When and How can we do Image Retrieval

Forget about solving the general problem

General Strategy

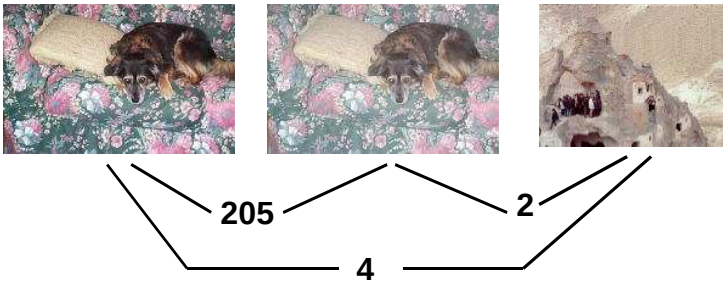
- Try to solve **specific CBIR** problems that satisfy certain feasibility criteria.
- Look for features beyond the obvious.

SIFT

- SIFT (Scale Invariant Feature Transform) key-points [Lowe 2004] seem helpful as a tool for 2D (but not 3D) matching.
  - When two **images** appear almost identical to the human observer they have a large number of matched key-points.
- SIFT key-point matching seems to be the best technique we have now.

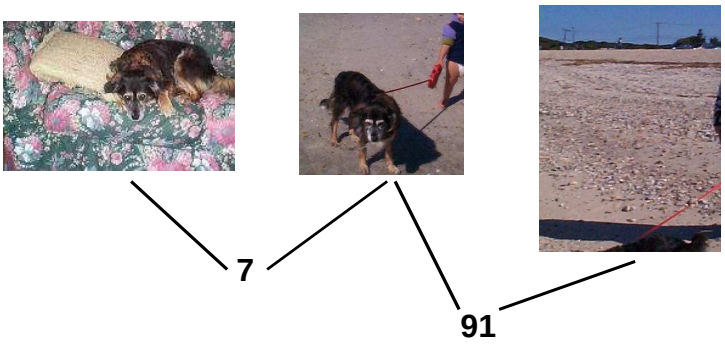
Number of matched SIFT Key-points – I
SIFT does well in 2D matching

Meeting a Challenge



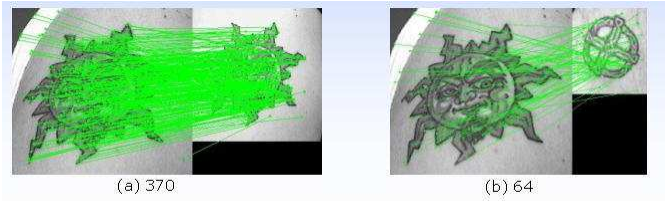
Results thanks to Ms. Jung-Eun Lee, a student of Prof. A.K. Jain at Michigan State University

Number of matched SIFT Key-points – II
SIFT fails in 3D object matching



Results thanks to Ms. Jung-Eun Lee, a student of Prof. A.K. Jain at Michigan State University

Application of SIFT to Biometrics



Numbers of matching points between (a) similar and (b) different images

Courtesy Prof. Anil K. Jain from his presentation on “Scars, Marks, and Tattoos: Soft Biometric for Victim and Suspect Identification” [LJJ08]

What is SIFT

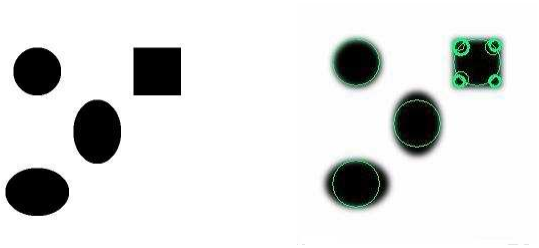
- Key points occur when the “support” of the blurring Gaussian kernel matches local curvature
- SIFT is not affected by variations either in size or in luminance.

Understanding SIFT - 1

Original Shape

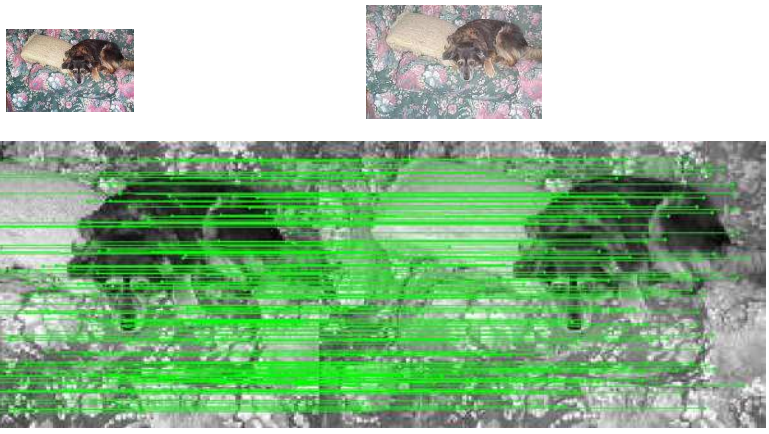
SIFT keypoints are centers of the green circles. On the left several keypoints are shown, on the right only the two strongest.

