Effects of Image Compression on Extracted Sped Up Robust Feature Quality

Damien Cerbelaud
Christopher Tsai
November 19, 2008
EE 398 – Image and Video Compression

Introduction

- Motivation: Smartphone applications (MAR)
- Capability: Networks, Recorders, Chips
- Rate reduction of JPEG images
  - Reduce size through resizing (downsampling/decimation)
  - Different JPEG-DCT quantization coefficients (new Q matrix)
  - Preserve features for image matching
  - Test: comparing compressed Query vs. compressed Database

Presentation Preview

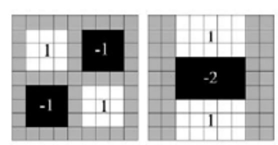
- Standard JPEG Quantization

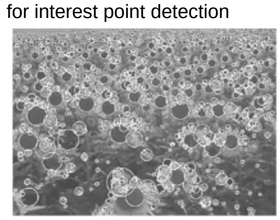
$$T = Q \cdot \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

- Reduction of image size – What is the best resolution?
  - 480 × 640 → 360 × 480 → 240 × 320 → 120 × 160 → Obliteration
- Effect of Q on quality of extracted SURF features
- Effect of Q on matching accuracy in database of 133 images
- Modifications to T for better features and match rate

SURF: Speeded Up Robust Features

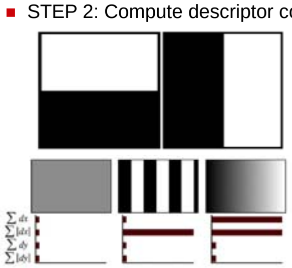
- Scale-invariant [sic], Rotation-invariant
- Two-step Algorithm [Bay, Tuytelaars, Van Gool]
  - Interest point detection through filtering/convolution
  - Feature classification using descriptor vector
- STEP 1: Fast discrete filtering for interest point detection





SURF: Speeded Up Robust Features

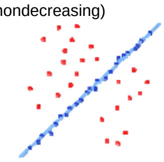
- STEP 2: Compute descriptor coordinates using Haar wavelets



- Variety of Haar wavelet sizes to more finely generate features

RANSAC: Random Sample Consensus

- Iterative method of finding best model for a set of data
  - 1.) *n* data points randomly selected to resolve free parameters
  - 2.) Generate 3D affine model on sample points
  - 3.) Test other points against this model
  - 4.) All points with small error are inliers
  - 5.) Compute average error of all inliers
  - 6.) Re-estimate model with inliers included
  - 7.) Repeat steps 3-6 until error is tolerably small (or nondecreasing)
- Advantages: Robustness – high accuracy
- Disadvantage: Unbounded convergence time
- SURF: Identified feature belongs to model
  - Some descriptors are wildly inaccurate
  - RANSAC eliminates some spurious features



Experimental Design: Resolution

- Uncompressed, resized, recompressed using JPEG
- JPEG with Quantization Matrix $T' = Q\ T$
- Vary resolution, Q factor, T-coefficient distribution
- Extract features and plot against image size
- Image size (Kbytes) as rate indicator
- Database vs. Database & Database vs. Query

JPEG-Recommended Perceptual Matrix

$$T = Q \cdot \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

- Matrix chosen for human visual perception
- DCT coefficients to which our eyes are most sensitive are quantized most finely (smaller step sizes 16, 11, 12, …)
- Asymmetric due to irregular monitor pixel sizes

A Brief Survey of Quantization Matrices

$T' = Q\ T$

- Wide range of Q: {0.1, 0.25, 0.5, 1, 2, 4, 8, 16, 32}
- **Fair comparison:** For each Q, all quantization matrices T normalized to the same geometric mean (equivalent rate).

$$M_{rec} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

Low Frequency Quantization Matrix

$T' = Q\ T$

- Wide range of Q: {0.1, 0.25, 0.5, 1, 2, 4, 8, 16, 32}
- **Fair comparison:** For each Q, all quantization matrices T normalized to the same geometric mean (equivalent rate).

