

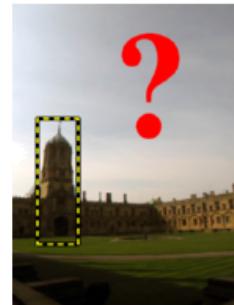
# **Total Recall: Automatic Query Expansion with a Generative Feature Model for Object Retrieval**

O. Chum, et al.

presented by Yangzihao Wang

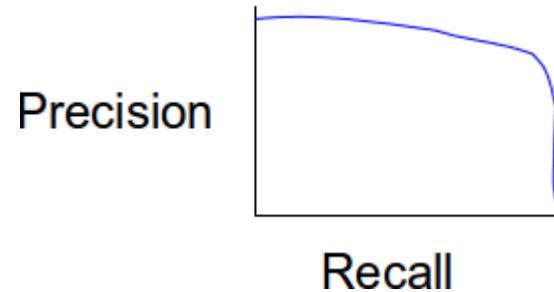
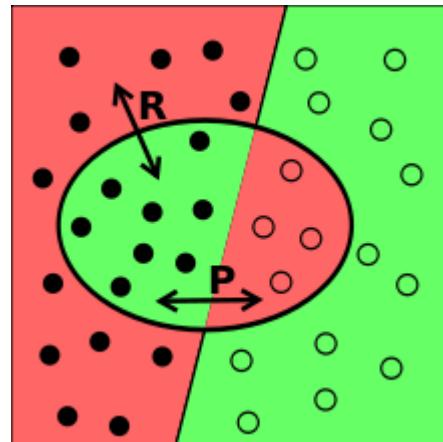
# Problem to Solve

Given a query image of an object, retrieve all instances of that object in a large (1M+) image database.



# What is Total Recall?

- Precision-Recall Curve



- Average Precision
  - area under the precision-recall curve

# Approach

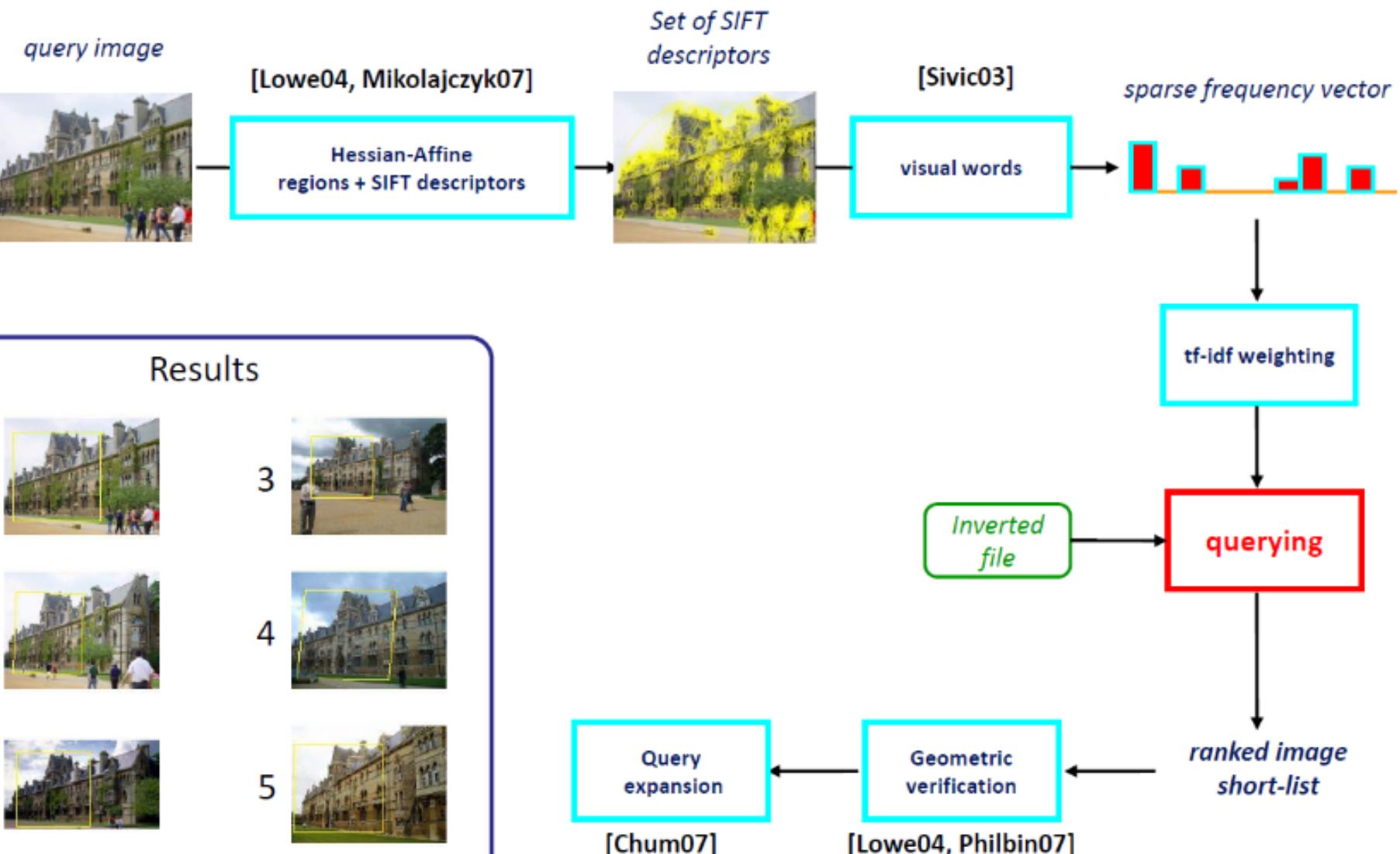


figure credit: Relja Arandjelovic

# Real-time Object Retrieval

- Generate descriptor
  - Hessian-Affine Region+SIFT
- Quantization
  - Approx k-means based on random trees
  - forming visual words
- Search Engine
  - inverted file(maps words to docs in which they occur)
  - tf-idf weighting scheme

