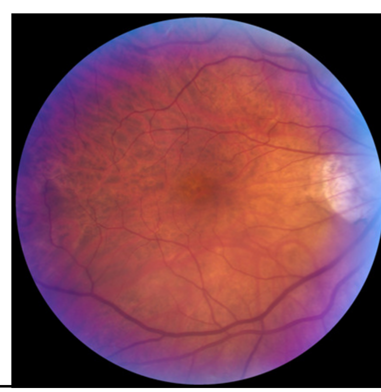
Project Paper Seminar

Vincent Ng

March 10th, 2011

Retina Project Overview

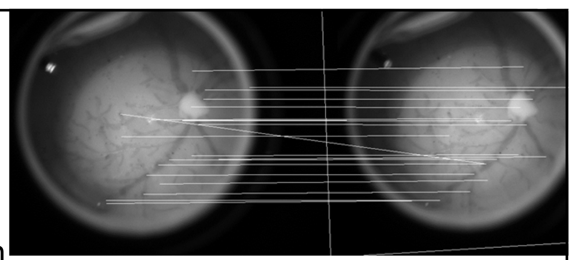
- Registration
  - Preoperative images to intraoperative images
  - Overlay of landmarks on live microscopic feed
- Requirements
  - Image Matching/Tracking



Paper Selection and Relevance

- Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool, ***"SURF: Speeded Up Robust Features",*** Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346--359, 2008
- SURF – algorithm for feature detection and descriptor generation

Background



- Computer Vision
  - Object recognition
  - Video tracking
- ***Finding point correspondence between 2 images***
  - Feature detection, feature descriptor, feature matching
  - SIFT (Scale Invariant Feature Transform)
    - Slow for live video implementation
    - "Predecessor" / Influence
    - Hessian (location), Laplacian (scale) – approx via Difference of Gaussians

Problems

- Feature Detector
  - Repeatability
  - Robust
  - Distinct
  - Scale, rotation invariant
- Feature Descriptor
  - Robust to noise, detection displacement, geometric/photometric deformations
  - Fewer dimensions = faster, but less distinct

SURF Theory: Feature Detector

- Feature Detector
  - Corners, blobs, T-junction
  - Integral Images
  - Hessian-matrix (choose where determinant is maximum)
  - Approximation
    - "Fast Hessian detector"
    - Small loss in repeatability

Given a point $\mathbf{x} = (x, y)$ in an image I, the Hessian matrix $\mathcal{H}(\mathbf{x}, \sigma)$ in $\mathbf{x}$ at scale σ is defined as follows

$$\mathcal{H}(\mathbf{x}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) \\ L_{xy}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) \end{bmatrix}, \qquad (2)$$

where $L_{xx}(\mathbf{x}, \sigma)$ is the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ with the image I in point $\mathbf{x}$, and similarly for $L_{xy}(\mathbf{x}, \sigma)$ and $L_{yy}(\mathbf{x}, \sigma)$.

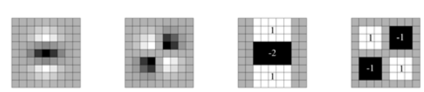


Fig. 2. Left to right: the (discretised and cropped) Gaussian second order partial derivative in y- (L_{yy}) and xy-direction (L_{xy}), respectively; our approximation for the second order Gaussian partial derivative in y- (D_{yy}) and xy-direction (D_{xy}). The grey regions are equal to zero.

Images and formulas taken from SURF paper

SURF Theory: Feature Detector

- Feature Detector
  - Scale Invariant

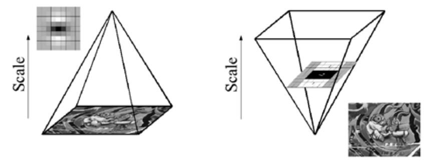


Fig. 4. Instead of iteratively reducing the image size (left), the use of integral images allows the up-scaling of the filter at constant cost (right).

Images and formulas taken from SURF paper

SURF Theory: Feature Descriptor

- Feature Descriptor
  - Describing smaller-scale features within interest neighborhood
  - Closely mimics SIFT
  - Rotation invariant
  - Haar Wavelet Responses

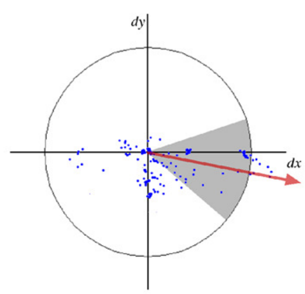


Fig. 10. Orientation assignment: A sliding orientation window of size $\frac{\pi}{3}$ detects the dominant orientation of the Gaussian weighted Haar wavelet responses at every sample point within a circular neighbourhood around the interest point.

