

Performance Evaluation

Ardy Goshtasby

*Wright State University and
Image Fusion Systems Research*

Criteria for evaluation

- **Accuracy:** The difference between estimated and true values.
- **Reliability:** The rate of success of an algorithm.
- **Robustness:** The degree of stability of the accuracy or reliability of an algorithm under variations in its input parameters.
- **Computational complexity:** Speed

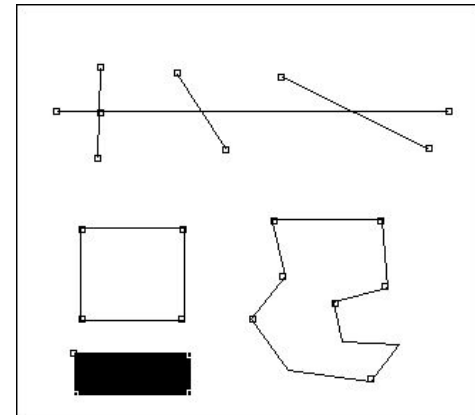
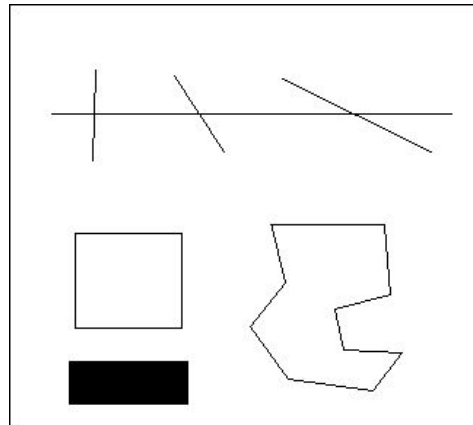
The performance of an algorithm depends on the performances of its components.

Feature selection accuracy

- Maximum, mean, or root-mean-squared distance between positions of obtained point features and true point features.

An example: Find points where edges or lines meet.

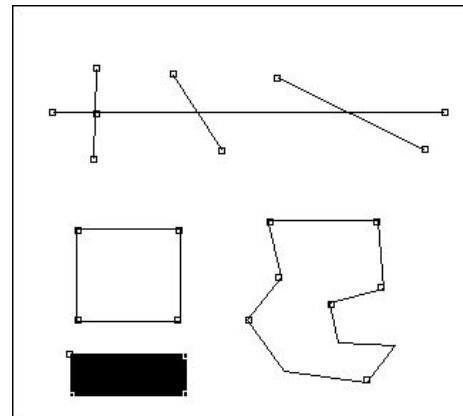
MAX=1.4 pixels
MEAN=0.6 pixels



Feature selection reliability

- **True positive probability:** The rate of detecting the features.
- **False negative probability:** The rate of missing the features.
- **False positive probability:** The rate of detecting false features.

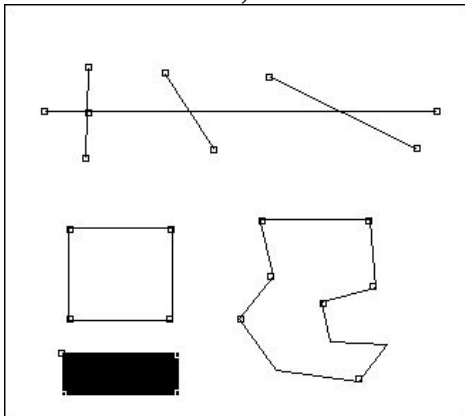
TPP: 16/21
FPP: 8/21
FNP: 5/21



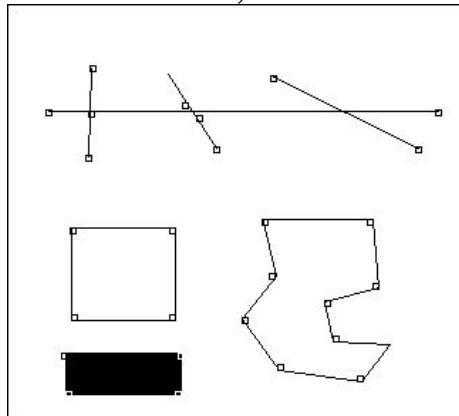
Feature selection robustness

- The rate of change in accuracy or reliability of detected features as resolution, noise, etc., in input are changed.

