

Computer Vision for Augmented Reality

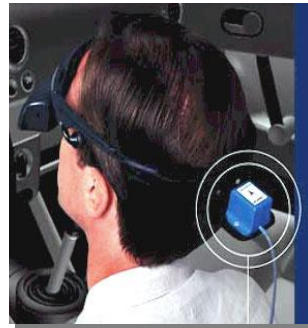
Vincent Lepetit

CVLab, EPFL, Switzerland

Which Tracking Technology for Augmented Reality ?



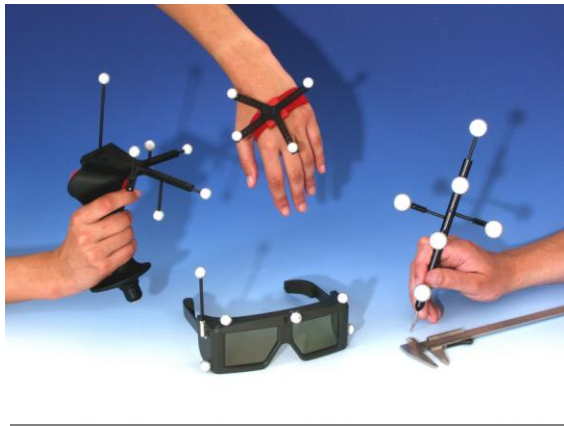
Flock of Birds (Ascension)



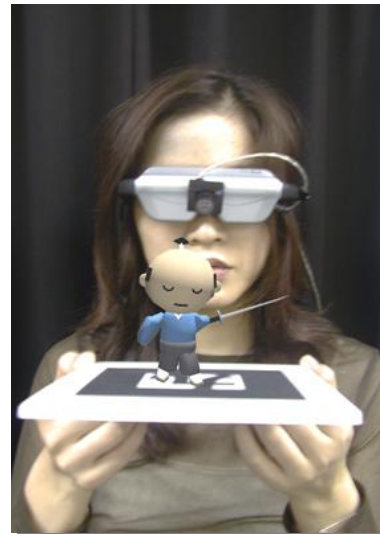
3D Bird (Ascension)



laserBird 2 (Ascension)



(A.R.T. GmbH)



ARToolkit



HiBall-3100 (3rd Tech)

Computer Vision with a single camera is probably the most practical technology for large public AR/VR applications.

"Ultimately, tracking in unprepared environments may rely heavily on tracking visible natural feature points [...]."

R. Azuma, Y. Baillet, R. Berhinger, S. Feiner, S. Julier, and B. MacIntyre. IEEE Computer Graphics and Applications. 2001.

But...

...no real AR application uses Computer Vision yet.

Vision-Based 3D Tracking

QuickTime™ and a
decompressor
are needed to see this picture.

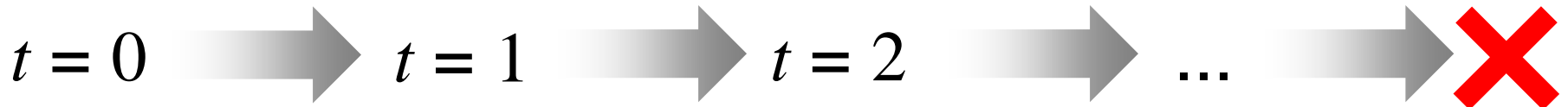
QuickTime™ and a
decompressor
are needed to see this picture.

QuickTime™ and a
decompressor
are needed to see this picture.


Recursive Tracking

QuickTime™ and a
decompressor
are needed to see this picture.


QuickTime™ and a
decompressor
are needed to see this picture.




3D Object Detection



$t = 0$




$t = 1$



$t = 2$



...



t

DEMO

QuickTime™ and a
decompressor
are needed to see this picture.

without the ability to detect the target object from scratch, a tracking system is not useful

if the target object has enough feature points, we now have the techniques to detect it in real-time

possible applications of
feature point-based detection

QuickTime™ and a
H.264 decompressor
are needed to see this picture.

