logo detection by SIFT matching

25–11-2009

Outline

Basics of the method

Three building blocks:

- **SIFT**: image $\Longrightarrow$ list of keypoints with descriptors
- **match**: two lists of keypoints with descriptors $\Longrightarrow$ list of pairs of closest points
- **adaptive multi-ransac**: list of pairs $\Longrightarrow$ list of affinities

possibly a figure for each function

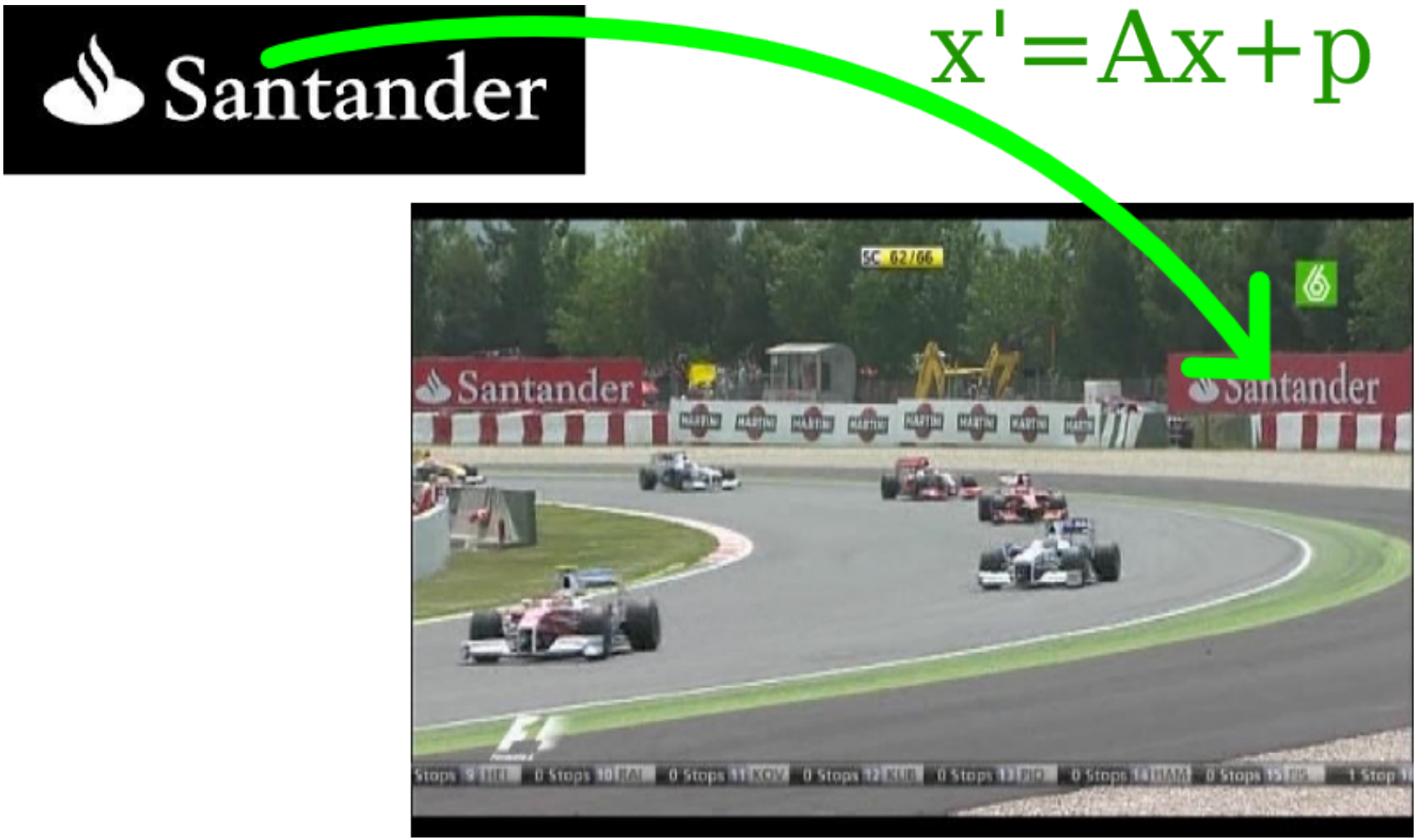
Orbit

logo $\Longrightarrow$ distorted versions of the logo $\Longrightarrow$ SIFT keypoints of all distorted versions

scheme describing the situation

Primary detection

Match frame keypoints against orbit keypoints



Secondary detection (single-frame tracking)

Match best detection against the rest of the frame (and neighboring frames)

Useful when there are several instances of a low-resolution logo, or to track detections in time.