Images and formulas taken from SURF paper

SURF Theory: Feature Descriptor

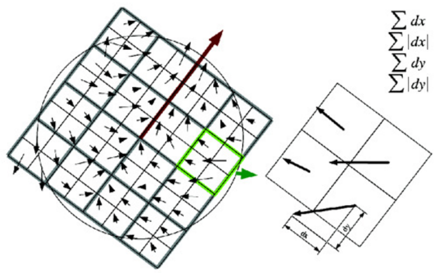


Fig. 12. To build the descriptor, an oriented quadratic grid with 4×4 square sub-regions is laid over the interest point (left). For each square, the wavelet responses are computed. The 2×2 sub-divisions of each square correspond to the actual fields of the descriptor. These are the sums dx, $|dx|$, dy, and $|dy|$, computed relatively to the orientation of the grid (right).

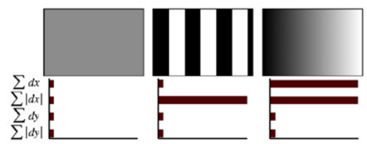


Fig. 13. The descriptor entries of a sub-region represent the nature of the underlying intensity pattern. Left: In case of a homogeneous region, all values are relatively low. Middle: In presence of frequencies in x direction, the value of $\sum |d_x|$ is high, but all others remain low. If the intensity is gradually increasing in x direction, both values $\sum d_x$ and $\sum |d_x|$ are high.

Images and formulas taken from SURF paper

SURF Theory: Feature Matching

- Feature Matching
  - Sign of Laplacian (trace of Hessian)

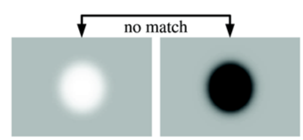
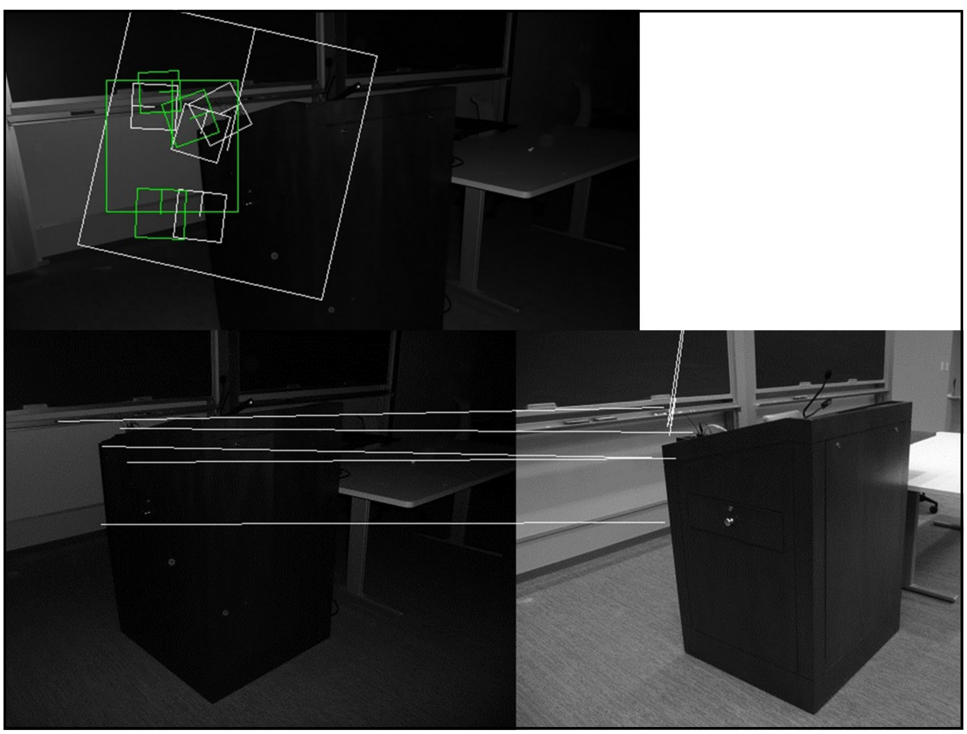


Fig. 15. If the contrast between two interest points is different (dark on light background vs. light on dark background), the candidate is not considered a valuable match.