C. Scherrer, J. Pilet, P. Fua, and V. Lepetit, The Haunted Book. *ISMAR'08*.



[Total Immersion]

3D objects

QuickTime™ and a
decompressor
are needed to see this picture.

M. Ozuysal, V. Lepetit, F. Fleuret, and P. Fua, Feature Harvesting for Tracking-by-Detection.
ECCV'06

3D environments

QuickTime™ and a
decompressor
are needed to see this picture.

[ARmusement]




QuickTime™ and a
decompressor
are needed to see this picture.


QuickTime™ and a
PNG decompressor
are needed to see this picture.

[Georg Klein & David Murray; CVPR'07 / ISMAR'07]


detection is not needed in every frame



$t = 0$




$t = 1$



$t = 2$

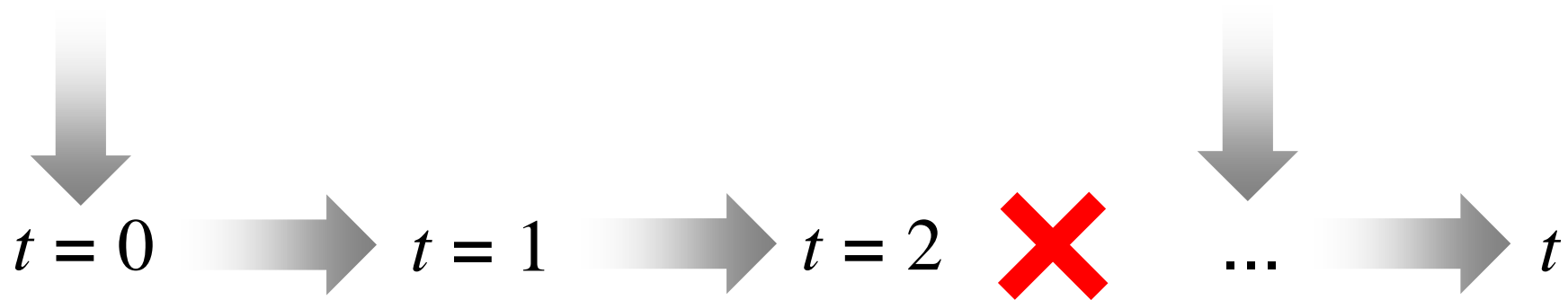


...