Tertiary detection (temporal tracking)

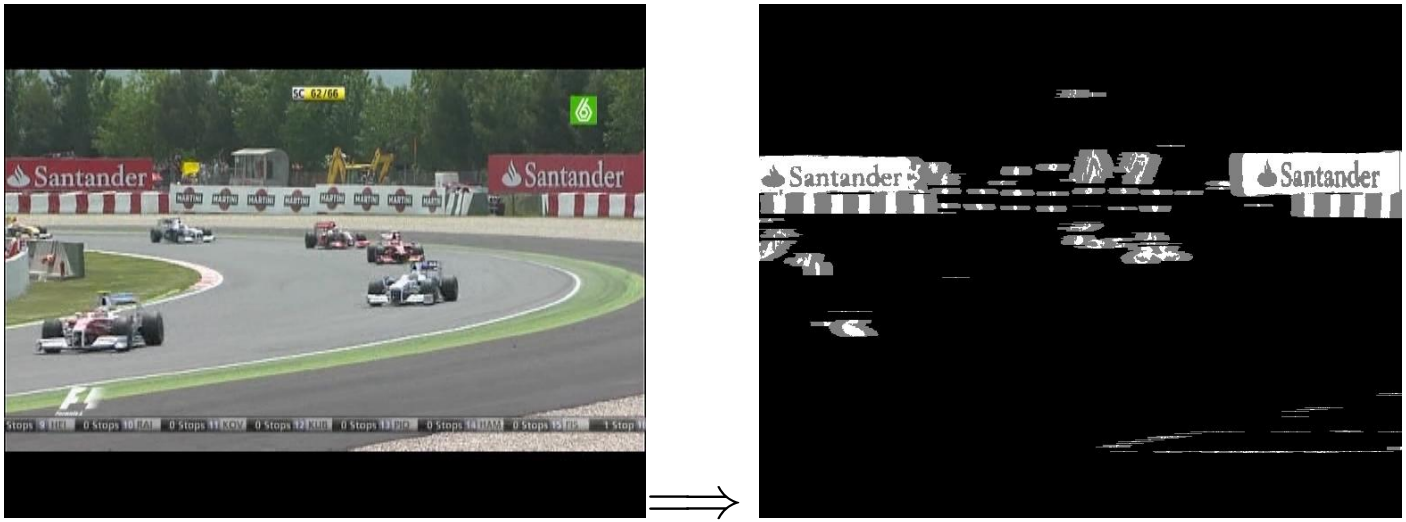
Match a detection to a nearby position onto the next and previous frames

$$\Longrightarrow$$

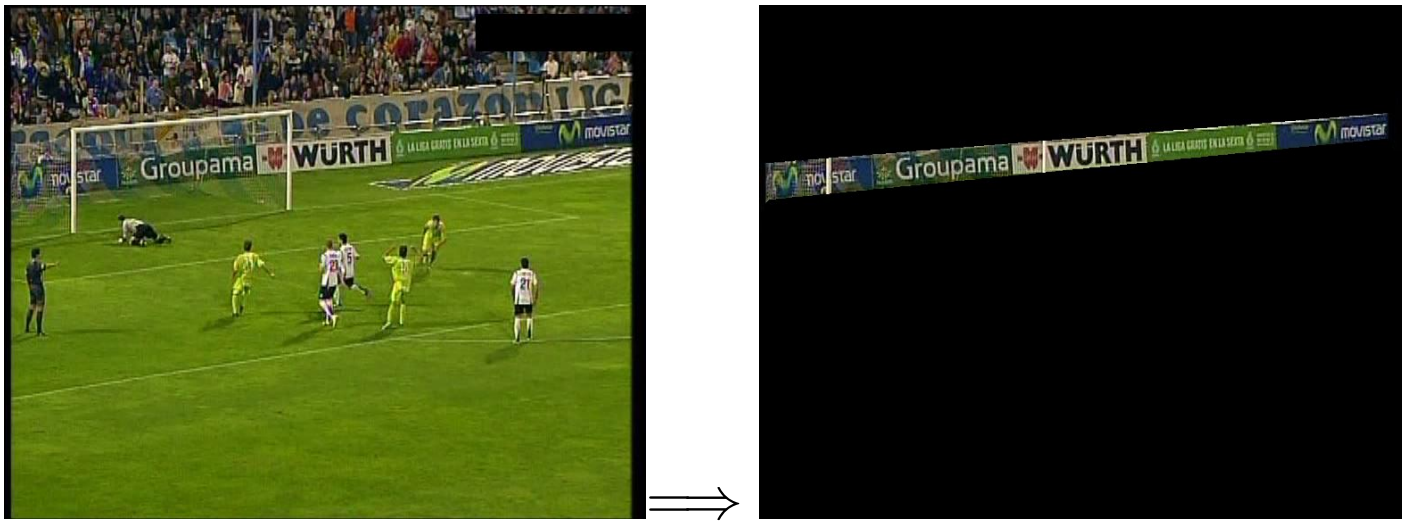
Easier than secondary detection because there are much fewer keypoints!

