

Multimedia Data Formats

Sascha Wernegger

June 13, 2016

Compression

- ImageMagick
- JXRLIB

Dataset Compression



Comparison



Estimate Compression Ratio

t

s : avg size = 391.659 bytes.

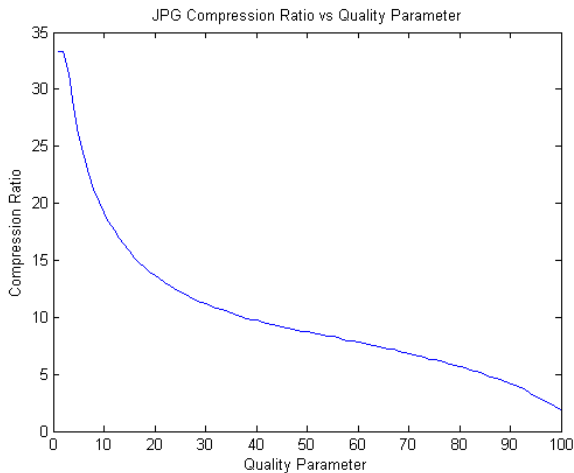
p : avg pixels = 765.969 pixel.

bpp : byte per pixel = 3.

r : avg raw size = $p * bpp = 2297908$.

e : estimated ratio = $r/s = 5.867$.

$$\text{Estimated Compressionratio} = \quad (1)$$



FeatureDetectors

- vlfeat
 - SIFT
 - PHOW(DSIFT)
- opencv
 - SURF
 - ORB

Benchmark

- VLBENCHMARK

Dataset

Oxford Buildings Dataset

- 5062 images
- compressed loss-less or with minimal loss in JPEG
- collected from Flickr by searching for particular Oxford landmarks
- manually annotated to generate ground truth for 11 different landmarks
- 5 queries per landmark
- total of 55 queries

Random sampled subsets to limit compression and benchmark speed!

Queries

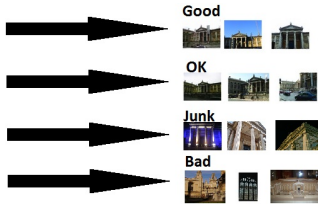
The query consists of a reference image and 4 query sets:

good A nice, clear picture of the object

ok More than 25% of the object is clearly visible.

junk Junk Less than 25% of the object is visible, or there are very high levels of occlusion or distortion.

bad Object not present



now similarity

between these image is measured

Generic Local Feature Extractor

Local Feature Frames

- search image for interest points
- define a frame for that point(points,circles,ellipses)

Descriptor

- compute descriptor using the frame

So we got n frames and n descriptors

Retrieval System

Ranking

- calculate KNN for the every reference descriptor
- vote with descriptor distance for the image
- normalize
- sort images after voting

Img 1 (good)



Img 2 (good)



Img 3 (ok)



Img 4 (bad)



Img 5 (good)



Img 6 (good)



Img 7 (good)



Img 8 (good)



Img 9 (good)



Img 10 (good)



Img 11 (good)



Img 12 (ok)



Img 13 (bad)



Img 14 (bad)



Img 15 (bad)



Img 16 (bad)



Img 17 (bad)



Img 18 (good)



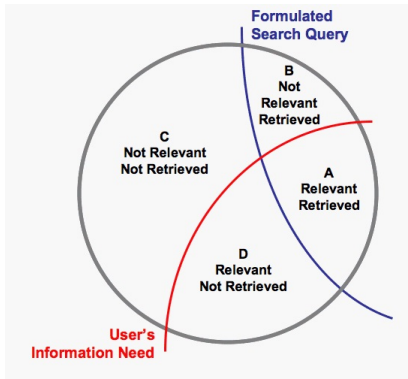
Img 19 (bad)



Img 20 (bad)



Recall Precision



Metrics for Measuring Classification Quality

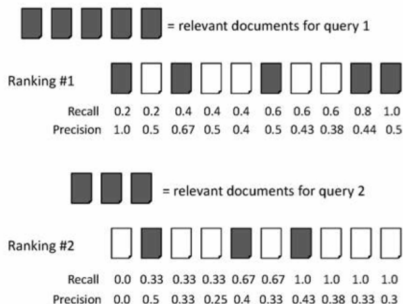
Point of View – Class 1

	Gold Class 1	Gold Class 2
Observed Class 1	TP	FP
Observed Class 2	FN	TN

$$\text{Precision} = \frac{tp}{tp + fp}$$

$$\text{Recall} = \frac{tp}{tp + fn}$$

Mean Average Precision



$$\text{average precision query 1} = (1.0 + 0.67 + 0.5 + 0.44 + 0.5) / 5 = 0.62$$

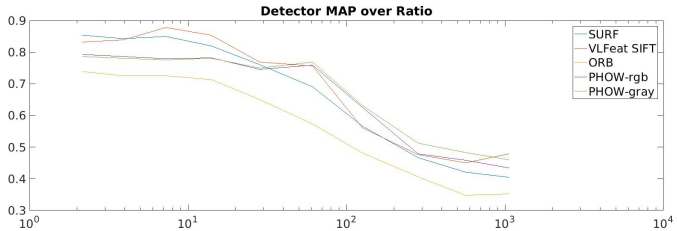
$$\text{average precision query 2} = (0.5 + 0.4 + 0.43) / 3 = 0.44$$

$$\text{mean average precision} = (0.62 + 0.44) / 2 = 0.53$$

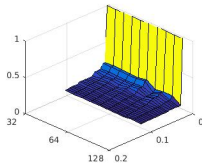
Mean Average Precision add

How use the four query classes

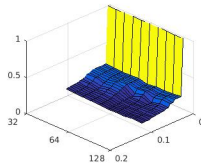
- good and ok images are relevant
- junk will be ignored
- bad will count as wrong



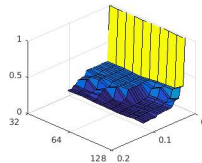
SURF



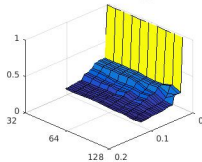
VLFeat SIFT



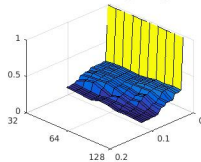
ORB



PHOW-rgb



PHOW-gray



Results

- plot of mAP over image file size
- plot query precision
- plot prc