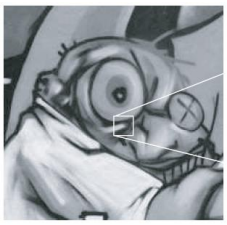
Detector repeatability study.

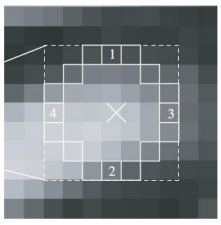
January 2010

This overview analyzes the performance and repeatability of the libmv keypoints detector.

Considered keypoints detector:

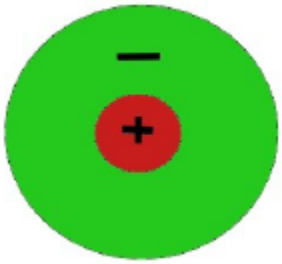
- Fast (detect corners),

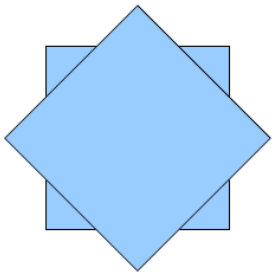




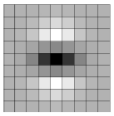
- Star (detect blob),

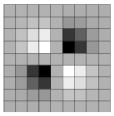
Symmetrical BiLevel approximation with Star Shapes.

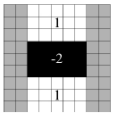


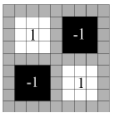


- Surf (libmv implementation, detect blob).



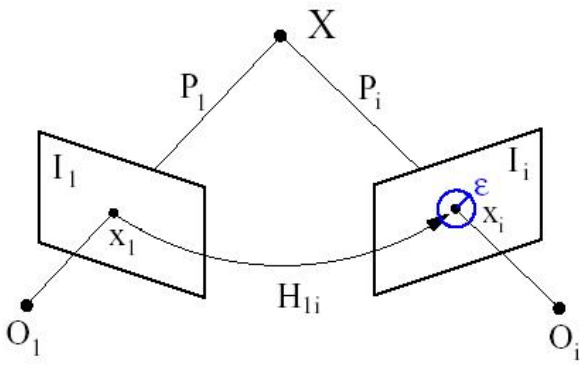






Basic method:

We consider one image as the reference. We detect keypoints over transformed version of the reference image. Like the transformation is known we can compute the residual between the expected position and the founded one.



For resume, consider I1 as the reference.

Compute Keypoints over a series of image Ii. Knowing the transformation H1i, we can compute the point that is the nearest of the expected one H1i* X1 in the image Ii.

A threshold ε is used to ensure position validity.

Output of the algorithm consists in:

- A percentage ∈ [0,100]: Does points computed in Ii are repeatable in I1?

$$\sum_{i=0}^{N} bool(\varepsilon < abs\{(H1i * x1) - xi\})$$

$$Number\ of\ keypoints\ that\ reproject\ in \pm \varepsilon$$

- A mean precision value in position ∈ [0, ε]: Does the repeatable point are near the expected position?

$$\sum_{i=0}^{N} abs\{(H1i * x1) - xi\}$$

$$Mean\ reprojection\ value\ of\ point\ that\ respect\ the\ Threshold\ Error\ \varepsilon$$

Dataset:

A series of images considered in a study “Local Invariant Feature Detectors: A Survey”.

- All Detectors - Survey: T. Tuytelaars and K. Mikolajczyk , Local Invariant Feature Detectors - Survey. In CVG, 3(1):1-110, 2008. PDF
- http://homes.esat.kuleuven.be/~tuytelaa/FT_survey_interestpoints08.pdf

The dataset consist in the following series of images:

http://www.robots.ox.ac.uk/~vgg/data/data-aff.html

- Bikes (varying blur)
- Trees (varying blur)
- Graffiti (varying viewpoint)
- Wall (varying viewpoint)
- Bark (varying zoom+rotation)
- Boat (varying zoom+rotation)
- Leuven (varying light)
- UBC (varying JPEG compression)

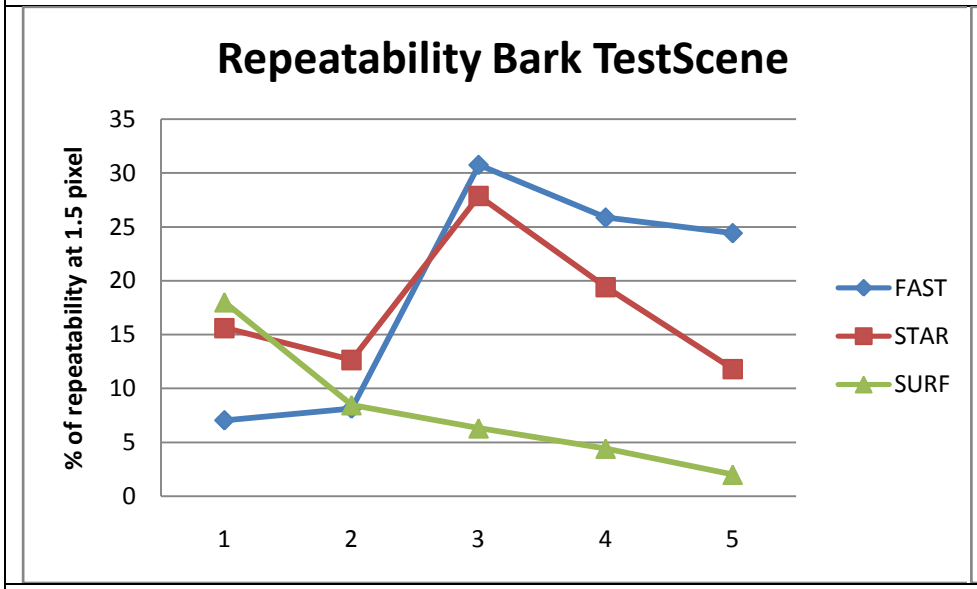
Condition of the test:

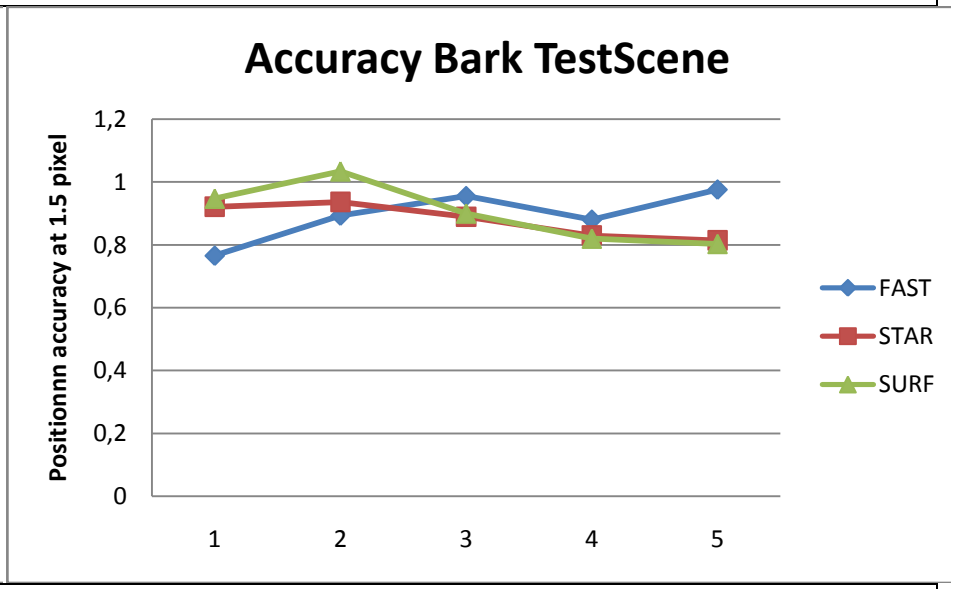
The test was run with the $\varepsilon = 1.5$, and for the three detectors listed at the top of this document (FAST/STAR/SURF). So we will be able to compare the results in term of repeatability and position accuracy.

All the detector will be setup with their default value. (No tuning are allowed).

Results:

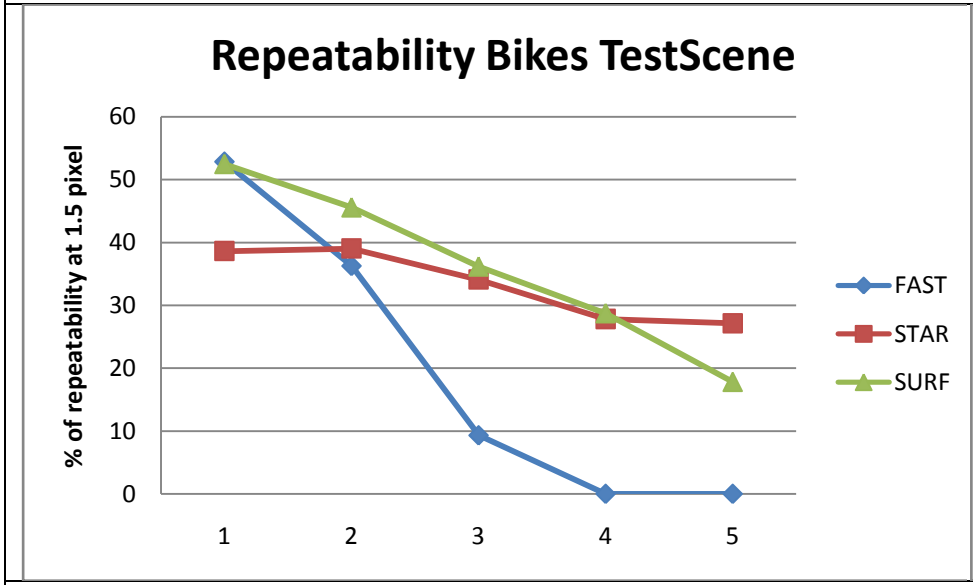
Bark:

