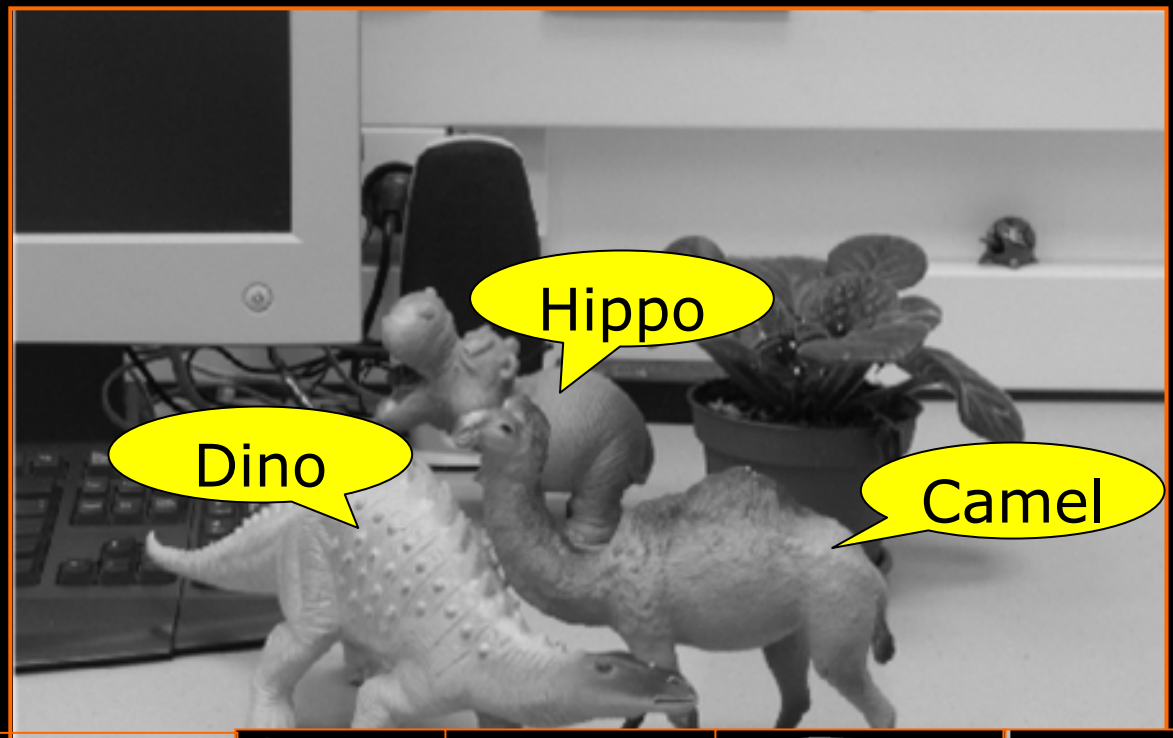
Indexing with Unknown Illumination and Pose

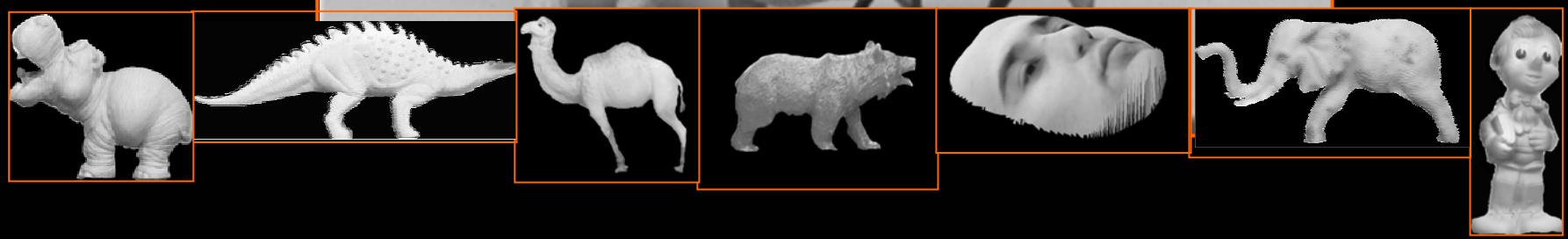
Ira Kemelmacher Ronen Basri
Weizmann Institute of Science