# Main Contributions

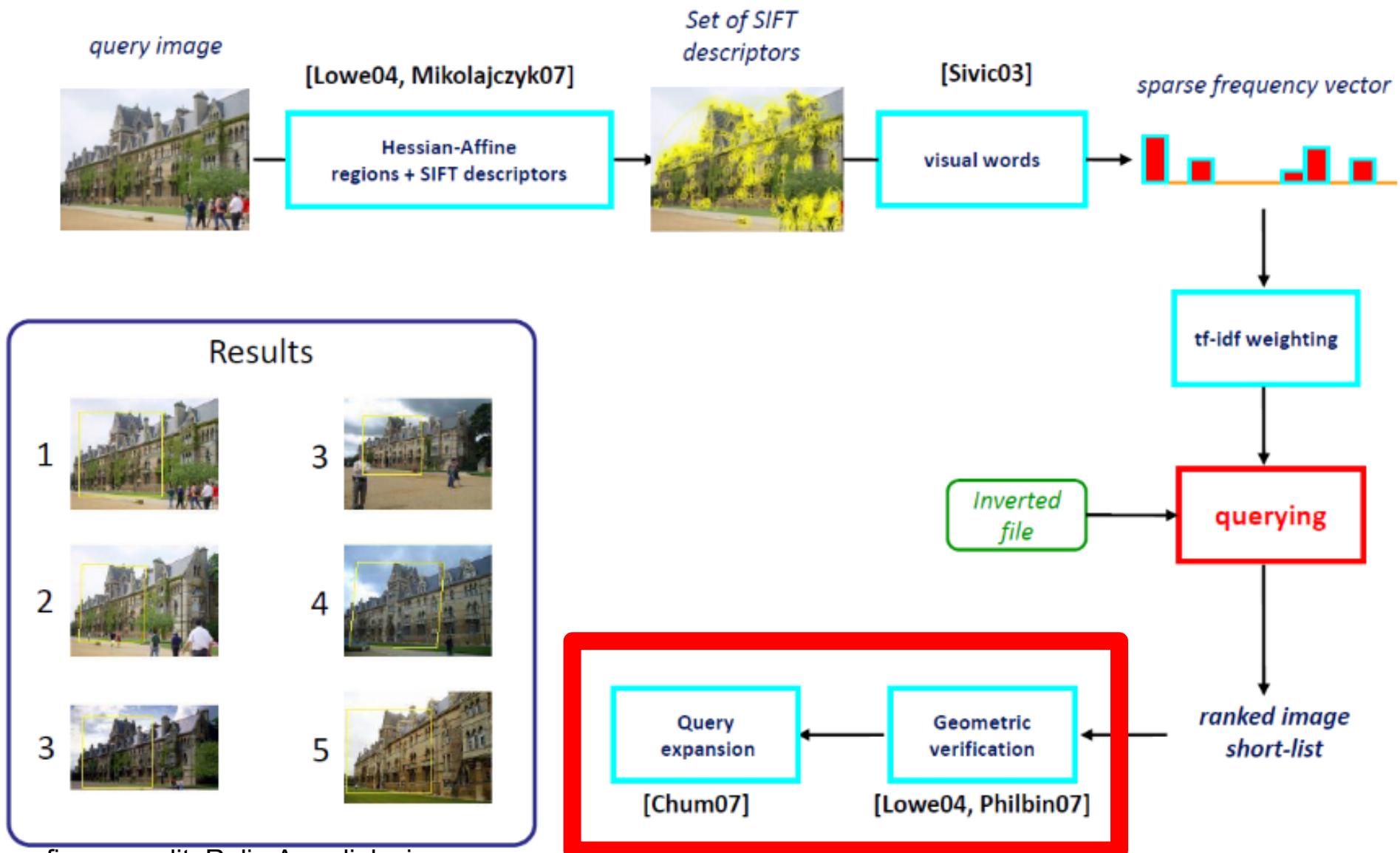
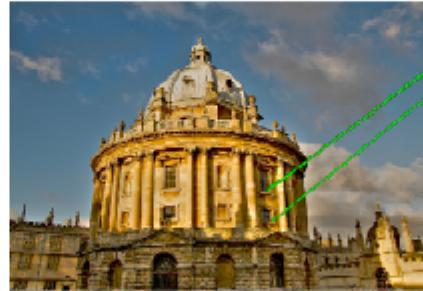
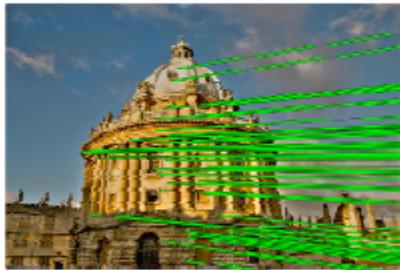


figure credit: Relja Arandjelovic

# Spatial Verification

It is vital for query-expansion that we do not...