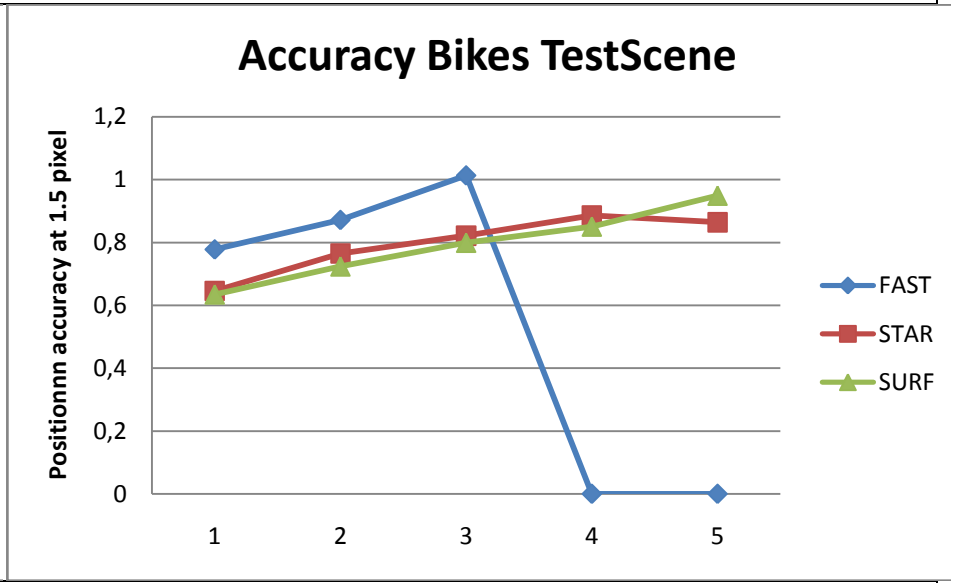




We see that Fast and Star perform the best over the surf libmv detector.
For the accuracy we see that we get similar result.
Note: Our detector only output INT coordinates, so we cannot expect subpixel precision. So a precision near 1 pixel is correct.
Notice the fact that FAST is computed at only one scale but it handle not so bad the point vs point matching, but as long the scale will be setup always to 1, the associated descriptor will be insufficient to give a correct discriminate patch.

Bikes:

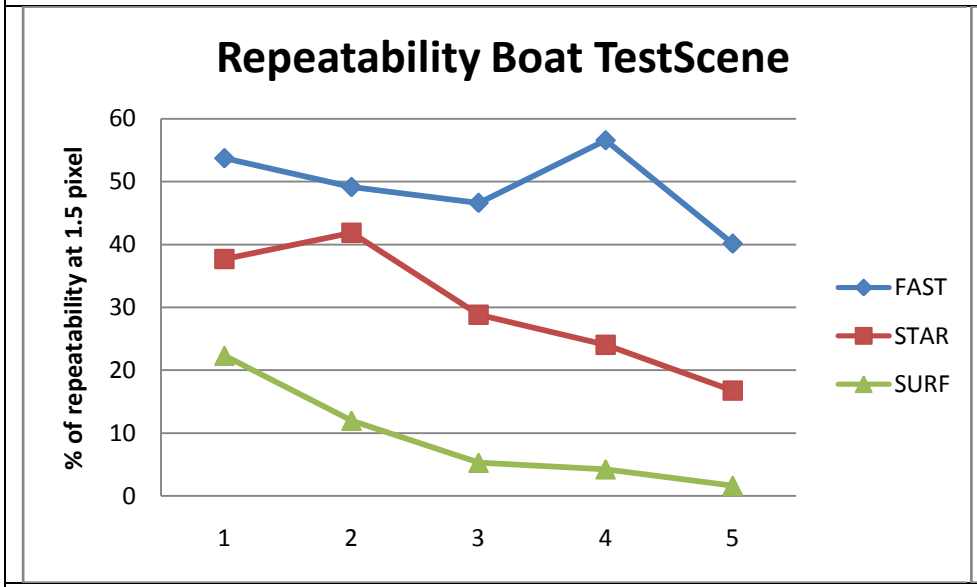


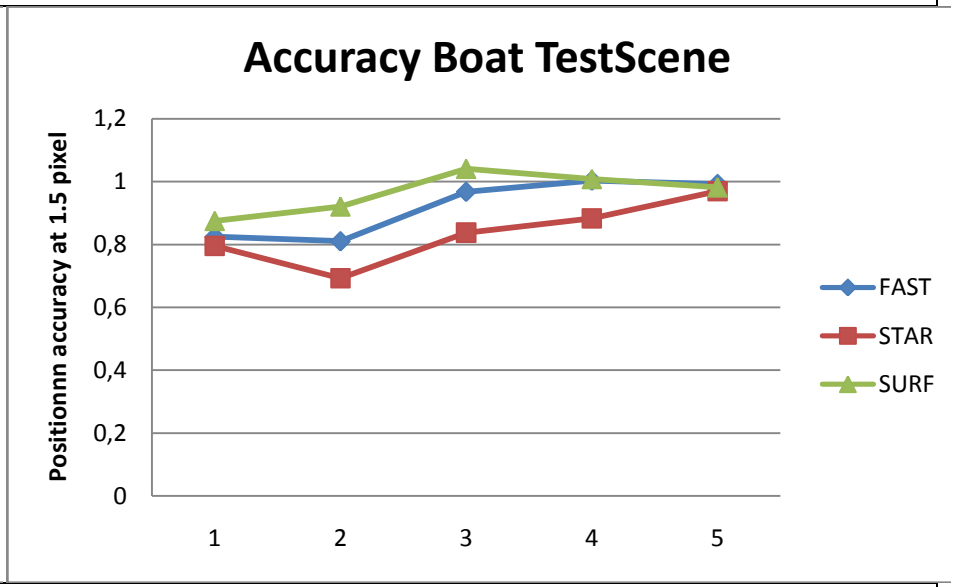


As expected in this scene, Fast perform the worse, because a corner detector is not robust against blur… A blob detector could give better result in handling blurred area…

We notice for the accuracy that we find near the same value between the STAR and SURF… As we expected again… In fact the STAR detector is a SURF simplified copy. As it's a static scene, value are under 1 pixel, so the position are not so bad estimated with the consideration that we do not perform sub pixel estimation.

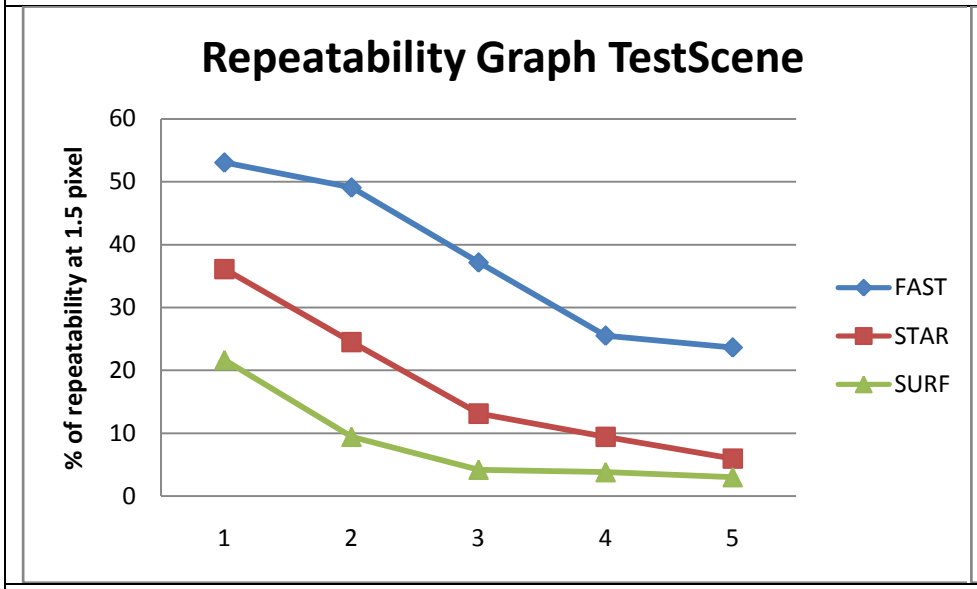
Boat:

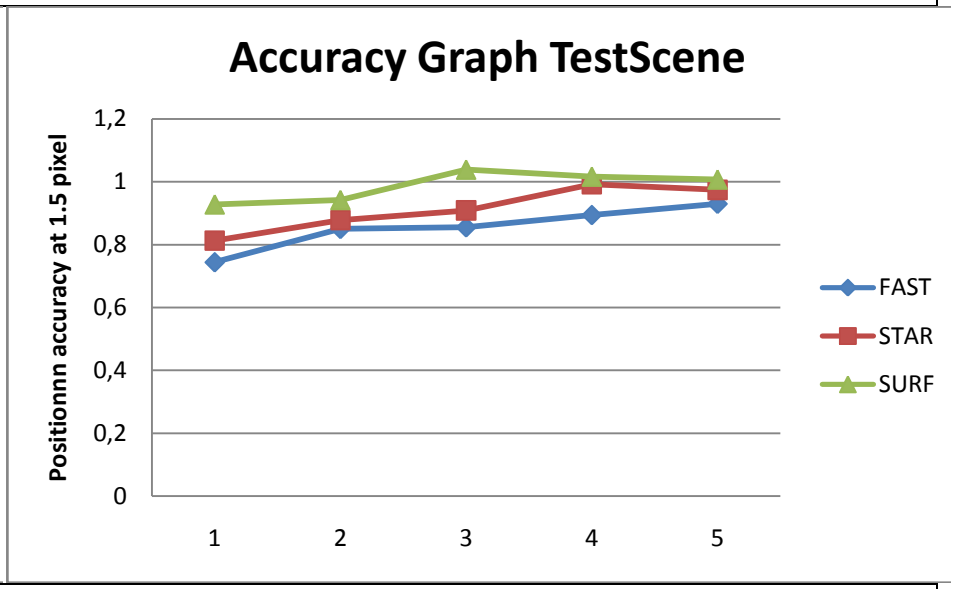




The scene present a lot of structure, it's why FAST perform the best on this test scene.
STAR gives better result than SURF approach…
For the accuracy we notice a similar curve shape for all the detectors. So the rotation angle is not handling perfectly for all detectors.
With surprise STAR get the best accuracy (the minimum value), we remind that STAR is an approximation of a Symmetrical circle BiLevel with star shape so normally it cannot handle very well all rotation angles with the same precision.