The task – Shape Indexing

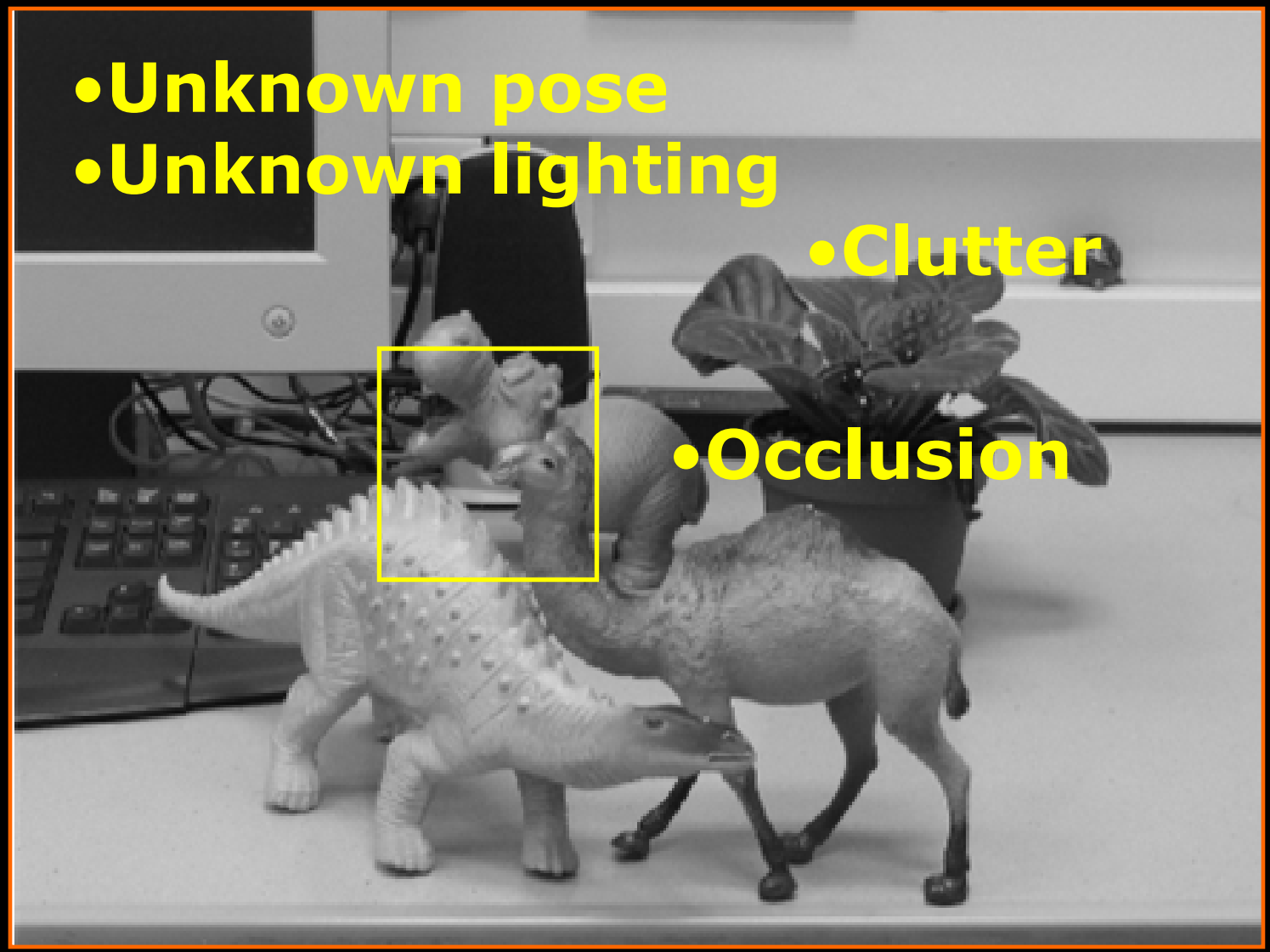
1. Recognize

2. Recover pose+lighting





Why is it hard?



