]

send 2 KB encrypted query information to the cloud

receive 14.28 KB/encrypted image (300 x 200 pixels) from the cloud

Prototype

- Cloud server:
  - A laptop running OS X
  - 2.5GHz Intel Core i7 CPU
  - 16GB Memory
- Client
  - Samsung S5 phone
  - Snapdragon 801 chip with 2.5GHz Quad-core CPU
  - 2GB RAM
- Communication
  - WIFI router

OUTLINE

- Motivation and challenges
- Selected image retrieval schemes
- Our solution
- Evaluations
- Future work
- References

Future work

- Using larger image datasets, e.g. ImageNet.
- More efficient security solution.
  - 128 bit binary code into 2 KB encrypted information.
- Using deep learning model to generate a binary code directly.

Conclusion

- Challenges in the field of image retrieval.
- Related papers and their limitations.
- Propose our solution.

References

[1] Jia Deng, Alexander C. Berg, and Li Fei-Fei. "Hierarchical semantic indexing for large scale image retrieval." *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011.

[2] Jiawei Yuan, Shucheng Yu, and Linke Guo. "SEISA: Secure and efficient encrypted image search with access control." *Computer Communications (INFOCOM), 2015 IEEE Conference on*. IEEE, 2015.

[3] Lan Zhang, Jung Taeho, Cihang Liu, Xuan Ding, Xiang-Yang Li, Yunhao Liu, "POP: Privacy-preserving Outsourced Photo Sharing and Searching for Mobile Devices." Distributed Computing Systems(ICDCS), 2015 IEEE Conference on. IEEE, 2015.

[4] Chechik, Gal, et al. "Large scale online learning of image similarity through ranking." Journal of Machine Learning Research 2010

[5] Fischer, Philipp, Alexey Dosovitskiy, and Thomas Brox. "Descriptor matching with convolutional neural networks: a comparison to sift." arXiv preprint arXiv:1405.5769 (2014).