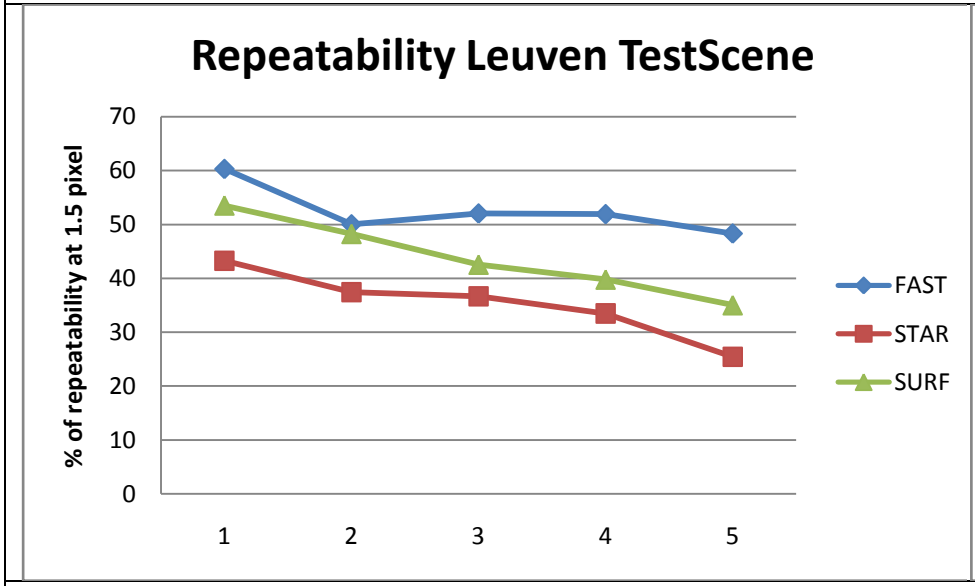
Graph:

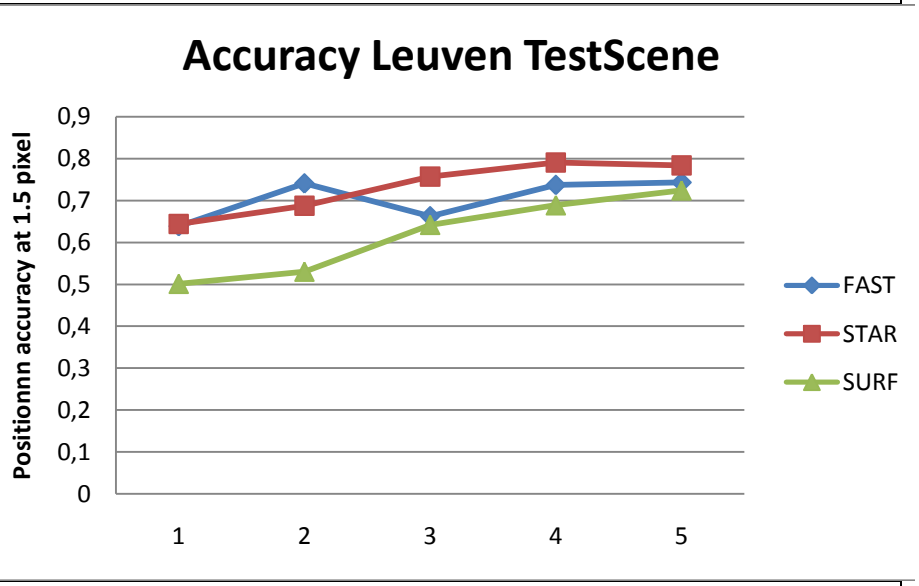




We notice all the graph have the same shape. If viewpoint value augment the number of keypoint decrease and so the matching is not so good.
Affine invariant approach on such type on image must give best result (we think to ASIFT, Affine Hessian …)

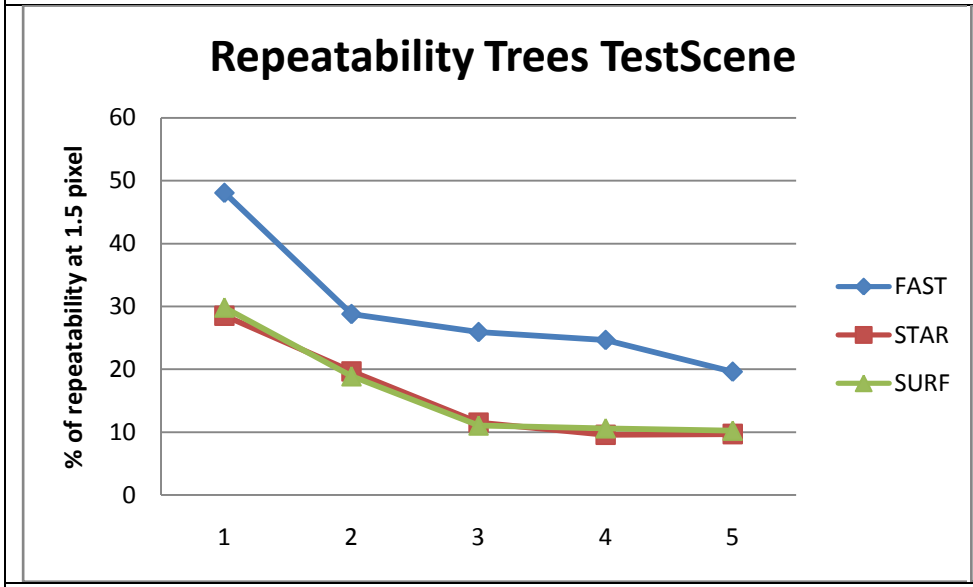
Leuven:

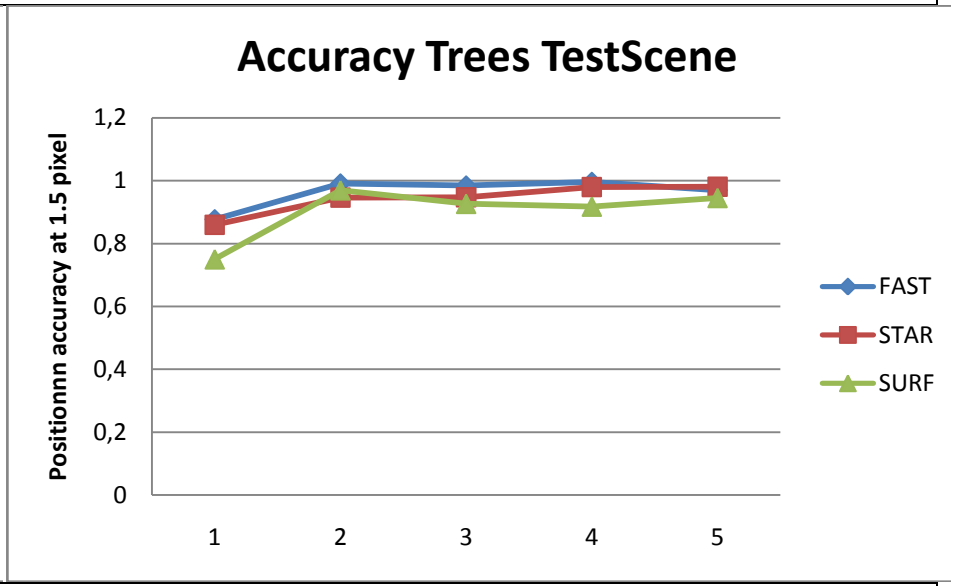




We notice that the detector do not suffer to much of small light variation. STAR is the most affected detector.
As the viewpoint do not change we notice that the precision is now under 0.8 (static scene).

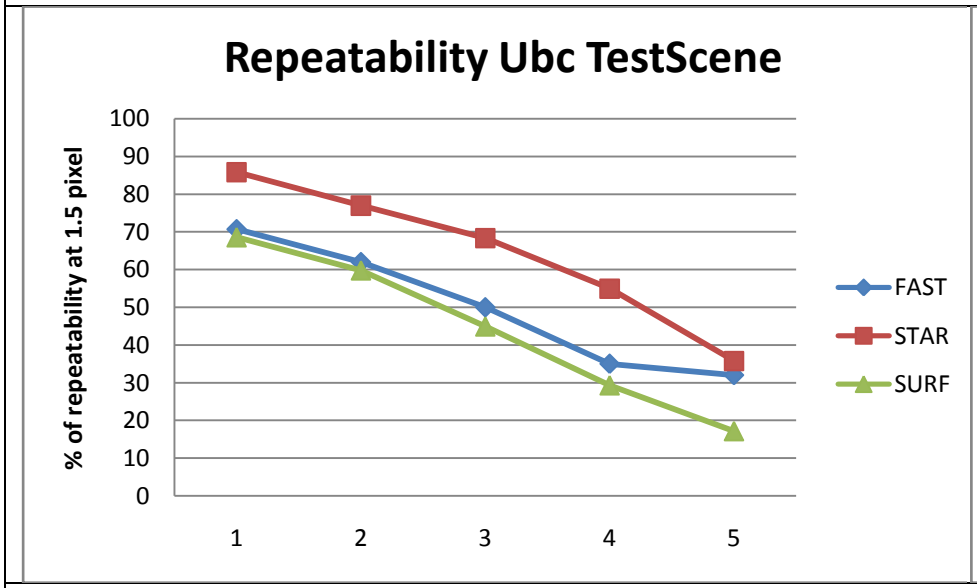
Trees:

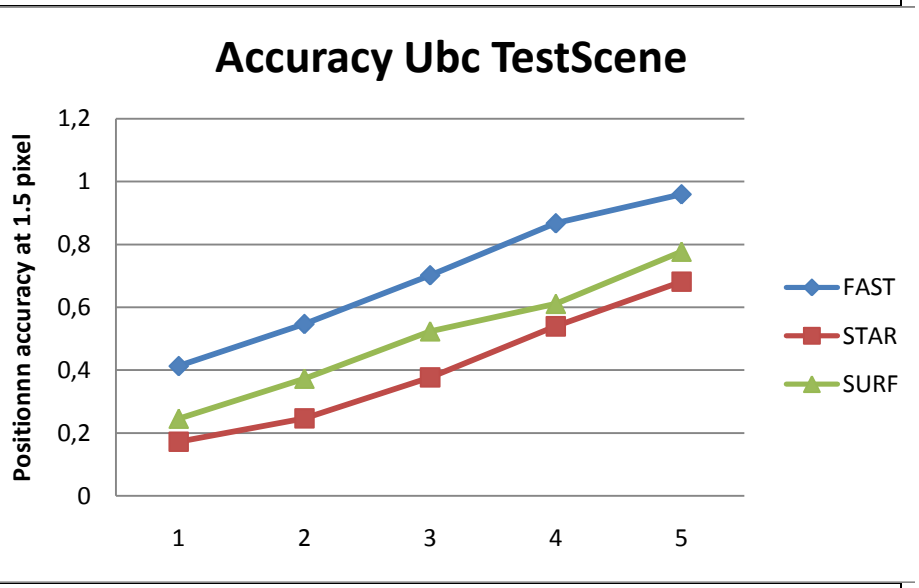




Here we expect the same as the Bikes test scene… But with the fact that the scene is very structured (due to the leaf) the FAST detector performs the best. We see equivalent accuracy precision. With no surprise the repeatability curve for SURF and STAR is the same.

Ubc:

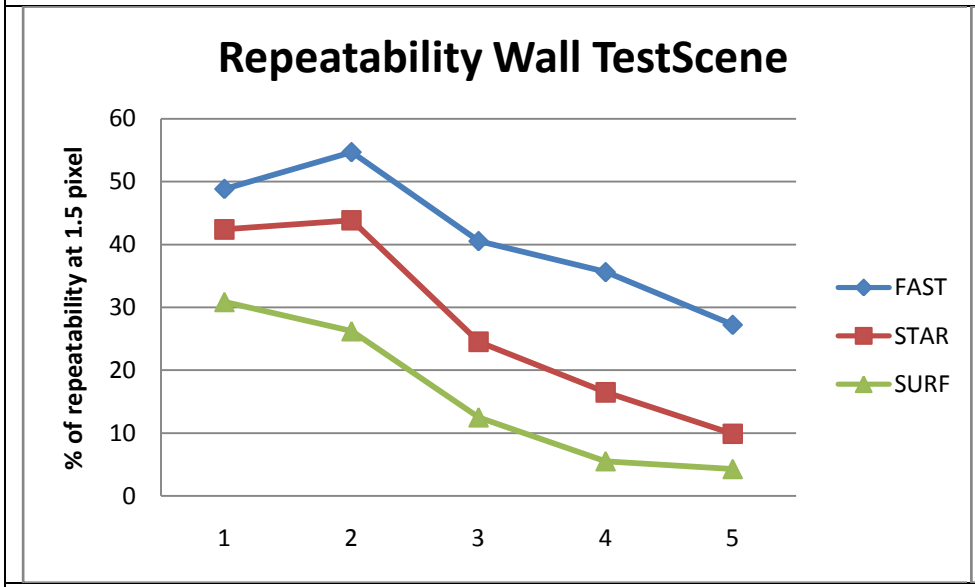


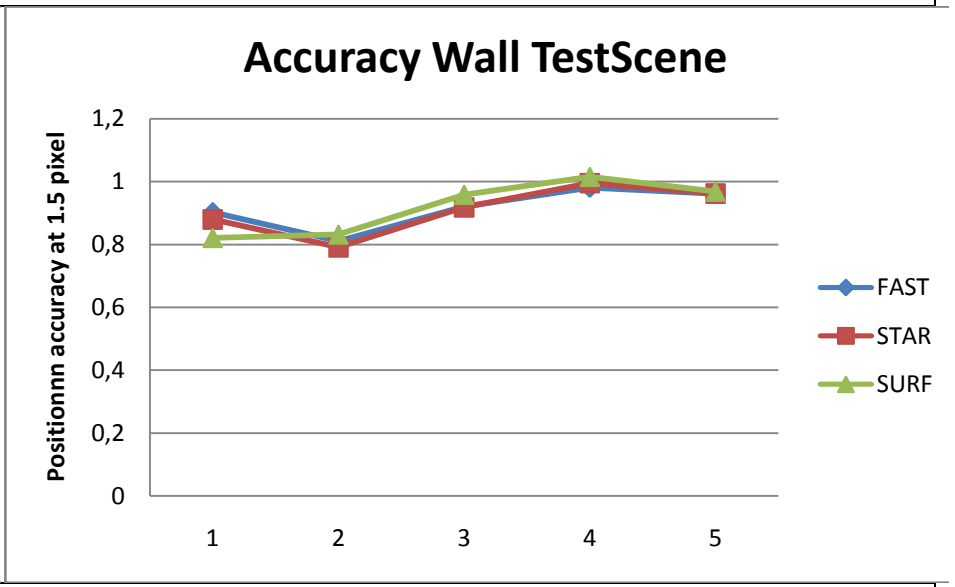


As long as JPEG compression give box artifact it will affect corner detector method. So we notice here that FAST corner detector perform the worst with SURF.