Assumptions

- Weak perspective projection
- 3D rigid transformation
- Lambertian model

Previous...

- Identification using alignment
  Fischler and Bolles, Huttenlocher and Ullman
- "3D to 2D invariants do not exist"
  Burns *etal*., Moses *etal.*, Clemens *etal.*
- Indexing faster than alignment
  Jacobs, Wolfson *etal.*

Previous Indexing Methods

- Ignored intensity information
- Need many point or line features
- Restricted to polyhedral objects

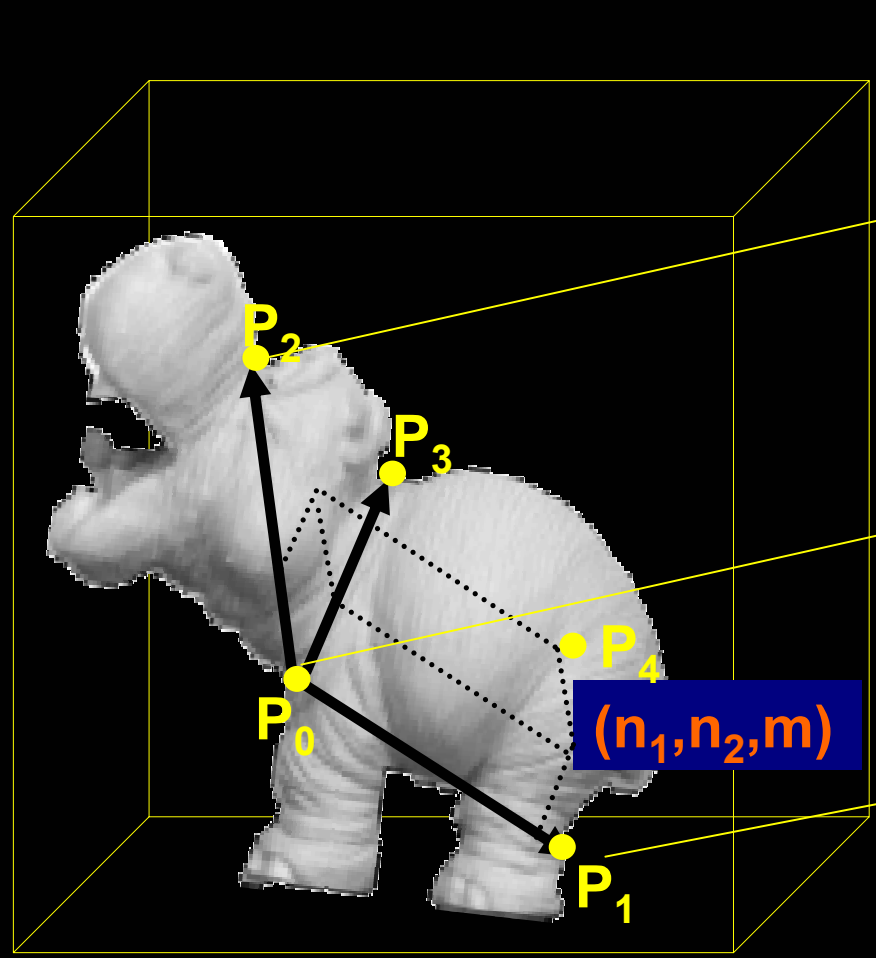
Our algorithm…

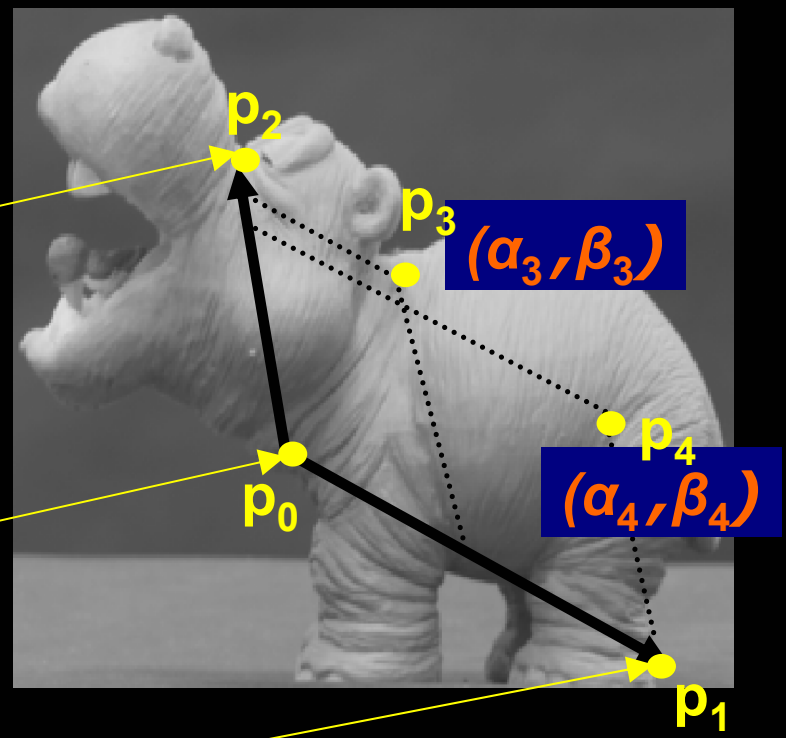
- Handles both *pose* and *lighting*
- Uses *intensities* to filter out incorrect matches
- Still relies on point features but only *very few* are needed
- *General* objects

Indexing with pose - Affine model

(Jacobs '96)