$$M_{low} = \begin{bmatrix} 6 & 5 & 6 & 11 & 19 & 36 & 52 & 68 \\ 6 & 7 & 10 & 16 & 26 & 66 & 76 & 78 \\ 9 & 10 & 14 & 26 & 49 & 79 & 106 & 94 \\ 10 & 15 & 23 & 36 & 72 & 140 & 143 & 122 \\ 14 & 22 & 45 & 80 & 111 & 201 & 210 & 172 \\ 22 & 40 & 75 & 102 & 149 & 215 & 260 & 232 \\ 50 & 81 & 119 & 156 & 210 & 278 & 306 & 284 \\ 80 & 130 & 159 & 193 & 252 & 253 & 289 & 306 \end{bmatrix}$$

High Frequency Quantization Matrix

$T' = Q\ T$

- Wide range of Q: {0.1, 0.25, 0.5, 1, 2, 4, 8, 16, 32}
- **Fair comparison:** For each Q, all quantization matrices T normalized to the same geometric mean (equivalent rate).

$$M_{high} = \begin{bmatrix} 32 & 20 & 16 & 24 & 30 & 44 & 47 & 45 \\ 22 & 20 & 21 & 25 & 30 & 59 & 50 & 36 \\ 23 & 19 & 22 & 29 & 42 & 52 & 52 & 34 \\ 20 & 23 & 26 & 31 & 48 & 70 & 54 & 33 \\ 23 & 25 & 39 & 53 & 56 & 77 & 60 & 36 \\ 26 & 35 & 49 & 51 & 57 & 63 & 57 & 36 \\ 45 & 54 & 59 & 59 & 60 & 60 & 50 & 34 \\ 53 & 61 & 57 & 53 & 53 & 40 & 35 & 26 \end{bmatrix}$$

Image with No Compression

Image Compressed with Q = 0.1

Image Compressed with Q = 0.5

Image Compressed with Q = 1

Image Compressed with Q = 2

Image Compressed with Q = 4

Image Compressed with Q = 8

Image Compressed with Q = 16

Image Compressed with Q = 32

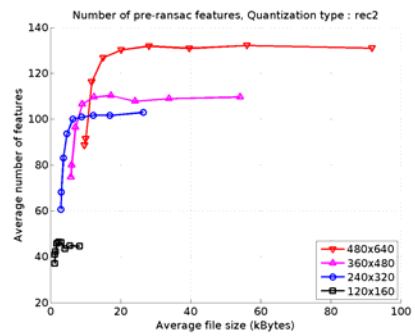
A Brief Survey of Resolutions

- The lower the quantization coefficient, the finer the detail

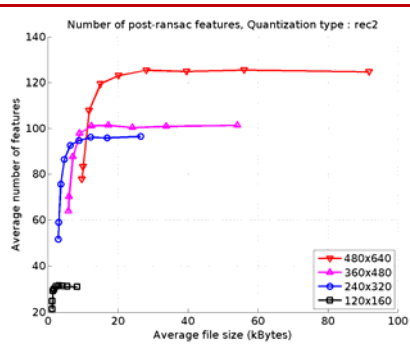
- Decimation decreases resolution, information
- High resolutions → too many minute, inessential features
- Low resolutions → too much blurring, key feature loss
- Compromise = Reduce rate, remove minute features
- GOAL: Find the ideal resolution for robust SURF extraction

- CONCLUSION: Effects of resizing are much more pronounced than the effects of varying quantization matrix

Effects of Resizing: Database v. Database



Effects of Resizing: Database v. Database

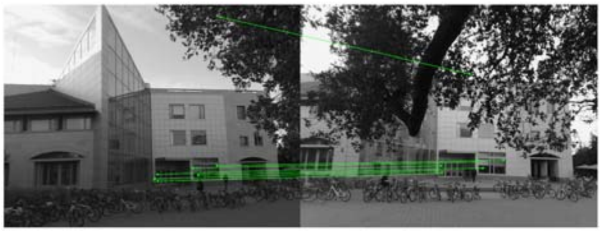


Effects of Resizing: Database v. Database

- More features detected as resolution increases
- Fewer features detected as resolution decreases
- Much more uniform/fair across features than quantization
- Q affects *type* of feature compressed: blobs v. edges
- Size affects number and type – removes minute features