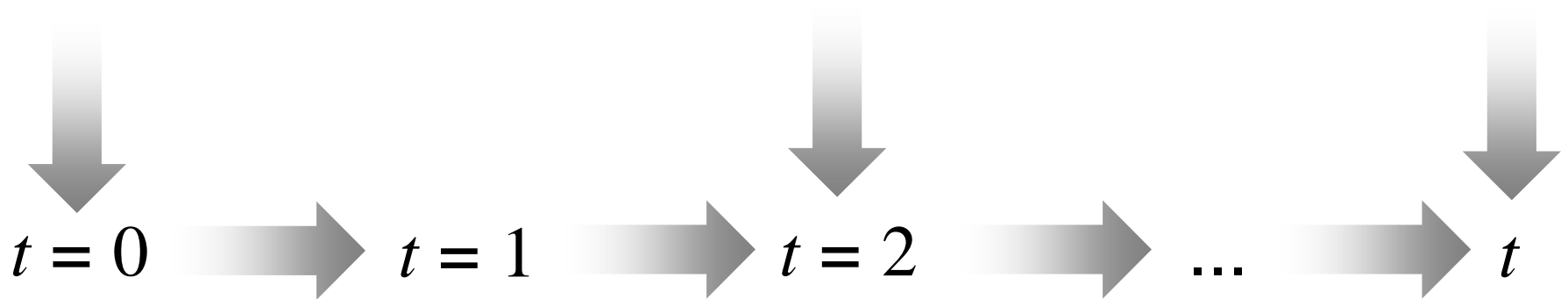


t

detection is not needed in every frame



detection is not needed in every frame



QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

Y. Park, V. Lepetit, and W. Woo, Multiple 3D Object Tracking for Augmented Reality.
ISMAR'08.

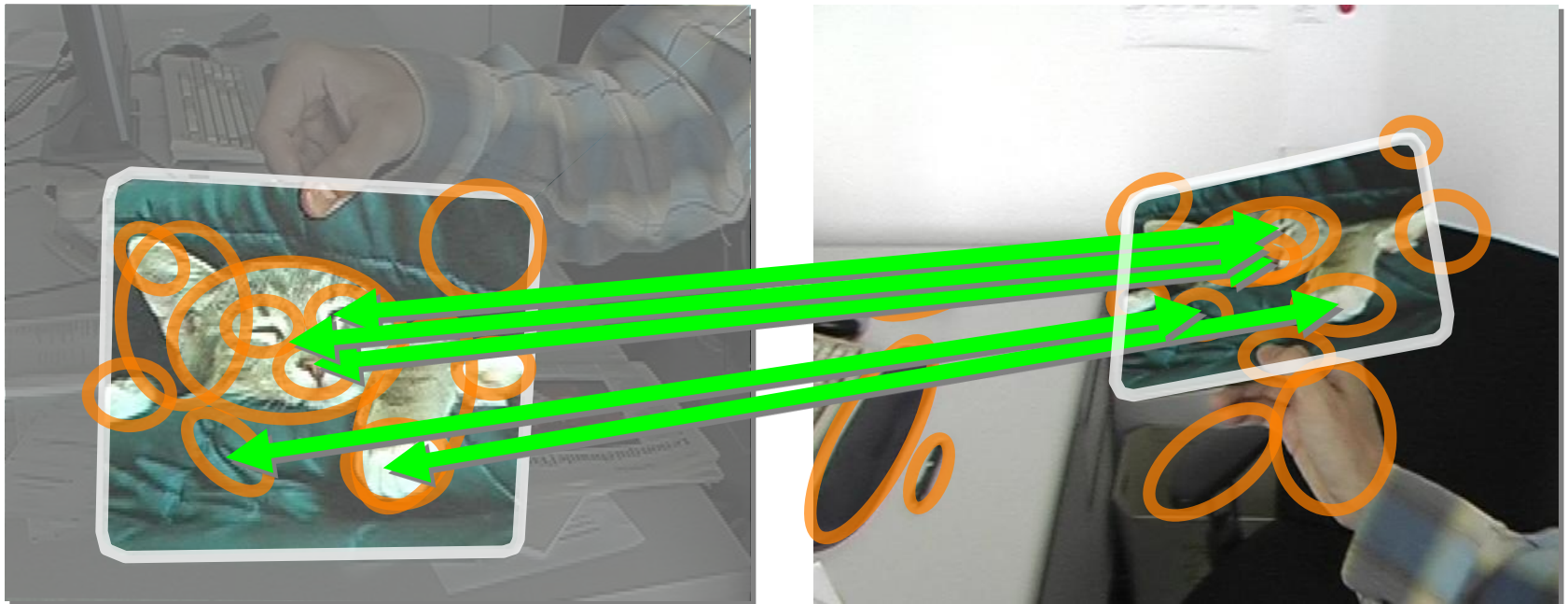
how it works

3D Object Detection

Keypoint detection (Harris, extrema of Laplacian, affine regions,...);

Keypoint recognition (descriptor matching or classification);

Robust pose estimation (RANSAC+P3P, ...).