$p_i = \boldsymbol{A}P_i + \boldsymbol{t}$ ➔ 8 DOF -> 5 points

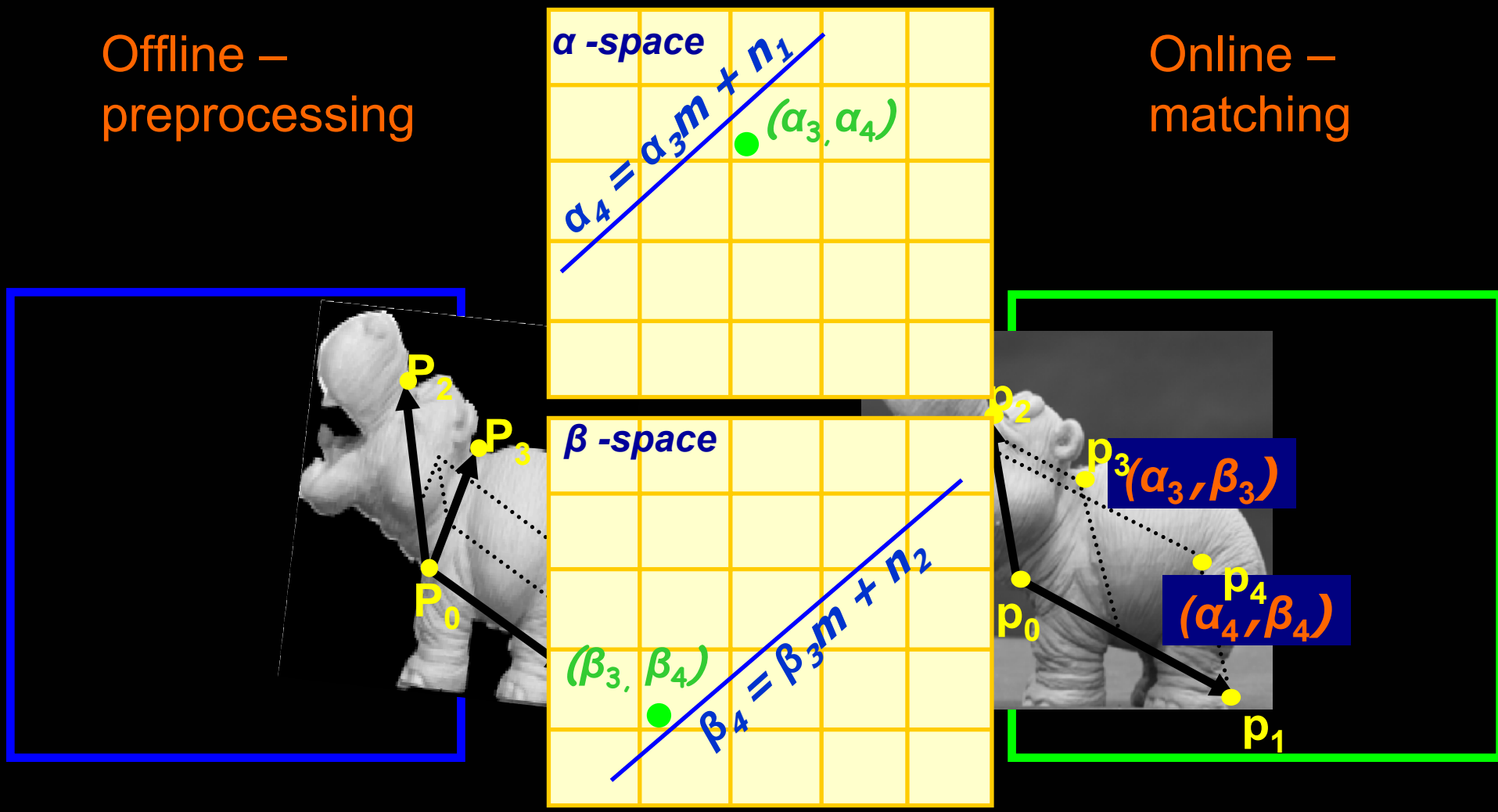




Representation in two 2D tables

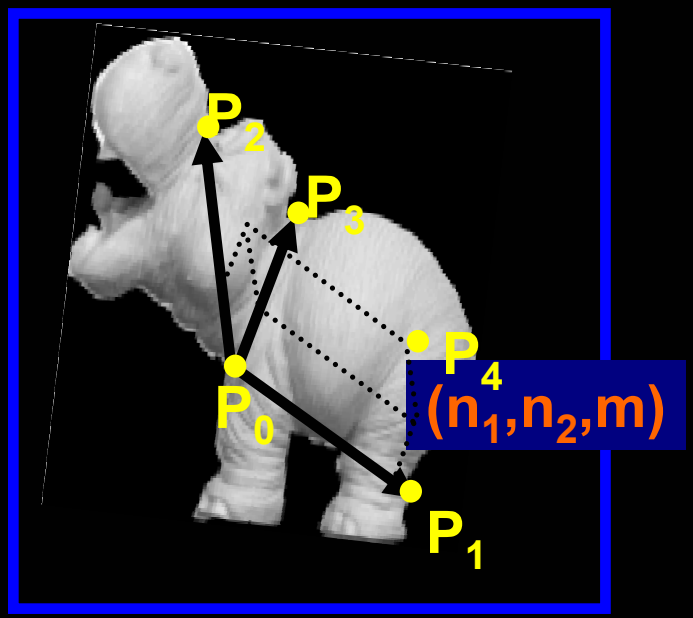
Offline – preprocessing

Online – matching