- ...expand using false positives, or
- ...use features which occur in the result image, but not in the object of interest



# Spatial Verification

- Use hypothesize and verify procedure to estimate homography between query and target
- > 20 inliers = spatially verified result

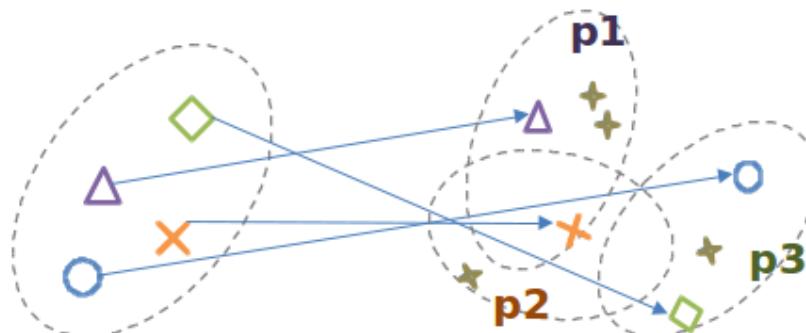


figure credit: Michael Isard

# Query Expansion

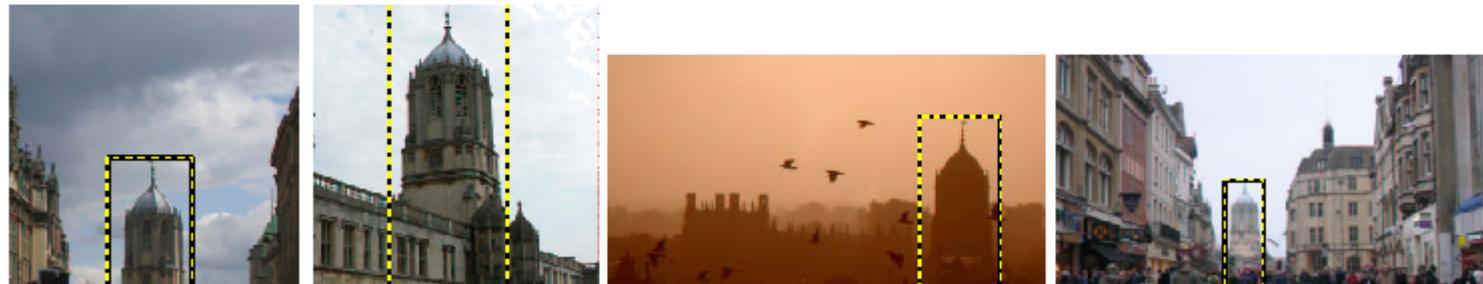
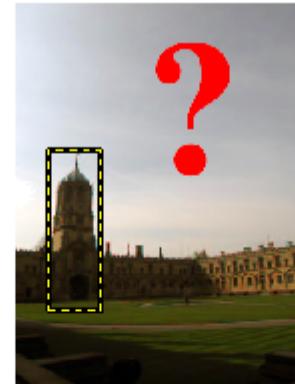
- Evaluate the original query  $Q_0$
- Construct latent model
- Issue new query/queries



# Generative Model

## Query Expansion Baseline

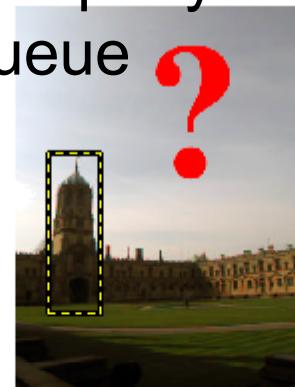
- Find top 5 (unverified) results from original query
- Average the term-frequency vectors
- Requery once
- Append these results



# Generative Model

## Transitive Closure Expansion

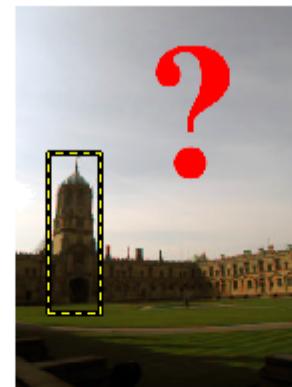
- Create a priority queue based on number of inliers
- Get top image in the queue
- Find region corresponding to original query
- Use this region to issue a new query
- Add new verified results to queue
- Repeat until queue is empty



# Generative Model

## Average Query Expansion

- Obtain top ( $m < 50$ ) verified results of original query
- Construct new query using average of these results
- Requery once
- Append these results



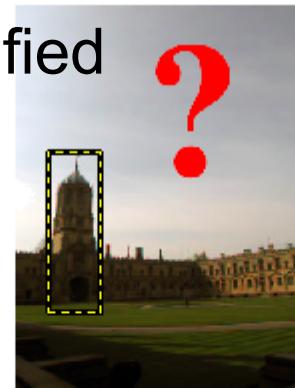
$$d_{\text{avg}} = \frac{1}{m+1} \left( d_0 + \sum_{i=1}^m d_i \right)$$