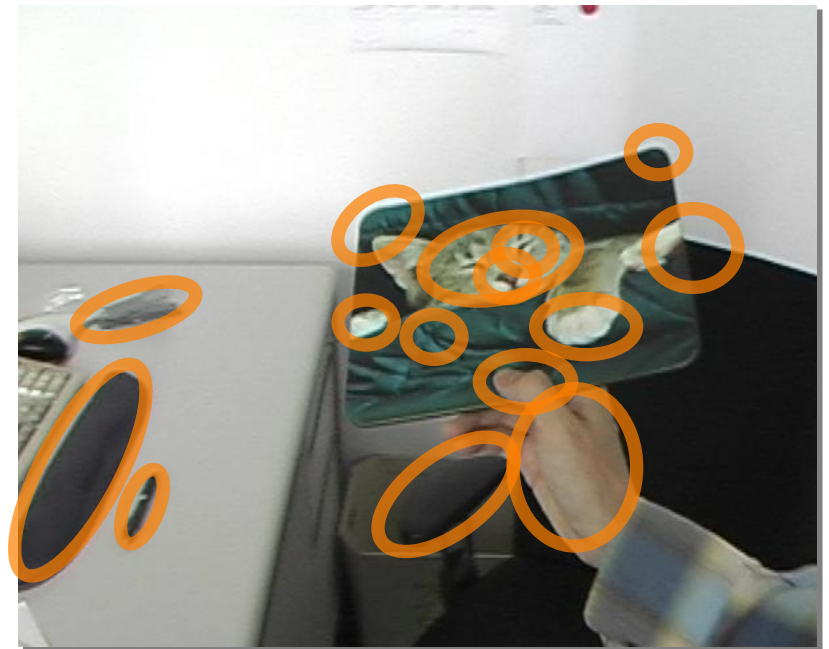
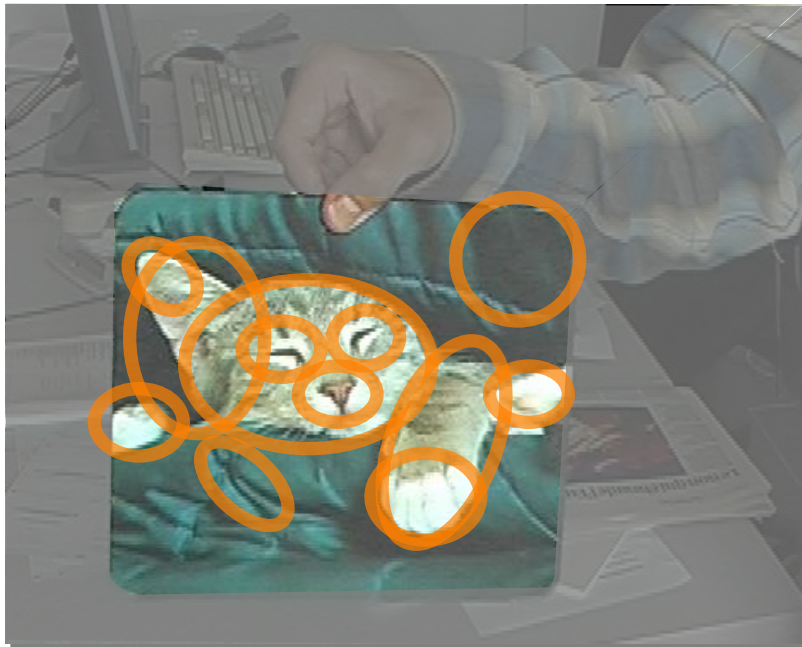
Registered image(s) of
the object to detect