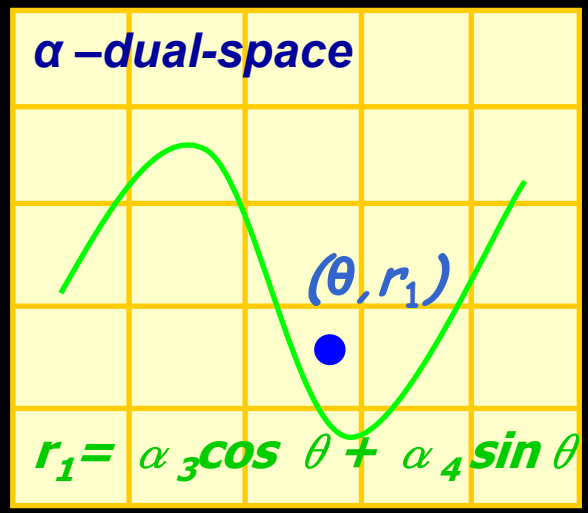


Modifications – still two 2D spaces

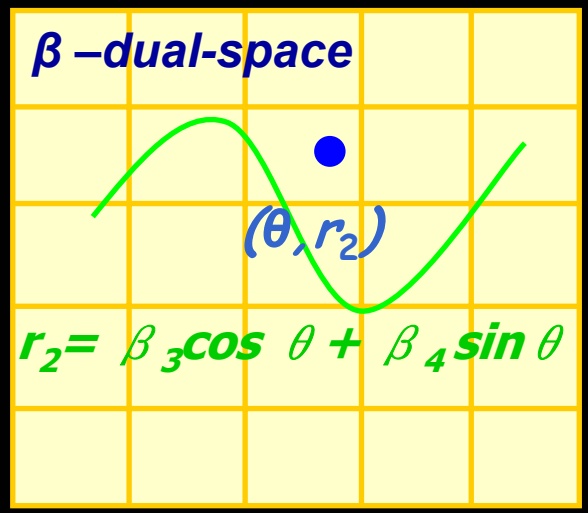
Offline – preprocessing



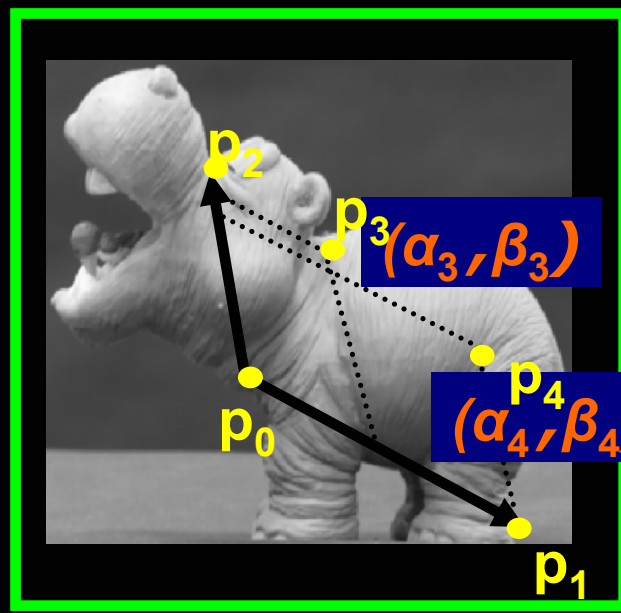
α –dual-space