Images and formulas taken from SURF paper



Experiment

- Comparison (speed, repeatability, reliability)
- Camera calibration for 3D reconstruction
- Object recognition experiment

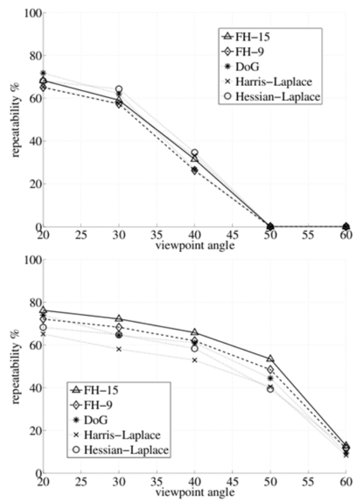
Key Result – Feature Detection, Speed

detector	threshold	nb of points	comp. time (ms)
FH-15	60000	1813	160
FH-9	50000	1411	70
Hessian-Laplace	1000	1979	700
Harris-Laplace	2500	1664	2100
DoG	default	1520	400

Table 1
Thresholds, number of detected points and calculation time for the detectors in our comparison. (First image of Graffiti scene, 800×640)

Images and formulas taken from SURF paper

Key Result – Feature Detection, Repeatability



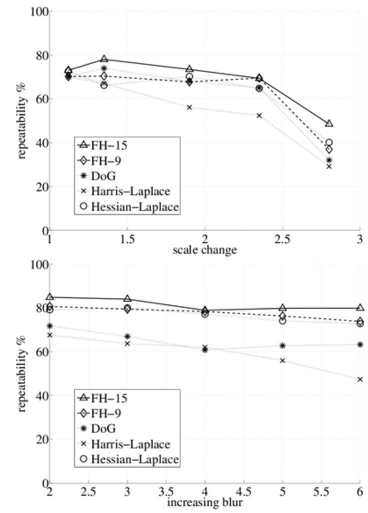
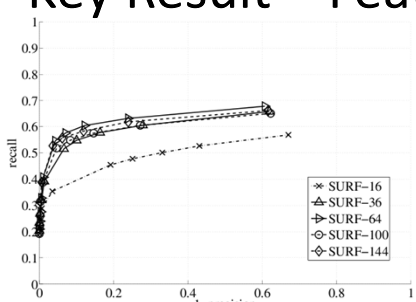


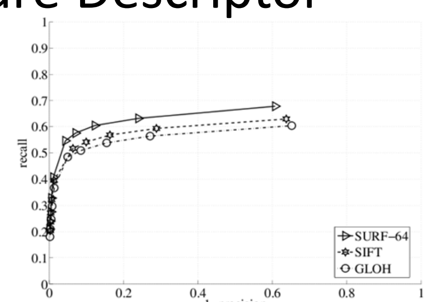
Fig. 16. Repeatability score for the Graffiti (top) and Wall (bottom) sequence (Viewpoint Change).

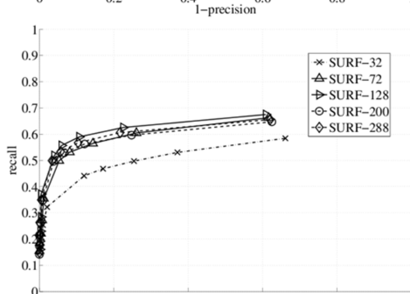
Fig. 17. Repeatability score for the Boat (top) and Bikes (bottom) sequence (Scale Change, Image Blur).

Images and formulas taken from SURF paper

Key Result – Feature Descriptor







Recall precision for nearest-neighbor ratio matching

Images and formulas taken from SURF paper

Critique

Good

- Described previous related work/algorithm
- Extended SIRF/other related detector-descriptors
- Justification of theory
- Balanced Speed vs. Performance
- Relevance to other areas of research

Bad

- Faster, faster, faster –> In-Depth Analysis of Performance?
- What ways is SURF better/worse than SIFT?
  - Need larger sample testing

Next Steps

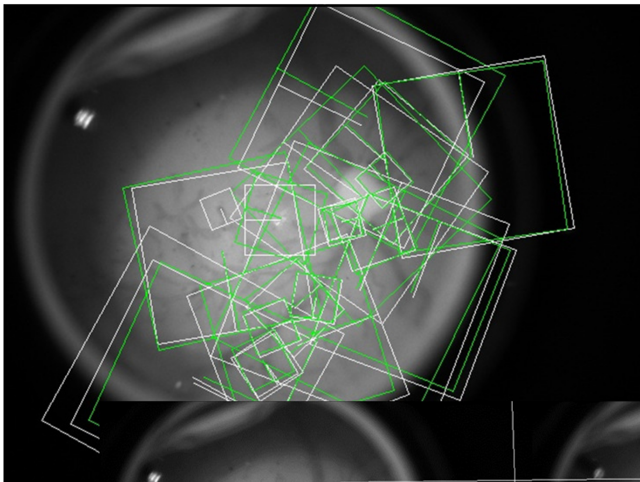
- GPU Parallelization - Done
- Feature Descriptor : 64 vs. 128 elements
- Testing on image distortions
- Implications for image matching by comparing descriptors
- No color info used

Assessment / Relevance to Project

- Speed vs. Accuracy
  - Live tracking?
- Robustness for current data
  - Scale invariant
- Customizable – Hessian threshold, description size
- Feature matching?

Conclusion

- Balance between Speed vs. accuracy
  - Approximation, reduction in operations
- Valid replacement for SIFT



Questions?