Input image

Standard Approach to Keypoint-Based Object Detection

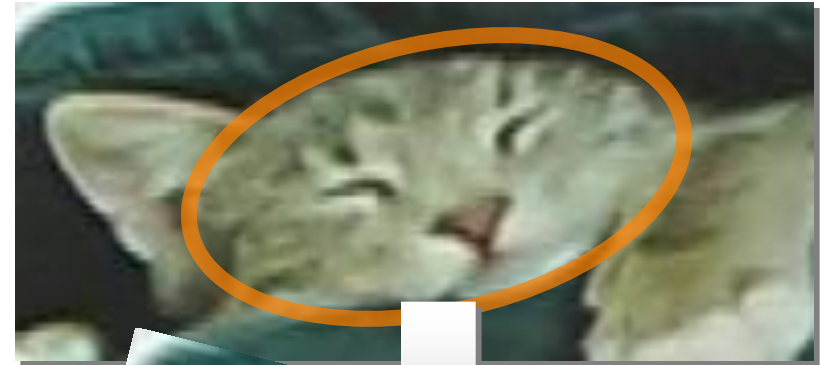
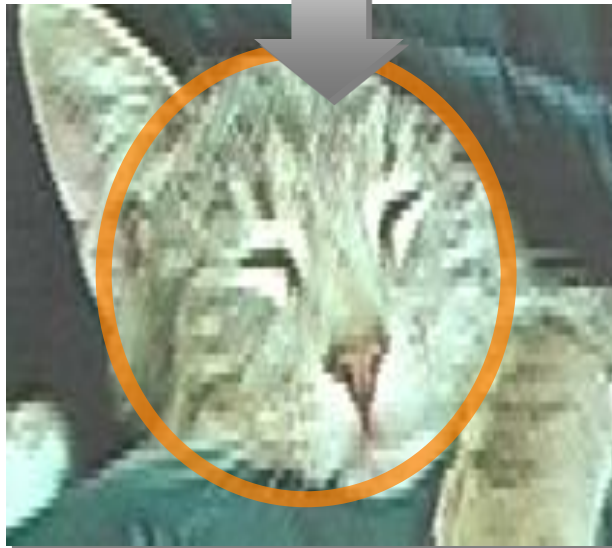
Standard Approach