# Generative Model

## Recursive Avg Query Expansion

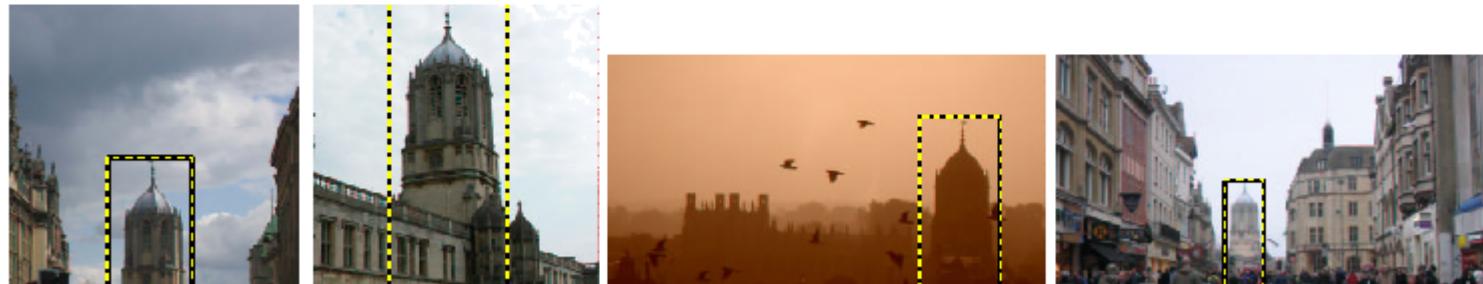
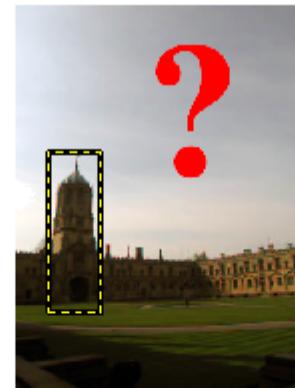
- Improvement of average query expansion method
- Recursively generate queries from all verified results returned so far
- Stop once 30 verified images are found or once no new images can be positively verified



# Generative Model

## Multi-resolution Expansion

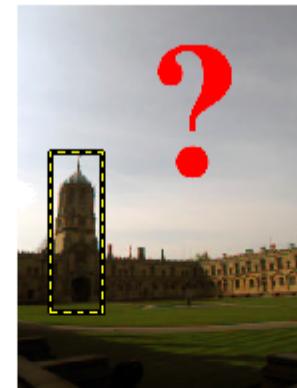
- For each verified result of original query calculate relative change in resolution required to project verified region onto query region



# Generative Model

## Multi-resolution Expansion

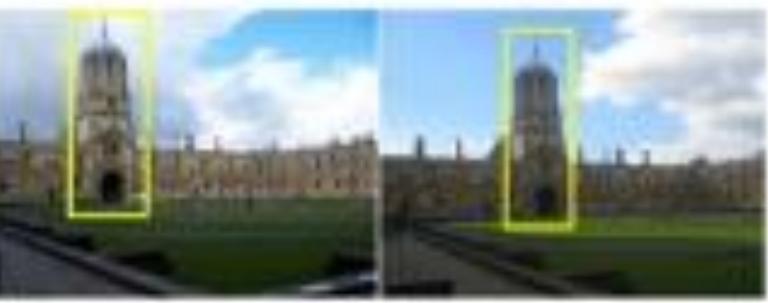
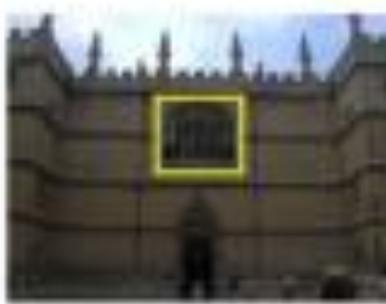
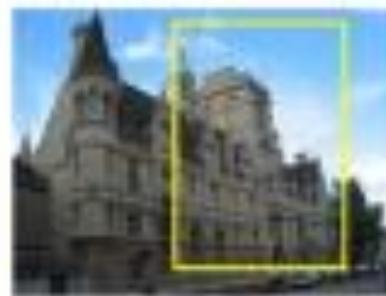
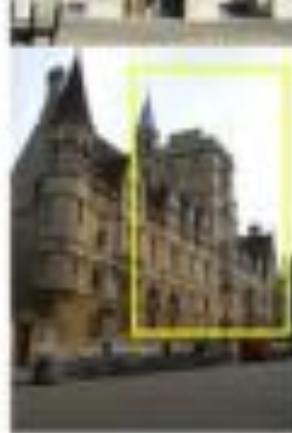
- Place results in 3 bands: $(0, \frac{4}{5})$ ,  $(\frac{2}{3}, \frac{3}{2})$ ,  $(\frac{5}{4}, \inf)$
- Construct average query for each band
- Execute independent queries
- Merge results



# Experiments

## Datasets:

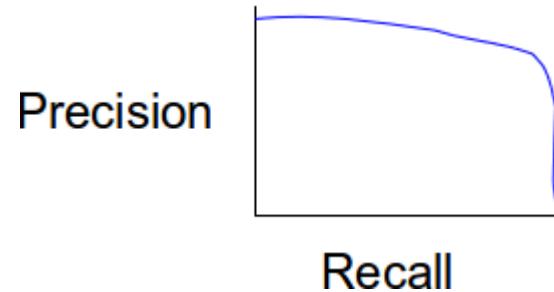
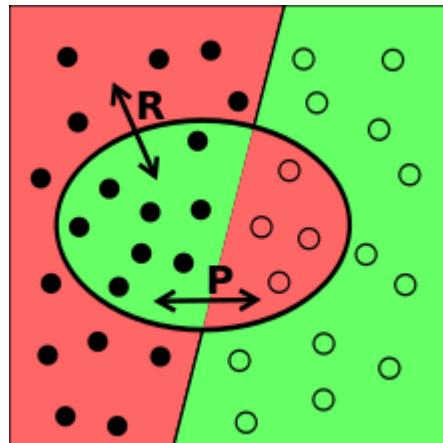
- Oxford dataset: 5k high res. with ground truth labels of Oxford famous landmarks
- Flickr1 (high-res popular tag)
- Flickr2 (medium-res popular tag)
- For experiment, using:
  - D1: Oxford+Flickr1
  - D2: Oxford+Flickr1+Flickr2