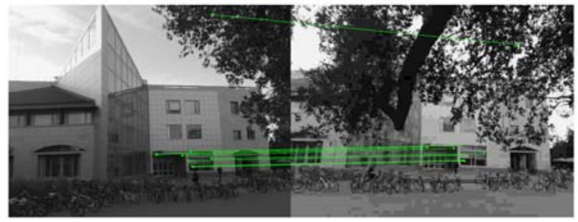
Target size s	Dimensions	Quantization factor
$\geq 11Kb$	480×640	$1 \leq q \leq 8$
$8Kb \leq s \leq 11Kb$	360×480	$q = 4$
$2Kb \leq s \leq 8Kb$	240×320	$2 \leq q \leq 32$
$s \leq 2Kb$	120×160	$1 \leq q \leq 32$

Query v. Database: Feature Reappearance



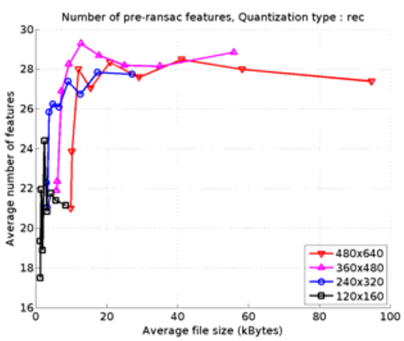
Q = 1

Feature Reappearance



Q = 8

Effects of Resizing: Query v. Database

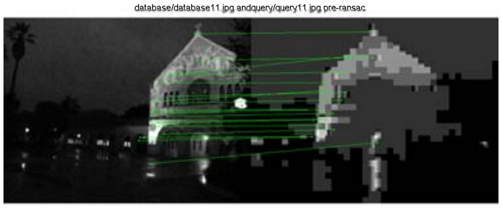


Effects of Resizing: Query v. Database

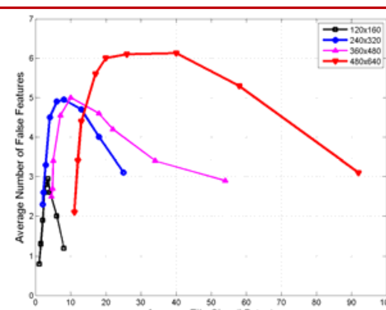
Effects of Resizing: Query v. Database

- Generally fewer features…approaching RANSAC minimum
- Quality and authenticity of features depreciated
- More false features (see diagonal lines)
  - Discordant locations
  - Nonsensical content (match made on similar contrast variations)

database/database11.jpg andquery/query11.jpg pre-ransac

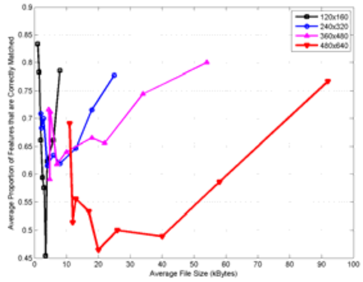


Effects of Resizing: False Features!



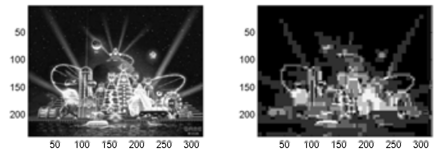
- Intermediate resolutions respond better to compression

Effects of Resizing: False Features!



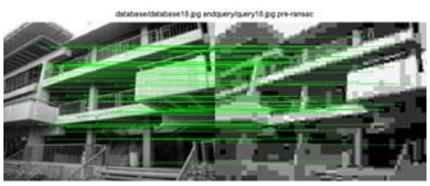
- Proportion used in RANSAC is suggestive of matching accuracy
- Intermediate resolutions again prevail over the extremes

Low Resolution Image Compression



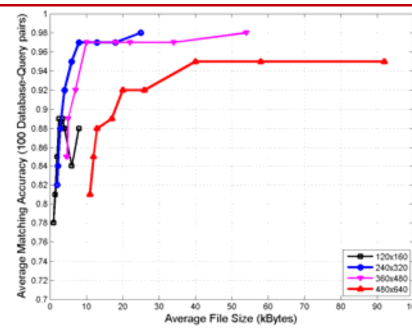
- Already suffering from dearth of pixels
- Even slightest quantization will blur/merge features
- Lost features are irrecoverable, not used in RANSAC
- Surviving features: **large areas** & **general structures**
- PROBLEM: Not enough features to build in RANSAC