Understanding SIFT - 2

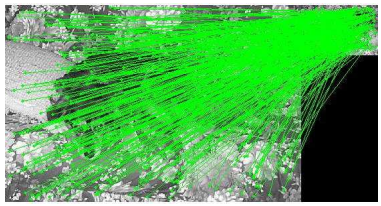


SIFT keypoints are centers of the green circles.

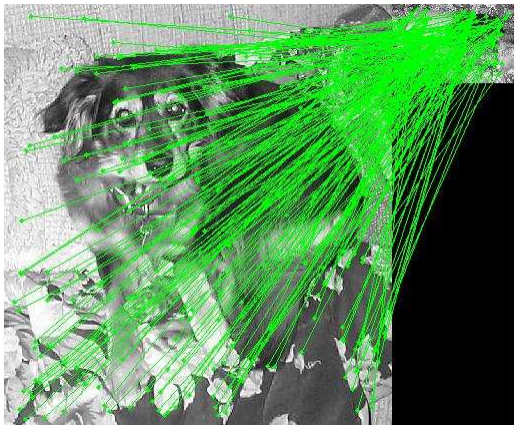
Examples of SIFT matching (AB)



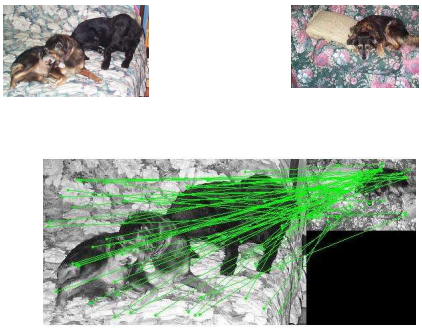
Examples of SIFT matching (FA)



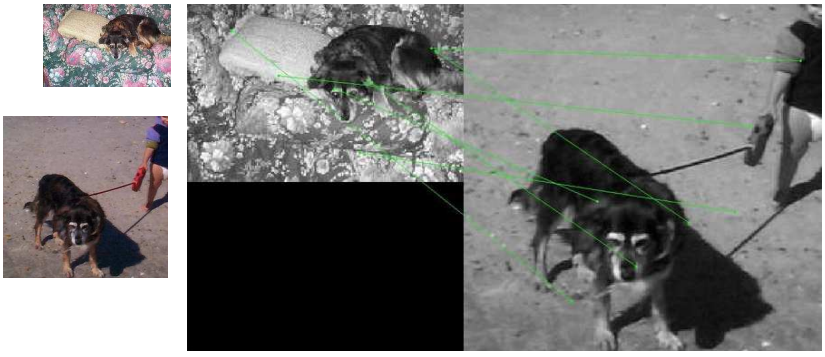
Examples of SIFT matching (GA)



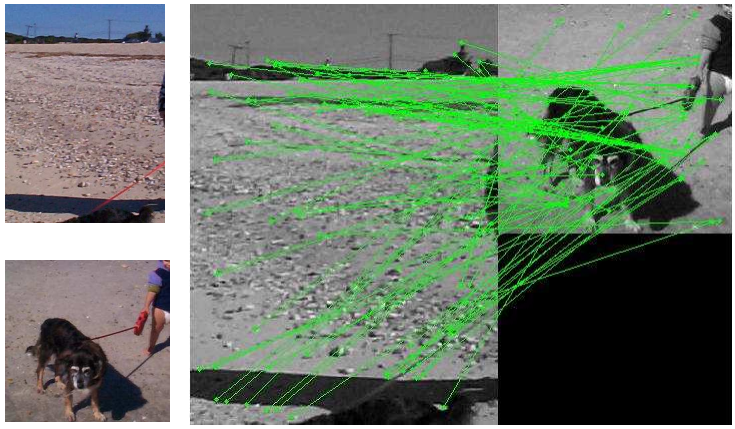
Examples of SIFT matching (HA)



Examples of SIFT matching (AJ)



Examples of SIFT matching (KJ)