# Experiments

## Evaluation:

- Precision-Recall Curve



- Average Precision
  - area under the precision-recall curve

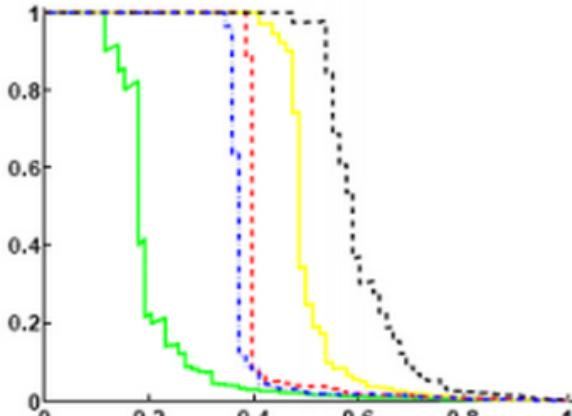
# Experiments



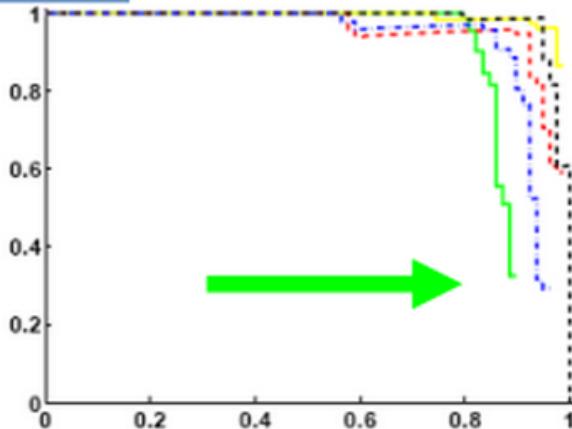
Query  
image



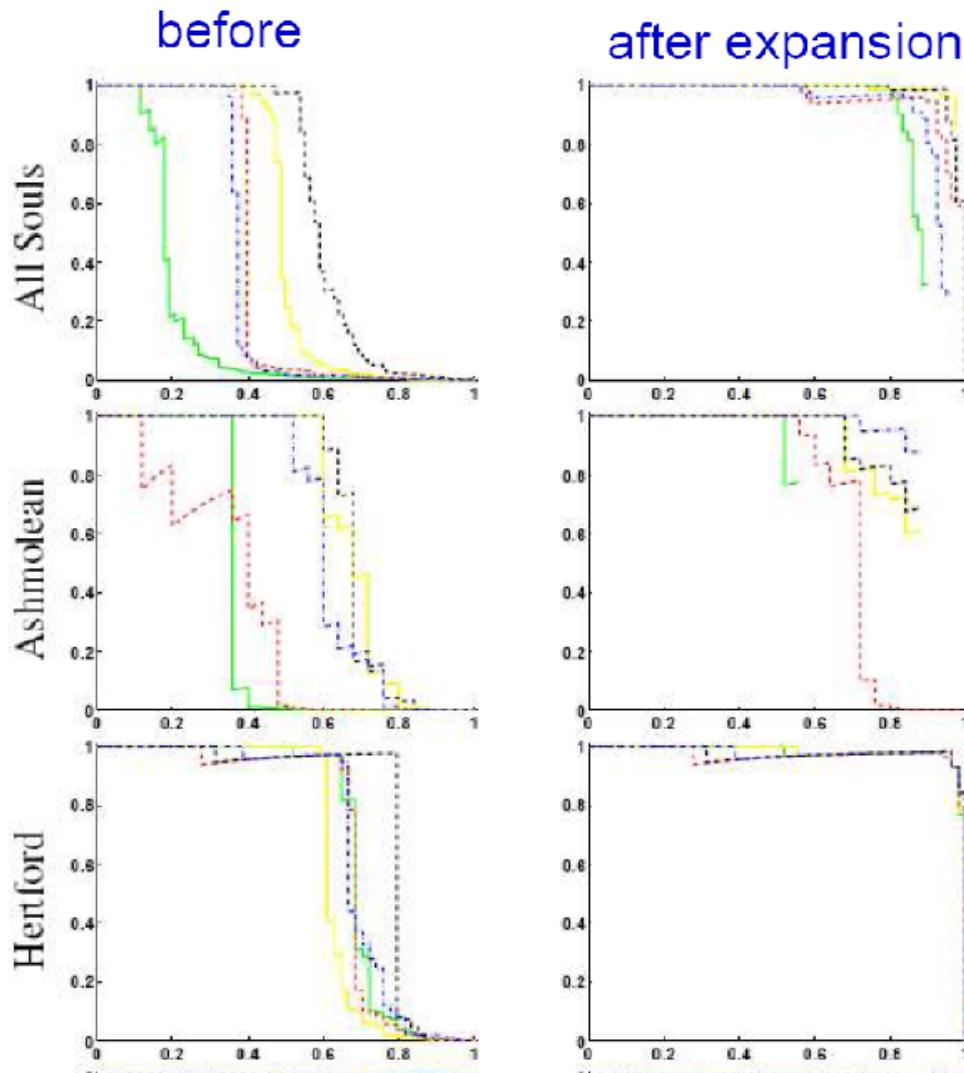
Original results (good)



Expanded results (better)



