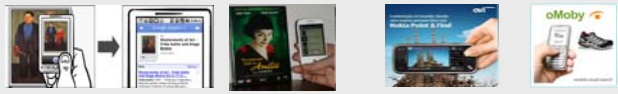
Mobile Product Search with Bag of Hash Bits and Boundary Reranking

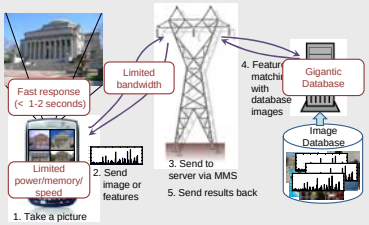
Junfeng He, Jinyuan Feng, Xianglong Liu, Tao Cheng, Tai-Hsu Lin, Hyunjin Chung, Shih-Fu Chang

Columbia University in the City of New York — DVMM lab

Mobile Visual Search: the next big thing?



Overall pipeline



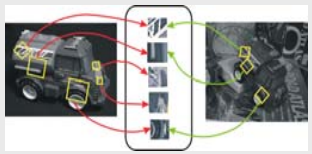
Main challenges

MVS calls for

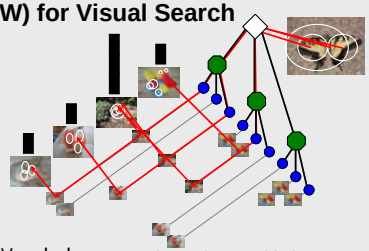
Distributed Optimization

- Client: light process for compact feature extraction
- Radio: transmit compact codes only
- Server: scalable indexing and object level matching over large databases

Bag of Words (BoW) for Visual Search



Search with BoW.
Picture from David Lowe



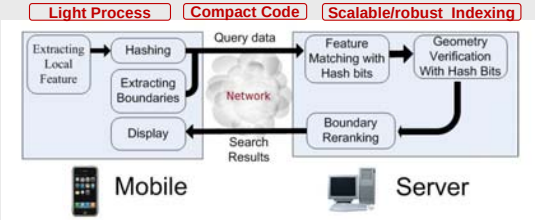
BoW via Vocabulary Tree. Picture from David Nister.

Bag of Words (BoW) for Mobile Visual Search?

- Sending image? transmit cost is too high
- Sending visual words?
  - Quantizing LFs is hard to do on mobile devices
  - LF matching via quantization in server side is too coarse

The Proposed Columbia MVS System –An Answer to the Distributed Optimization Challenge

Light Process | Compact Code | Scalable/robust Indexing



- Indexing/Search with **Bag of Hash Bits (BoHB)** instead of BoW
- Combine **global boundary features** with local features for reranking

Bag of Hash Bits

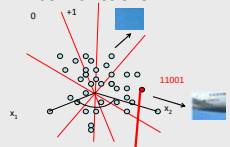
Tens of hash bits for each LF

Distributed Optimization !

- Client: only tens of inner products
  - Cheap computation/memory
- Radio: about 1KB for each query
  - Much less than sending images
- Server: More flexible/accurate match of local feature than BoW

Indexing/Search with Hash Bits

Hash functions:



Compute hash bits of database samples

X	x_1	x_2	x_3	x_4	....
h_1	0	1	1	0	...
h_2	1	1	1	0	....
...	...	...	...	...	...
h_k	...	...	...	...	...

00101... 11001... 10110... 0010... ...

Build hash table

hash codes	data index
00101...	X_1 X_{12}...
11001...	X_2 X_8 X_{23}...
10110...	X_3 ...
.....	