Trick 1: Use a mask

- Color mask (e.g., Santander logos are red)



- Shape mask (e.g., “U” detection)

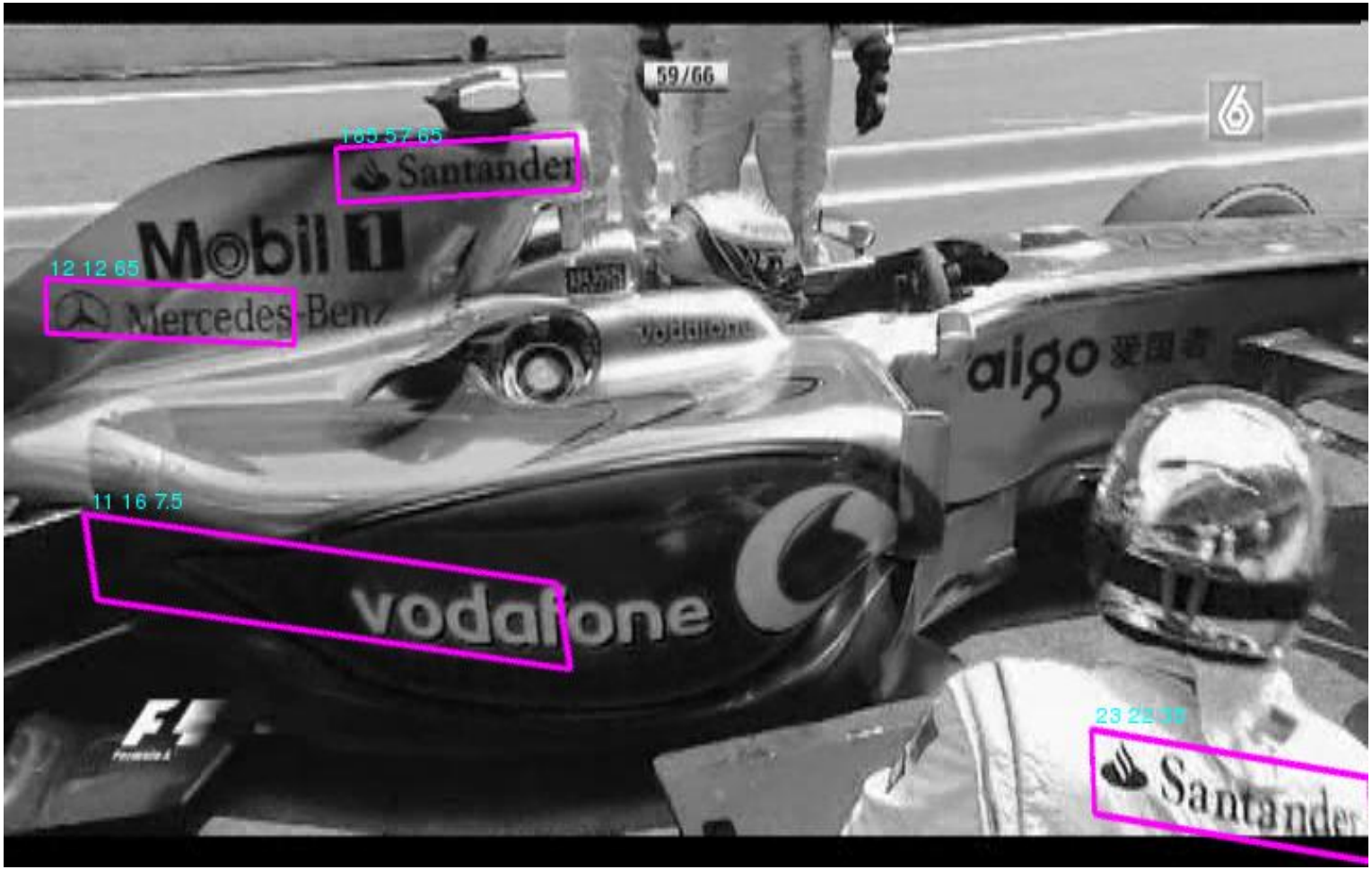


Trick 2: Exclusion principle

Remove overlapping detections (by picking only the best one among each overlapping class)

Trick 3: Multiple logos

- Most false positives are *other* logos (e.g., a “Santander” is detected when there is a “Vodafone”).
- Solution: Look for all possible logos, and apply the exclusion principle to the resulting detections.



List of quality indicators

Quality indicators are numbers associated to each detection. They can be thresholded to adjust the sensitivity.