# Experiments



# Experiments

Ground truth			<i>Oxford + Flickr1</i> dataset						<i>Oxford + Flickr1 + Flickr2</i> dataset					
	OK	Junk	ori	qeb	trc	avg	rec	sca	ori	qeb	trc	avg	rec	sca
All Souls	78	111	41.9	49.7	85.0	76.1	85.9	<b>94.1</b>	32.8	36.9	80.5	66.3	73.9	<b>84.9</b>
Ashmolean	25	31	53.8	35.4	51.4	66.4	74.6	75.7	41.8	25.9	45.4	57.6	<b>68.2</b>	65.5
Balliol	12	18	50.4	52.4	44.2	63.9	<b>74.5</b>	71.2	40.1	39.4	39.6	55.5	<b>67.6</b>	60.0
Bodleian	24	30	42.3	47.4	49.3	<b>57.6</b>	48.6	53.3	32.3	36.9	43.5	<b>46.8</b>	43.8	44.9
Christ Church	78	133	53.7	36.3	56.2	63.1	<b>63.3</b>	63.1	52.6	18.9	55.2	<b>61.0</b>	57.4	57.7
Cornmarket	9	13	54.1	60.4	58.2	74.7	74.9	<b>83.1</b>	42.2	53.4	<b>56.0</b>	65.2	68.1	<b>74.9</b>
Hertford	24	31	69.8	74.4	77.4	89.9	90.3	<b>97.9</b>	64.7	70.7	75.8	87.7	87.7	<b>94.9</b>
Keble	7	11	79.3	59.6	64.1	90.2	<b>100</b>	97.2	55.0	15.6	57.3	<b>67.4</b>	65.8	65.0
Magdalen	54	103	9.5	6.9	25.2	28.3	<b>41.5</b>	33.2	5.4	0.2	16.9	15.7	31.3	26.1
Pitt Rivers	7	9	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	100	90.2	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Radcliffe Cam.	221	348	50.5	59.7	88.0	71.3	73.4	<b>91.9</b>	44.2	56.8	86.8	70.5	72.5	<b>91.3</b>
Total	539	838	55.0	52.9	63.5	71.1	75.2	<b>78.2</b>	46.5	40.5	59.7	63.1	67.0	<b>69.6</b>