High Resolution Image Compression



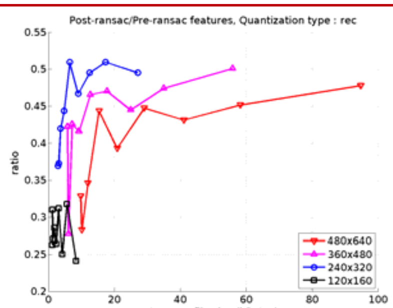
- More features → More spurious features
- Compression-induced blocking artifacts are features
- Surviving features: **strong edges** & **object detail**
- PROBLEM: High proportion of false features in RANSAC,
- No problem in fine quantization, but in mobile phones…

Matching Accuracy: Query v. Database



- 240 × 320 is optimal resolution for matching accuracy

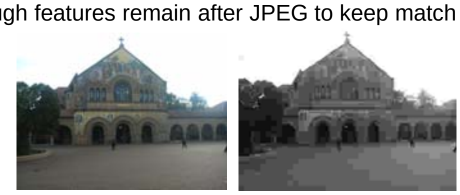
Post/Pre RANSAC Ratio: Query v. Database



- 240 × 320: Strong RANSAC robustness/feature preservation

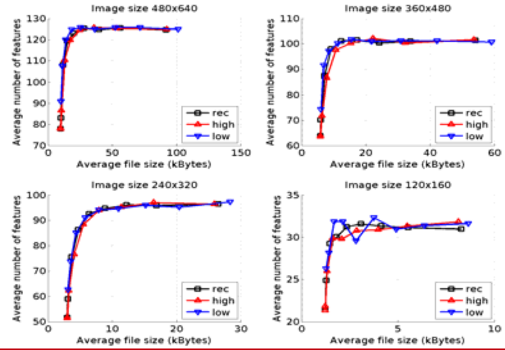
Effects of Image Resizing: Conclusion

- Robust common features are large background structures
- Can significantly downsample and still preserve these
- Intermediate res also free from minute “noise” details
- Enough features remain after JPEG to keep match



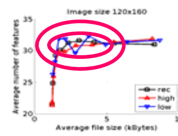
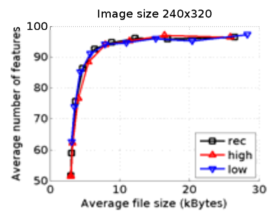
- Recommendation: **240 × 320** OR **360 × 480**

Modifying the Q Matrix: Database v. Database

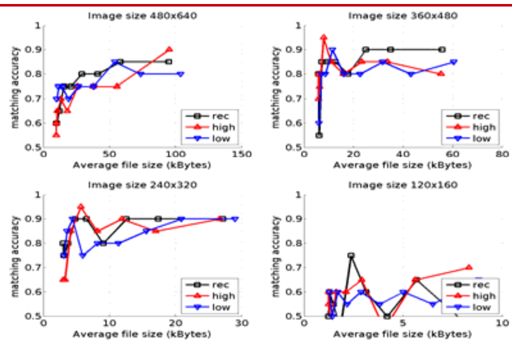


Modifying the Q Matrix: Database v. Database

- For all sizes, perturbing matrix changes little
- Low-Frequency-Enhancement performs best, but…
- …Gain in kilobytes never exceeds few percent
- Size unimportant
- Fluctuations indicate new spurious features

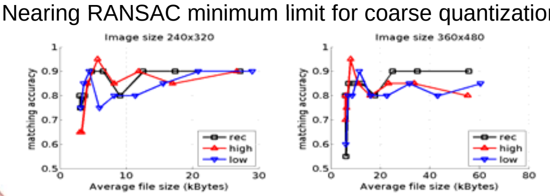



Modifying the Q Matrix: Query v. Database



Modifying the Q Matrix: Query v. Database

- For all sizes, perturbing matrix changes little
- Matching accuracy admits no "best choice"
- Fluctuation is result of having too few features per image and too few images in the test set
- Nearing RANSAC minimum limit for coarse quantization

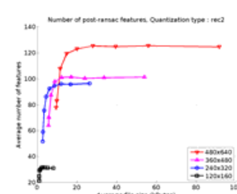


Conclusion

- Higher Q → Fewer Features, Lower Robustness, Accuracy
- Lower Resolution → Fewer Features, Larger-Scale Preserved
- Higher Resolution → Spurious Features, Detail Preserved
- Intermediate Balance of Features & Rate: 240 × 320, 480 × 640
- To reduce size while retaining features, decimation > quantization
- Scenario-dependent results