We notice that STAR performs well. It's expected because a blob detector can keep the same center if a blob area is growing… We see very good precision for the STAR detector on such type of scene.

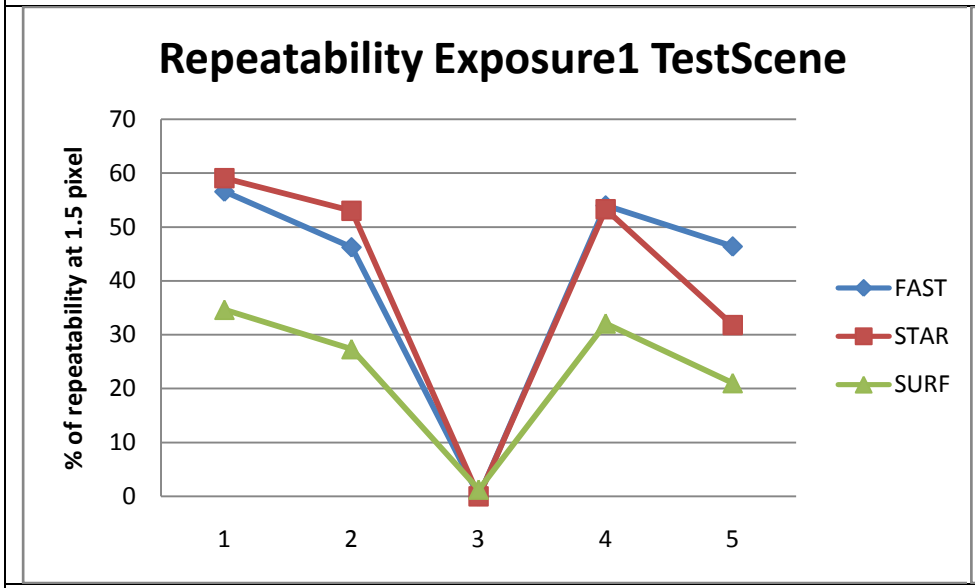
Wall:

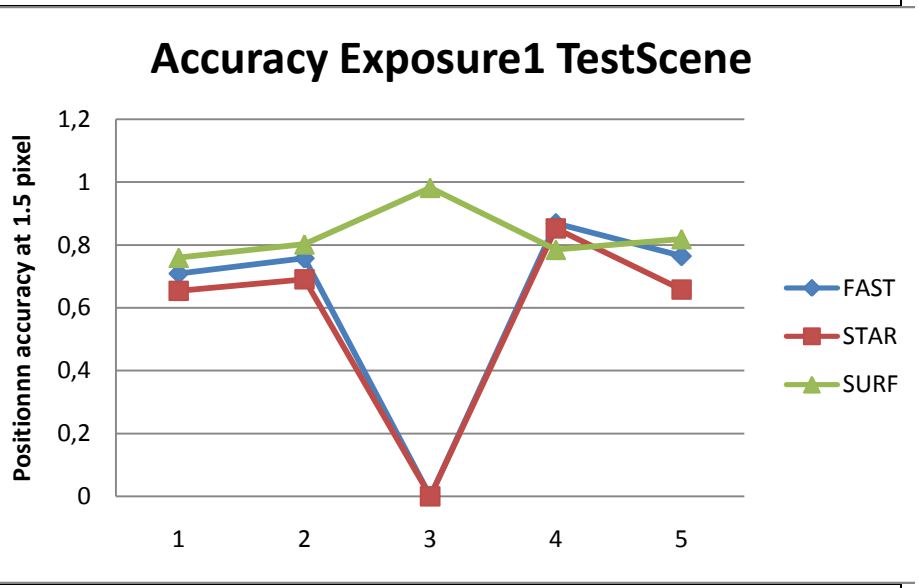




In this scene we expect the same as the Graph test scene.
We see that the curve decrease with the viewpoint become more important. We see that the precision remains the same for all the detectors.

Homemade sequence “Exposure”: 





On this scene: Different exposure value taken to realize a HDR images we get very big difference in exposure.
We see that for the three detectors they cannot handle the exposure 3 (two few light, or not tuned parameters).
Again STAR has the best precision accuracy.

Conclusion :

We have now a main and test set to test our detector over a camera ground truth.

We have seen that our detectors are not so bad and that FAST and STAR seems to be the winner in term of repeatability and position accuracy.

It's clear that the next step is to consider rotation invariance for FAST/STAR/SURF detector to have a complete Keypoint detector system. Subpixel accuracy could be a plus too.

TODO and future work:

Now we must take different task in option.

- Test the accuracy of scale estimation and robustness to rotation. For the moment STAR and SURF are scale invariant, and rotation is not estimated for the three detectors…

- Consider subpixel estimation to achieve best position accuracy in those test set.

- Consider to compute a minimal subset of keypoint on image rather than export the entire set of computed feature. I.E try to keep the 800 Keypoints that get the better score…