Offline Indexing

- Compute hash bits for database
- Build hash tables

Online Search

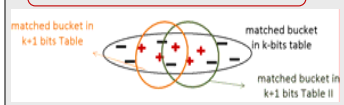
- Compute hash bits for the query
- Retrieve candidates by table lookup
- Rank candidates

Why Bag of Hash Bits (BoHB) is superior -- Or the Secrets to Make Hashing Successful for MVS

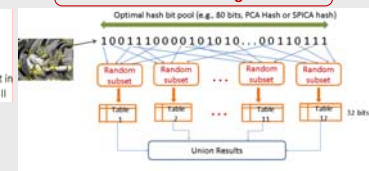
- Advantages of meaningful hamming distance between word index
  - BoW: word index is meaningless
  - BoHB: check words within a small hamming distance

- More flexible/robust indexing with multi hash tables via bit reusing
  - BoW: longer codes or multi indexing are prohibitively expensive
  - BoHB: flexible to support longer codes and multi hash tables

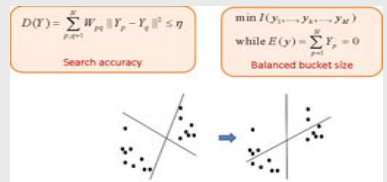
Multi tables with long codes give better search results



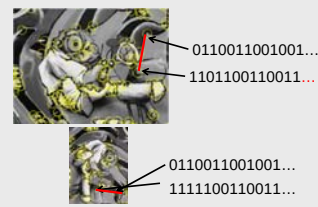
Multi table in very small bit budget? -- Bit Reusing



- Use compact hash bits instead of random LSH hash bits
  - Compact hash like SPICA hashing or PCA hashing
  - Not only preserve similarity/distance but minimize the dependence of bits

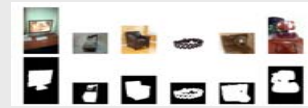


- Spatial reranking with hash bits
  - Length ratio similarity
  - Use hash bit hamming distance as soft weights in spatial verification



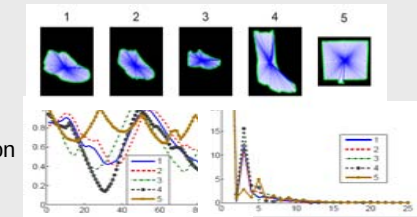
Extract Boundaries

- Mobile side: interactive segmentation like Grabcut
- Serverside: automatic AttentionCut



Boundary Feature -- Central Distance

- Distances from boundary points to center point
- FFT of distance vectors
- Translation invariant
- Scale invariant
  - Sampling +normalization
- Rotation invariant
  - Amplitude of FFT



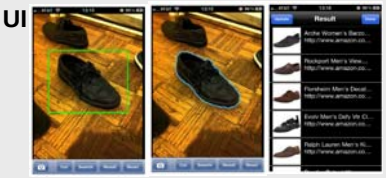
Reranking with Local Features and Boundary Feature

- Local feature (LF): capture details
- Boundary feature: represent the overall shape
- A fusion of LF spatial reranking + boundary reranking (L2 distance)

Datasets

- Two sets: 400K (300K) product images crawled from multiple online shopping companies like Amazon, Ebay, and Zappos
- Hundreds of categories; shoes, electrical devices, groceries…

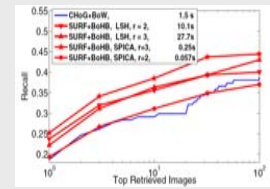
UI

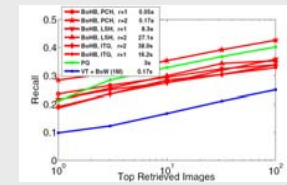


Speed

- SURF extraction: ~1s
- Compute Hash bits : < 0.1s
- Transmission: 80 bits per LF
- Search: ~0.4s on average
- Download/display: ~1-2s

Experiments





Conclusion

- BoHB achieves > 30% higher accuracy and >10 times faster search speed at comparable transmission data size than CHOG, the state of the art MVS
- BoHB outperforms most state-of-the-art visual search methods, like BoW via Vocabulary Tree, or Product Quantization (PQ)
- Boundary reranking is helpful to remove noisy candidates

Selected References

[1] V. Chandrasekhar, et.al. Mobile Product Recognition. MM, 2010.
[2] J. He, et. Al., Compact Hashing with Joint Optimization of Search Accuracy and Time. CVPR, 2011.
[3] V. Chandrasekhar, et. al. Chog: Compressed histogramof gradients a low bit-rate feature descriptor. *CVPR*, 2009