Step 1: Detection invariant to scale and rotation, or perspective transformation



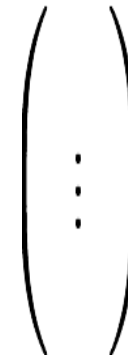
Standard Approach

Step 2: Patch rectification



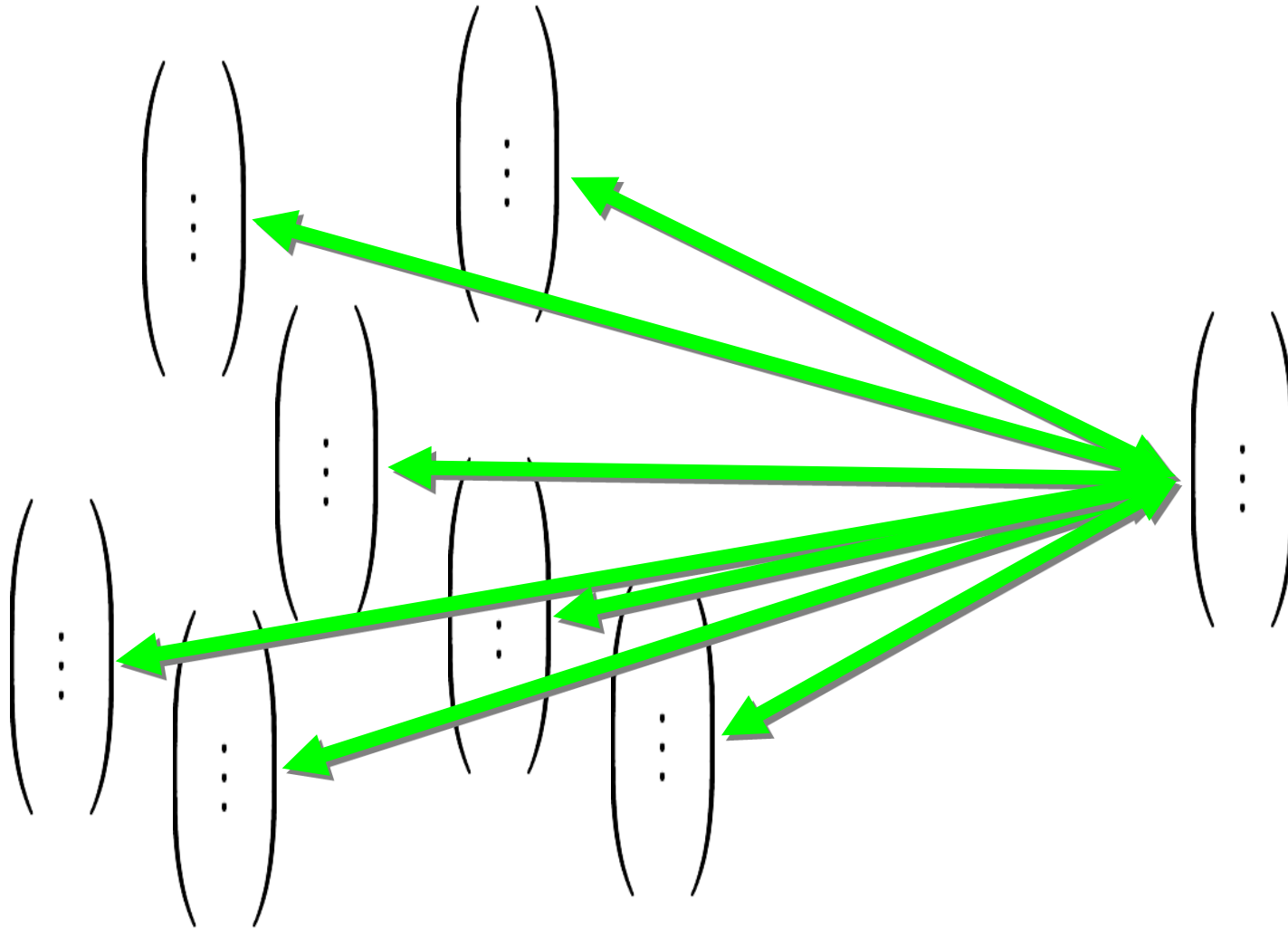
Standard Approach

Step 3: Build description vector



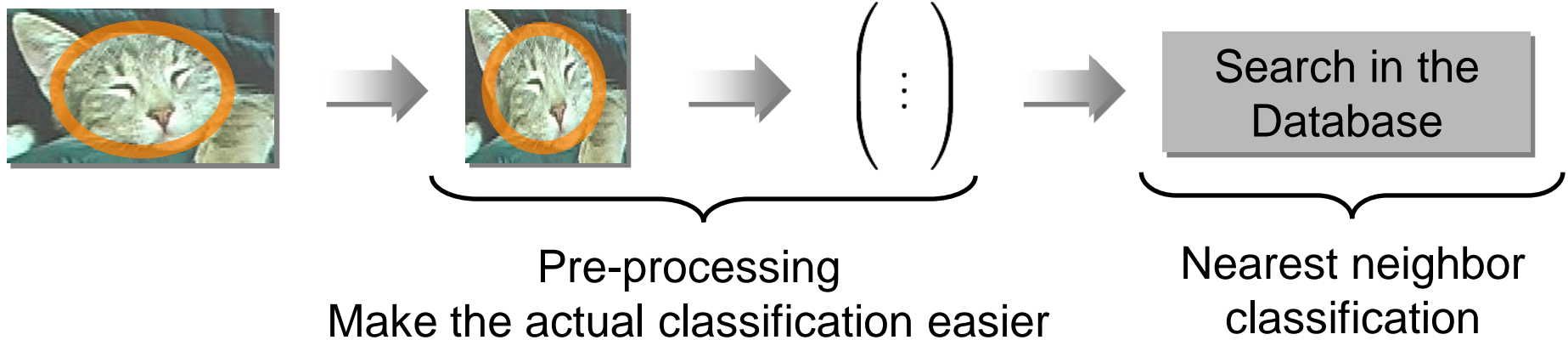
Standard Approach

Step 4: Match description vectors

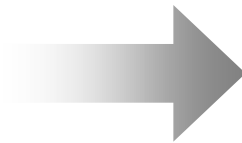
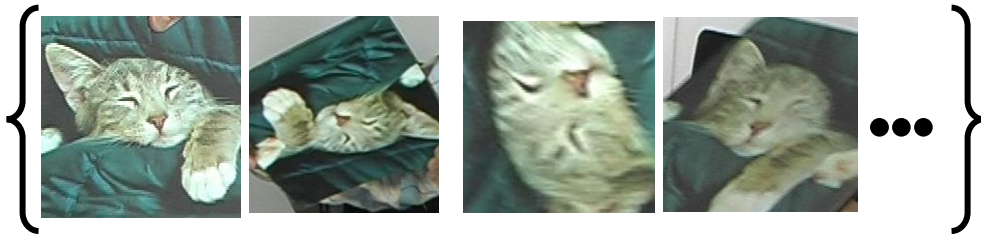


Keypoint Recognition

The standard approach is a particular case of classification:



Training phase

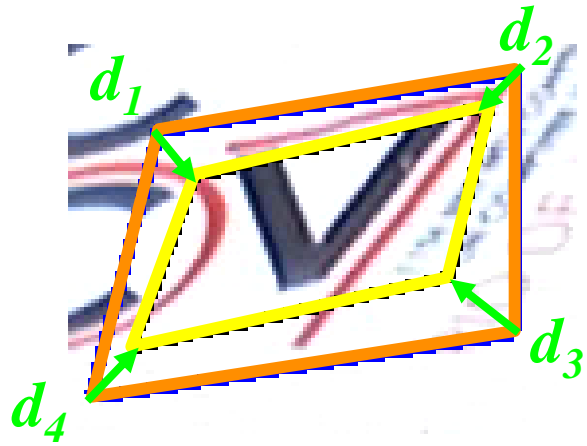


Classifier

Used at run-time to recognize the keypoints

M. Ozuysal, P. Fua, and V. Lepetit, Fast Keypoint Recognition in Ten Lines of Code. CVPR'07.

working with less textured objects



$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix} = A(I(H) - I^*)$$

QuickTime™ and a
decompressor
are needed to see this picture.

S. Hinterstoisser, S. Benhimane, N. Navab, P. Fua, V. Lepetit, Online Learning of Patch Perspective Rectification for Efficient Object Detection. CVPR'08.

what about non-textured objects ?

QuickTime™ and a
decompressor
are needed to see this picture.



[Ahmadi et al. Curac]