ori = original query

qeb = query expansion baseline

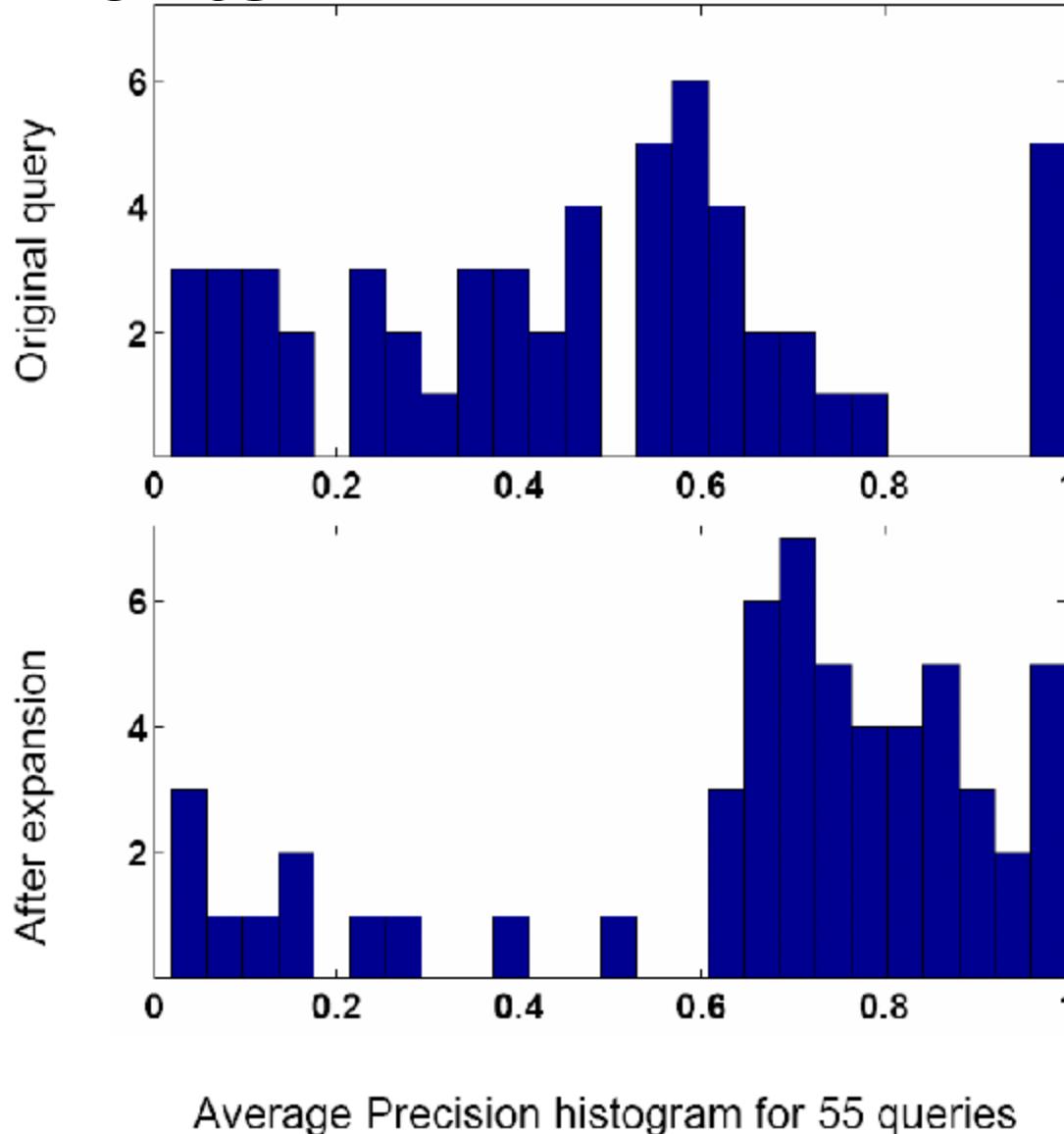
trc = transitive closure expansion

avg = average query expansion

rec = recursive average query expansion

sca = multiple image resolution expansion

# Experiments



# Experiments

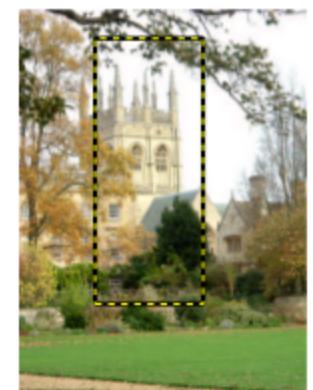
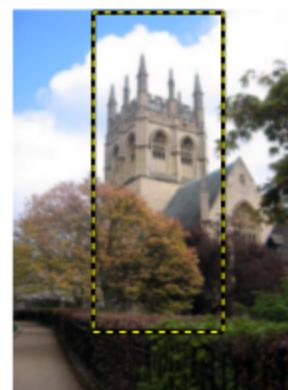
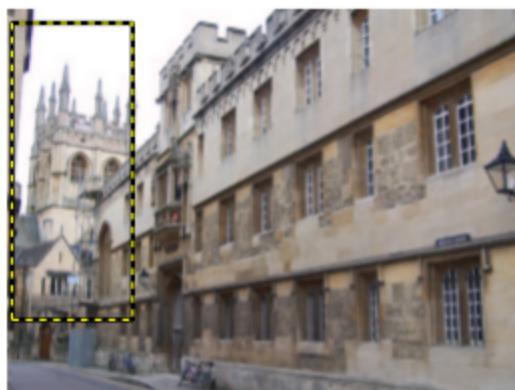
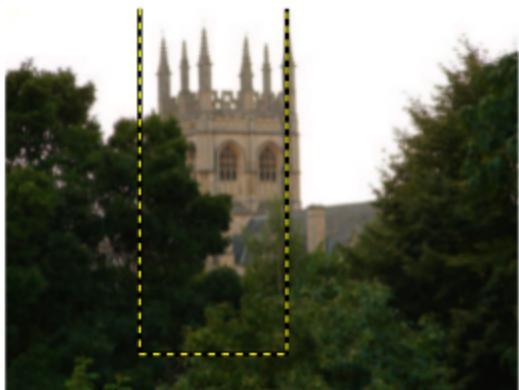
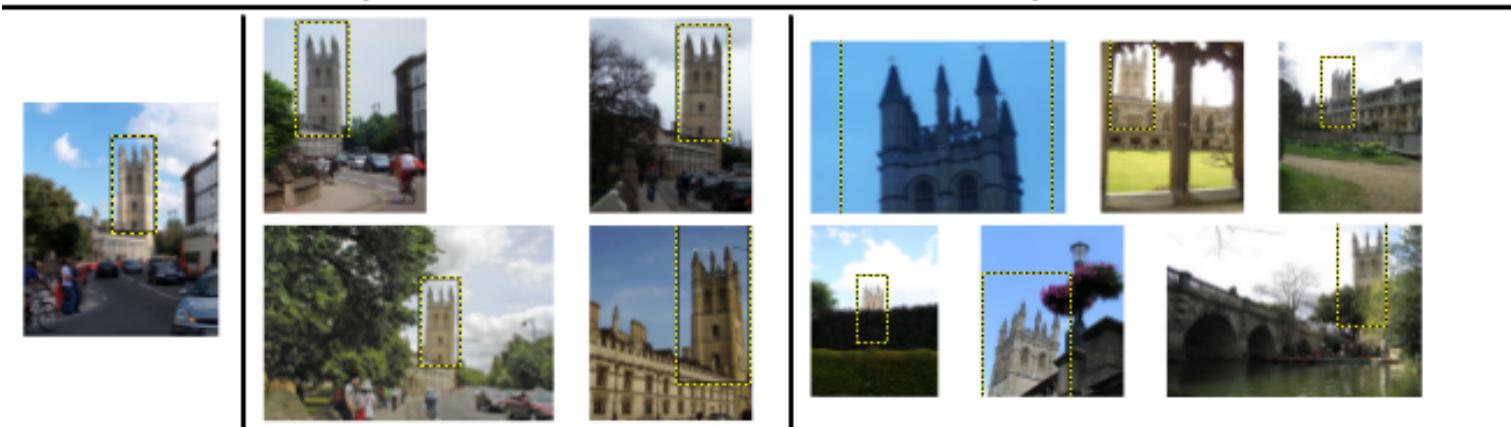
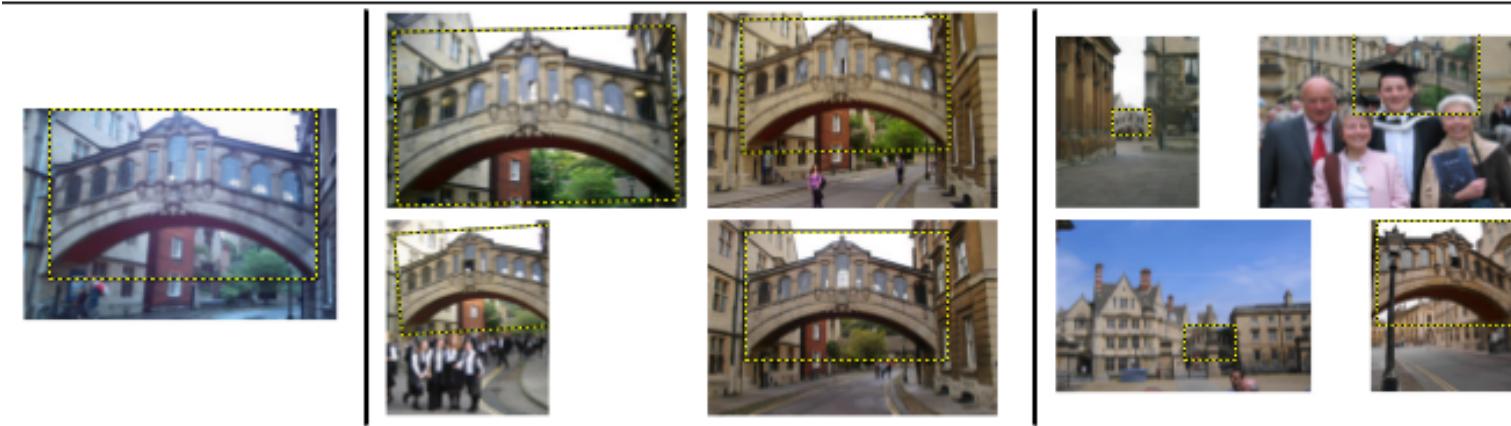
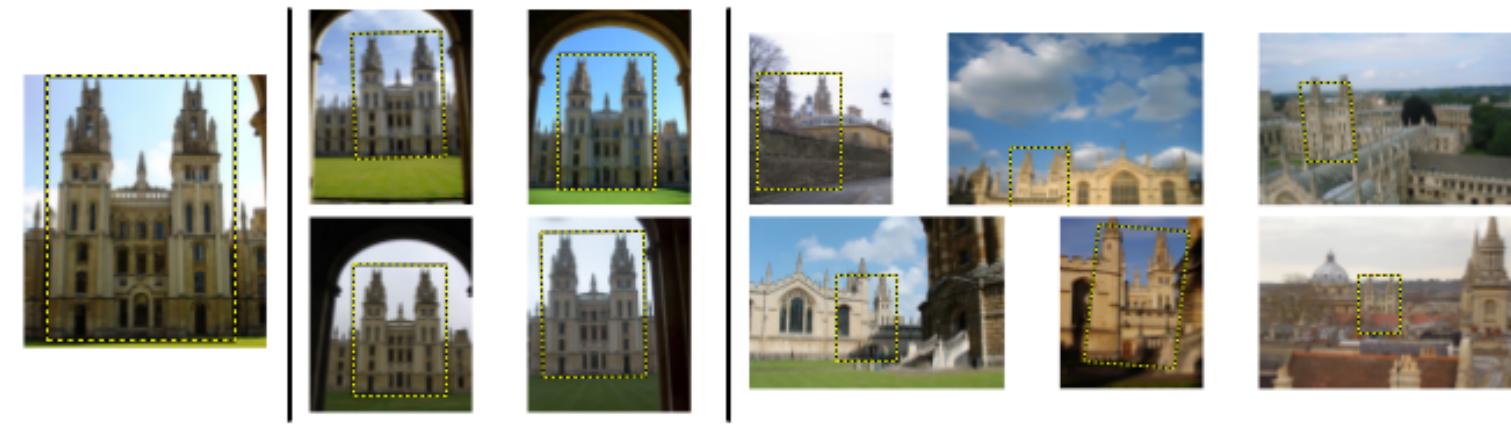