Presentation Review

- **DECIMATION – Enhances SURF Matching**
- Extreme reduction destroys feature information
- No reduction →"noise" features, unwieldy rates
- Compromise is ideal: Use 240 × 320
- **QUANTIZATION COEFFICIENTS**
- Coefficients alone: imperceptible
- With a **fixed size,** one can always achieve same performance with another quantization scheme and multiplication factor Q
- Common action on Q and size to get best tradeoff



A Glimpse into the Future

- Larger and different data sets → will smoothen fluctuations
- Experiments with more query images
- Probing the nullspace of SURF: where can we best compromise?
- Use of scale, angle, and position information from SURF extraction
- More systematic measurement of "**feature robustness**"
  - Which scales and positions of features are most resistant?
  - Which types of features respond most favorably to decimation?
  - Which types of features respond most favorably to quantization?

Bibliography

[1] Lowe, D. G., "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[2] Lowe, D. G., "Object Recognition from Local Scale-Invariant Features", Proc. Of the International Conference on Computer Vision, Corfu, Sept. 1999.

[3] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded Up Robust Features", in Proc. Ninth European Conference on Computer Vision, pp. 404-417, 2006.

[4] ITU-T and ISO/IEC JTC1, "Digital Compression and coding of continuous-tone still images", ISO/IEC 10918-1, ITU-T Recommendation T.81 (JPEG), Sept. 1992.

[5] G. Takacs, V. Chandrasekhar, N. Gelfand, Y. Xiong, W.-C. Chen, T. Bismpigiannis, R. Grzeszczuk, K. Pulli, and B. Girod, "Outdoors Augmented Reality on Mobile Phone using Loxel-Based Visual Feature Organization," submitted to IEEE Trans. Pattern Analysis and Machine Intelligence.

Acknowledgments

[1] Professors Bernd Girod and Markus Flierl

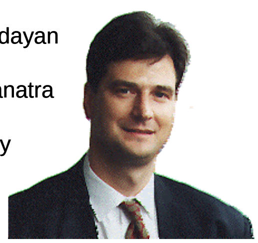
[2] Mentors David Chen and Vijay Chandrasekhar

[3] Teaching Assistant David Varodayan

[4] Peers June Zhang and Ivan Janatra

[5] SCIEN Lab, Stanford University

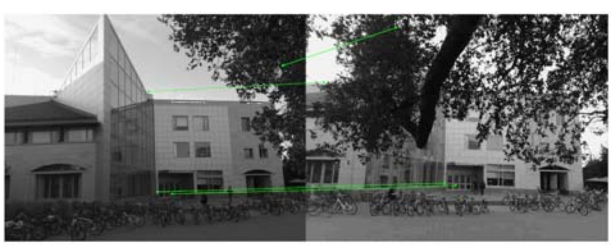
[6] Cristi Custuricu for JPEG in C



Feature Reappearance

Q = 2

Feature Reappearance

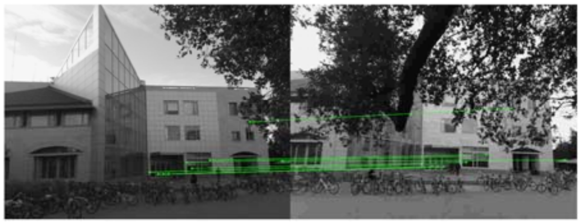


Q = 3

Feature Reappearance

Q = 4

Feature Reappearance



Q = 5

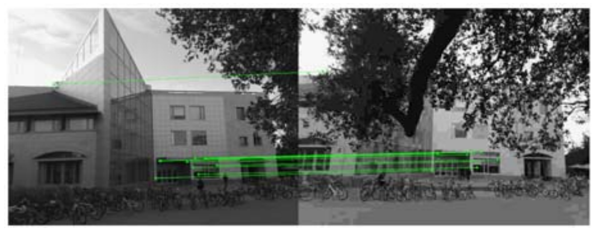
Feature Reappearance

Q = 6

Feature Reappearance

Q = 7

Feature Reappearance



Q = 8

Feature Reappearance

Q = 10