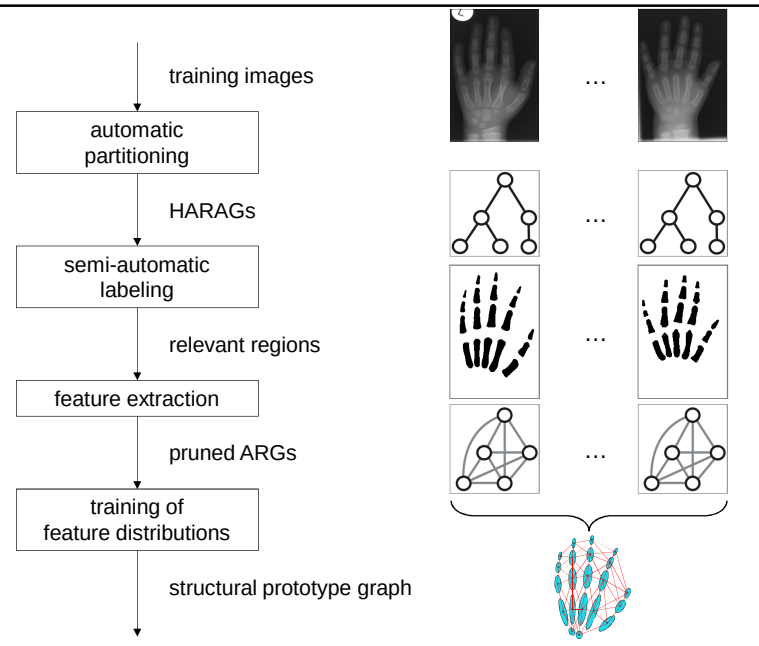


A Practical Application

- SIFT key points are the basis of a system for identification based on tattoos and scars.
- The system was developed at Prof. Anil Jain's laboratory at Michigan State University.
- It is used by various police departments.

HARAGs

- Matching of **H**ierarchical **A**ttributed **R**egion **A**djacency **G**raphs have been used in the matching of bones in hand X-rays by B. Fischer *et al* "Scene Analysis with Structural Prototypes for Content-Based Image Retrieval in Medicine", [FSGD08].
- The next slide is courtesy of Prof. Thomas M. Deserno of Aachen University, Germany.



Criteria for solvable CBIR problems

1. The mapping between semantics and image features is well defined.
2. Top level knowledge and/or context are known.
3. The accuracy requirements are well defined.
4. The computational requirements are well defined.
5. The matching of images requires careful scrutiny.

Mapping between Semantics and Image Features

- The representation by features should be possible so that only images with similar interpretation (and no others) are mapped onto the same set of features.
- Examples abound in the pattern recognition literature, including OCR and fingerprint recognition.
- If we deal with a specific application we may also be able to incorporate high level knowledge into the methodology.

Accuracy Requirements

- An application should have its own:
  - Absolute accuracy requirements.
  - Relative significance of false matches versus omitted true matches.
- A reminder from OCR
  - Older papers reported recognition rates of 97-98%, a rate useless in practice because it corresponds to over 50 errors per page. For a practical system the recognition rate should be at least 99.9% (2-3 error per page) and the *errors should be rejections (rather than substitutions)*. The tolerable rate for substitution errors is much lower.

Computational Requirements

- In many cases instant response is not needed.
  - In a medical application it may take well over an hour to produce an image, so waiting another hour to find matches in a database is not particularly onerous.
  - Auto-tagging can be done in the background, so speed is not critical.
- If a fast response is needed the database may be organized for fast retrieval.

The matching of images requires careful scrutiny

- Humans are not very good at careful scrutiny and it is likely that machines can match or even exceed human performance. (This has been an important factor in the success of automatic industrial inspection.)
- Medical, industrial, and biometric, amongst other applications, seem to fit in this category.

Where does Shopping CBIR fit?

- Users are supposed to submit a picture of a product and find through CBIR web sites where that product is sold.
- If the product is broadly defined, or if the list to be searched is small, CBIR may succeed, otherwise it will fail for the same reasons as Veggie-Vision.
  - Veggie-Vision could tell tomatoes from eggplants but not organic tomatoes from conventional tomatoes, so it never made it to the check-out counter.
- **The devil is in the details!**

Concluding Suggestions

- We should focus on matching 2D images rather than 3D scenes to avoid issues of pose, viewpoint, and large variations in scale.
- There are plenty of challenging applications needing 2D image matching: Medical, Biometric, Forensic, Industrial, Security, etc.
- It is better to really solve a special case of CBIR than pretend to solve the general CBIR problem.