Figure 5. Some false positive images for Magdalen Tower query. The tower shown is actually part of Merton College chapel.

# Experiments



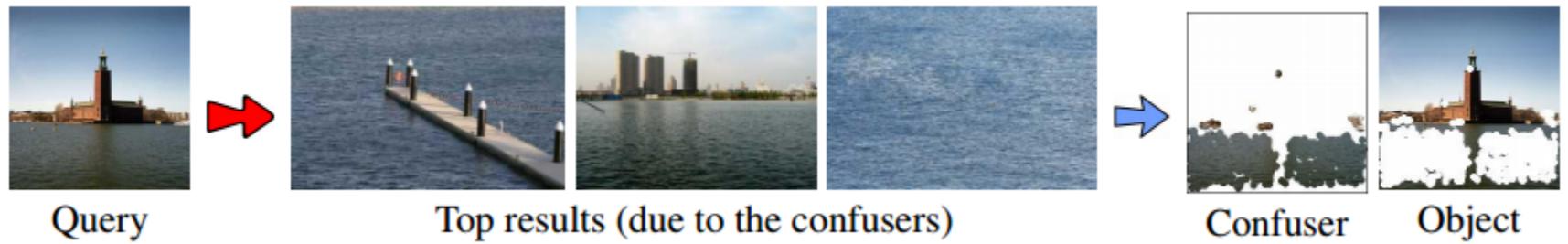
# Conclusion

Have successfully ported methods from text retrieval to the visual domain:

- Visual words enable posting lists for efficient retrieval of specific objects
- Spatial verification improves precision
- Query expansion improves recall, without drift

# Future Works

How to deal with confuser?



How to improve spatial verification?

- compare with previous verified images
- add spatial context outside the original region

# Question?