β –dual-space

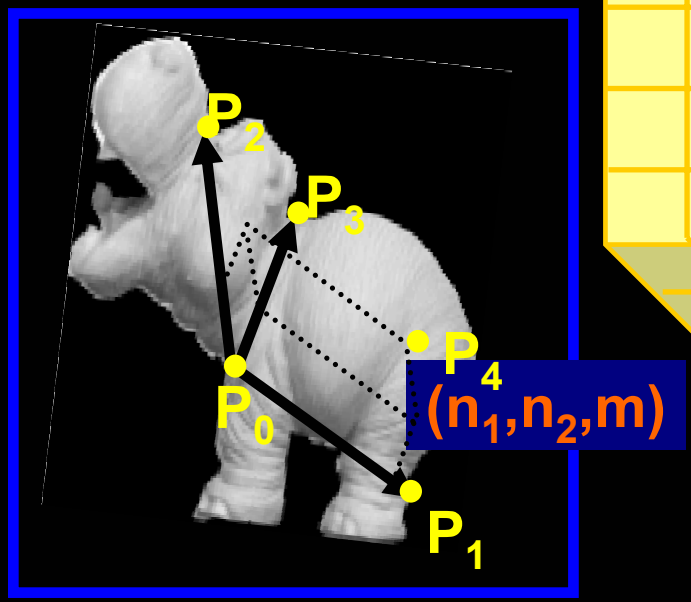


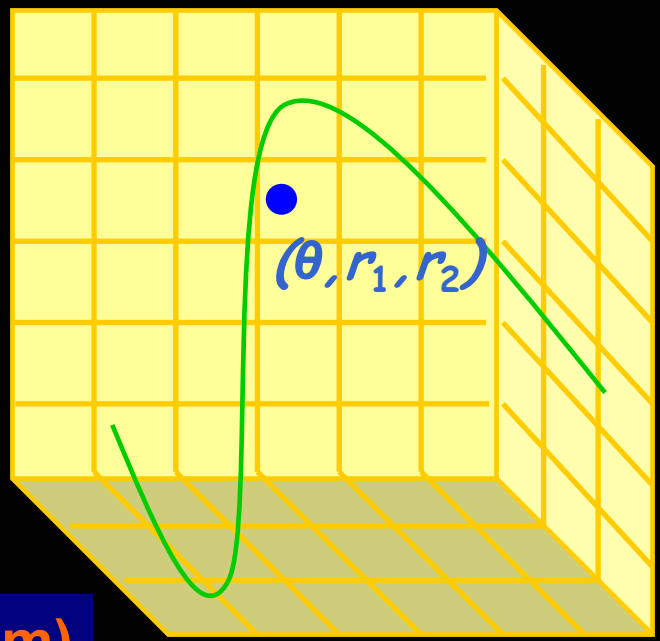
Online – matching



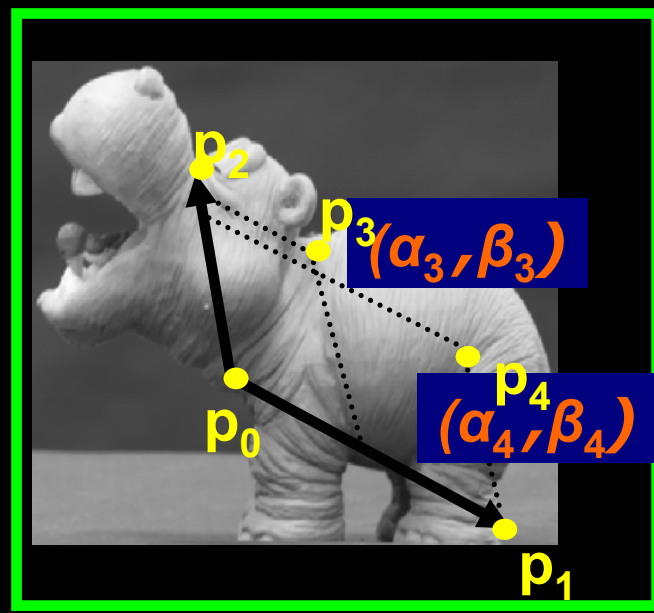
One 3D space

Offline – preprocessing