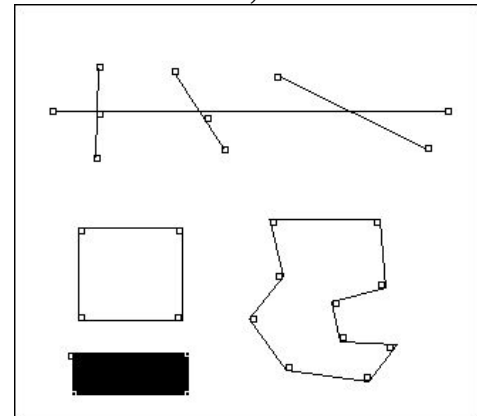
TPP: 16/21; FPP: 8/21



TPP: 20/21; FPP: 8/21



TPP: 20/21; FPP: 8/21



Corners detected at different resolutions.

Feature correspondence performance

- **Accuracy:** A correspondence is either correct or incorrect; therefore, speaking about the accuracy of a feature correspondence algorithm is meaningless.
- **Reliability:**
 - **True positive probability:** The rate of finding the correct correspondences.
 - **False positive probability:** The rate of finding incorrect correspondences.
 - **False negative probability:** The rate of missing the correct correspondences.
- **Robustness:** The rate of change of the reliability under variations in input parameters.

Transformation function performance

- Accuracy:** Given that all correspondences are correct and exact, the maximum, mean, or root-mean-squared distance between corresponding points in the images after registration describes the accuracy.
- If the images have only linear geometric differences, the transformation should be linear or one that is very close to it.
 - A transformation for non-rigid registration should be able to describe complex local geometric differences between images.

Transformation function performance

- **Reliability:** A transformation function always registers two images, therefore, speaking about the reliability of a transformation is meaningless.
- **Robustness:** The stability in accuracy of a transformation function under various types of geometric difference between images, density of point features, organization of point features, and noise in the correspondences.

Registration performance

- **Accuracy:** This depends on the accuracies of the point features and the transformation function.
- **Reliability:** This depends on the reliabilities of the selected features and the correspondences.
- **Robustness:** Stability in the accuracy or reliability of registration as different kinds of images are used.

Computational complexity

- The highest computational complexity of the components of the algorithm.
- The computational complexity of registration algorithms is required to be a linear function of the image dimensions for them to be useful in interactive applications.
- Algorithms with computational complexities higher than order two of the image dimensions are not very useful.

Validation data

- **Simulation or synthetic data:** Realistic images are used as the reference images. Sensed images are generated from the reference images using known transformation functions.
- **Phantom data:** Images of a synthetic scene that models the scene to be imaged are used.
- **Fiducial landmarks:** Markers that are easily identifiable in images are planted in the scene and images of the scene are obtained.
- **Natural landmarks:** Easily identifiable scene landmarks such as road intersections, centers of gravity of lakes, branching points of vessels.
- **Differential markers or corner points:** Points with predefined mathematical properties.

References

1. P. Jannin, J. M. Fitzpatrick, D. J. Hawkes, X. Pennec, R. Shahidi, and M. W. Vannier, Validation of medical image processing in image-guided therapy, *IEEE Trans. Medical Imaging*, **21**(12):1445–1449 (2002).
2. J. A. Schnabel, C. Tanner, A. D. Castellano-Smith, A. Degenhard, M. O. Leach, D. R. Hose, D. L. G. Hill, and D. J. Hawkes, Validation of nonrigid image registration using finite-element methods: Application to breast MR images, *IEEE Trans. Medical Imaging*, **22**(5):238–247 (2003).
3. J. M. Fitzpatrick, D. L. G. Hill, Y. Shyr, J. West, C. Studholme, and C. R. Maurer, Jr., Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain, *IEEE Transactions on Medical Imaging*, **17**(4):571–585 (1998).
4. C. R. Maurer, Jr., J. M. Fitzpatrick, M. Y. Wang, R. L. Galloway, J., R. M. Maciunas, and G. S. Allen, Registration of head volume images using implantable fiducial markers, *IEEE Trans. Medical Imaging*, **16**(4):447–462 (1997).
5. J. C. H. Wong, C. Studhomes, D. J. Hawkes, and M. N. Maisey, Evaluation of the limits of visual detection of image misregistration in a brain fluorine-18 fluorodeoxyglucose PET-MRI study, *European J. Nuclear Medicine*, **24**:642–650 (1997).