- Number of inliers
- Maximum error
- Meaningfulness
- Visibility
- Scale (computed from affinity)
- Tilt (computed from affinity)
- Scale (computed from orbit matches)
- Blur (computed from orbit matches)
- Color histograms . . .

Quality indicators computed from RANSAC

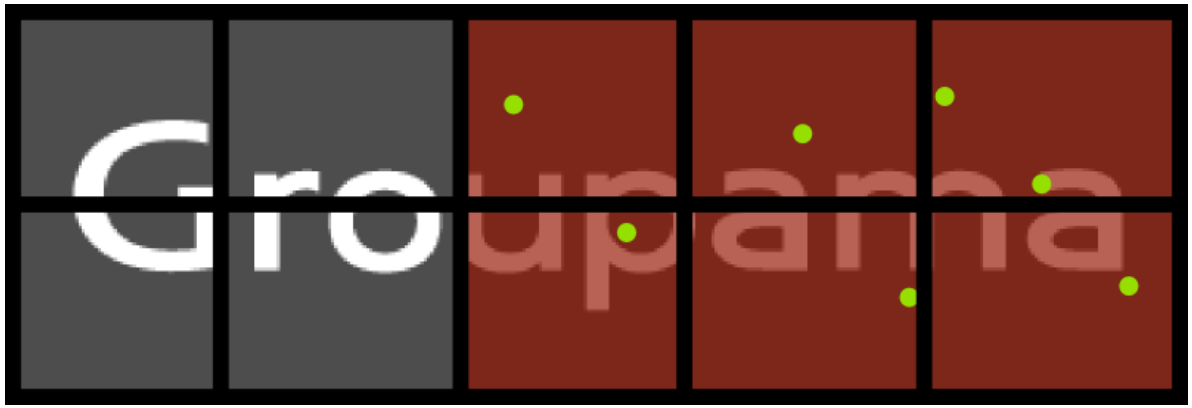
- Number of inliers n
- Maximum error ϵ
- Meaningfulness $= f(n, \epsilon)$

$$f(n, e) = -\log\left(\binom{N}{n}\binom{n}{3}(n-3)\epsilon^{2(n-3)}\right)$$

Interpretation: f is increasing in n and decreasing in ϵ

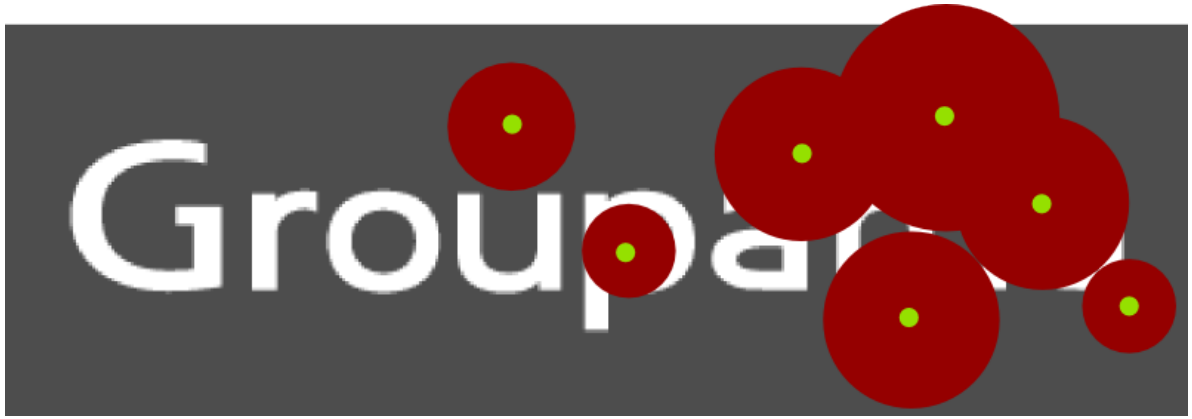
Quality indicators computed from keypoints

- Visibility (coverage of fixed rectangles)



(count number of rectangles with keypoints)

- Visibility (coverage by scaled keypoints)



(compute area of red-painted pixels)

Quality indicators computed from the affinity

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} p \\ q \end{pmatrix}$$

- Scale: $ad - bc$
- Tilt: φ
- Shear: α
- Rotation: θ
- Displacement: $\sqrt{p^2 + q^2}$ (useful only for tracking)

Quality indicators computed from the orbit

- Average scale of matched points
- Average blur of matched points

Quality indicators computed from the image

(not yet implemented)

- $\int_R |\nabla u|^2$ norm on detected rectangle (measures blur)
- Color histogram distance between detected rectangle and original logo
- Correlation between detected rectangle and original logo

Examples of quality indicators

(see annotated video frames)

Conclusion

Future work

- Decide which quality indicators are more informative.
- Understand the distribution of inliers within the detected rectangles. *(They are often distributed on a few clusters corresponding to one or two letters of the logo)*.