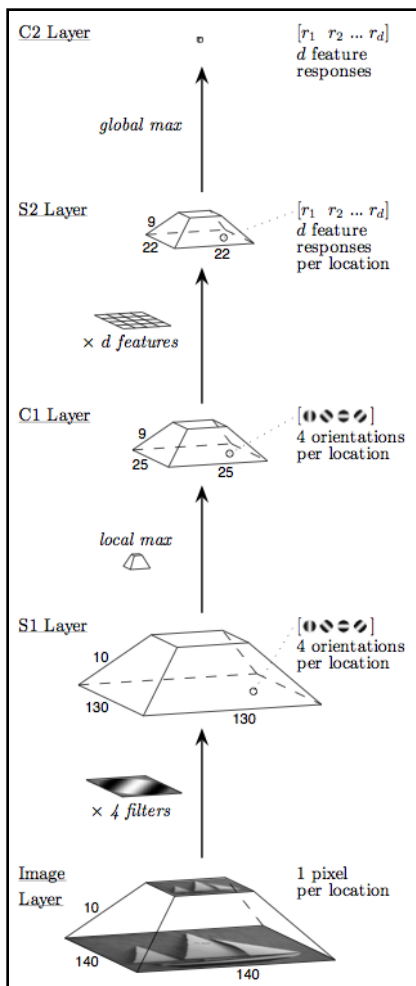


feature points are probably not adapted

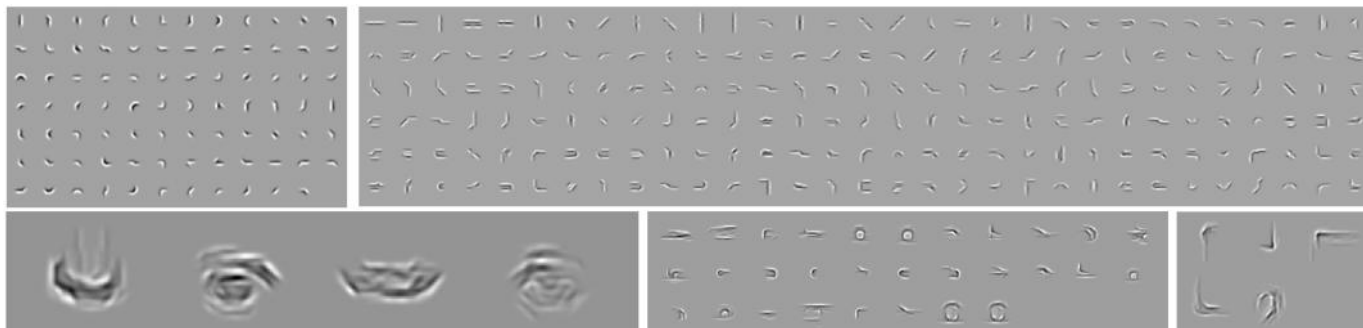
we need something else

look at object recognition community

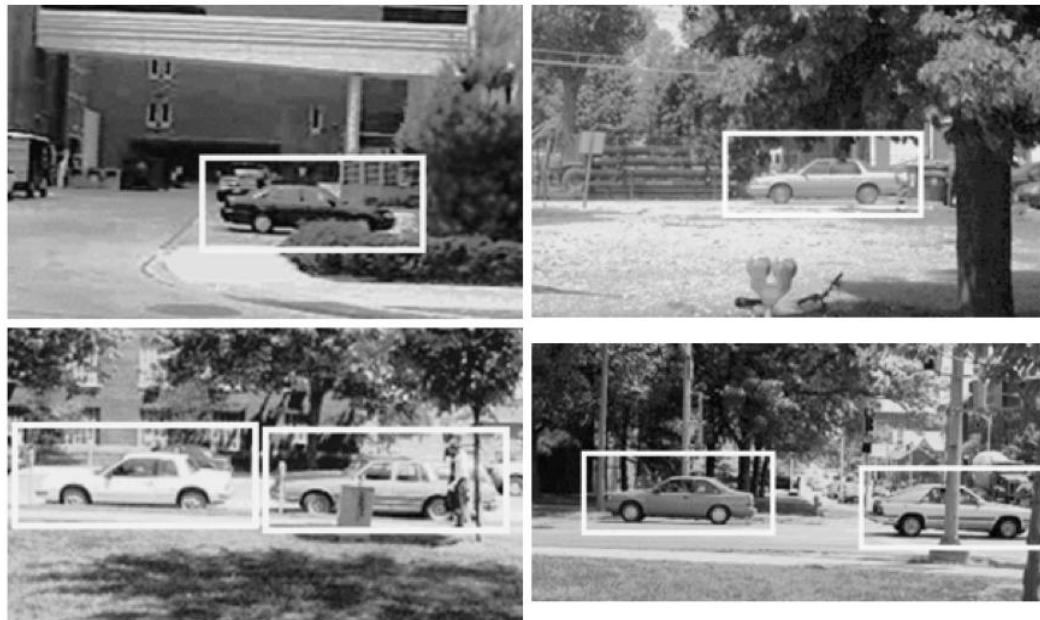
State-of-the-Art in Object Recognition



[Mutch&Lowe,CVPR'06]



[Fidler&Leonardis, CVPR'07]



[Mutch&Lowe,CVPR'06]

Thank you