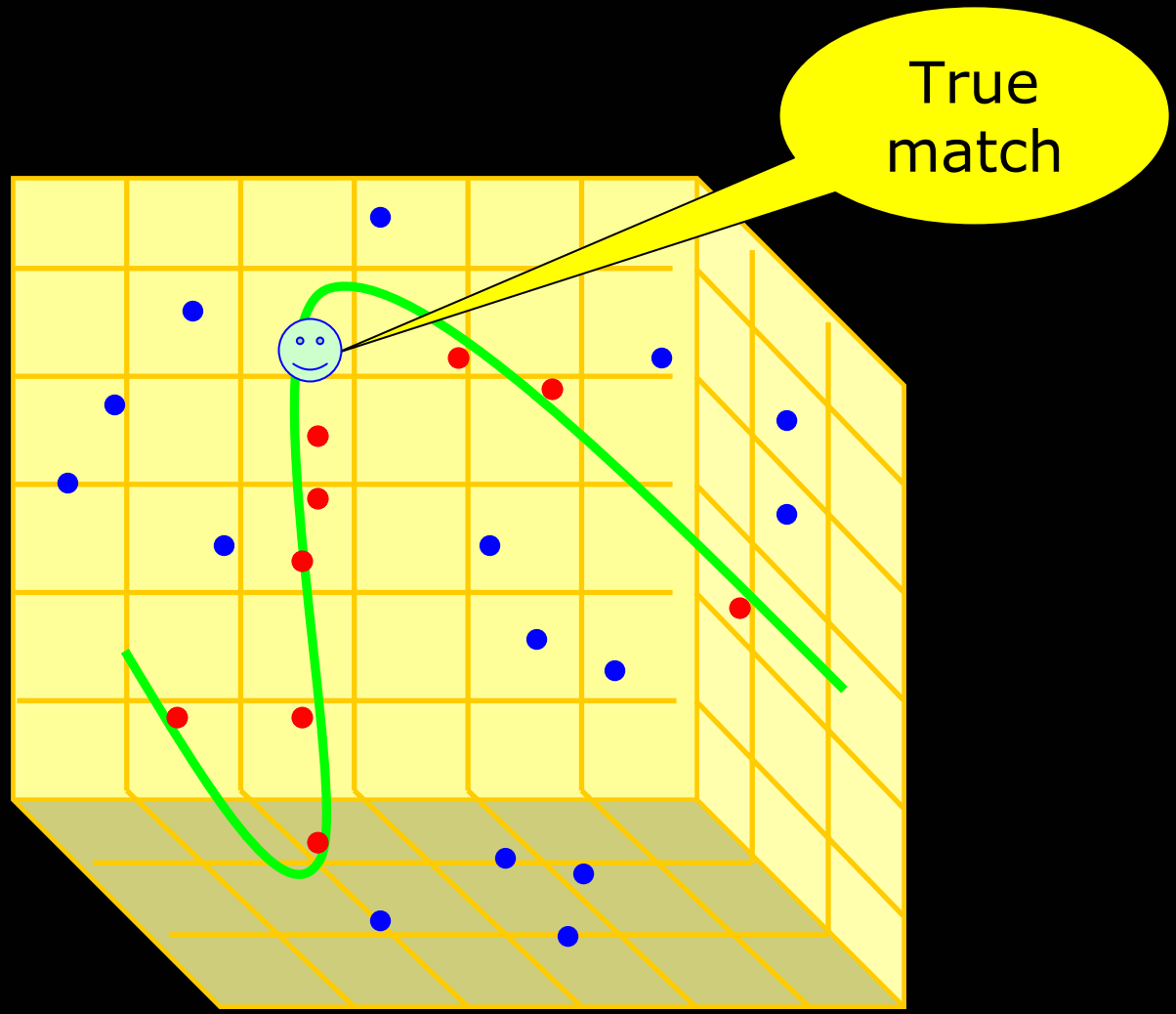




Online – matching



False Matches



How to eliminate the false matches

- Enforce rigidity using *inverse Gramian Test* - Weinshall, '93

$$\frac{\left|x^T B y\right| + \left|x^T B x - y^T B y\right|}{|x|\,\|B\|\,|y|} < \varepsilon$$

- Consistency with lighting → NEXT

Harmonic Images – Linear Basis for Lighting

(Basri and Jacobs '01, Ramamoorthi and Hanrahan '01)

ρ

ρn_z

ρn_x

ρn_y

$\rho(3n_z^2 - 1)$

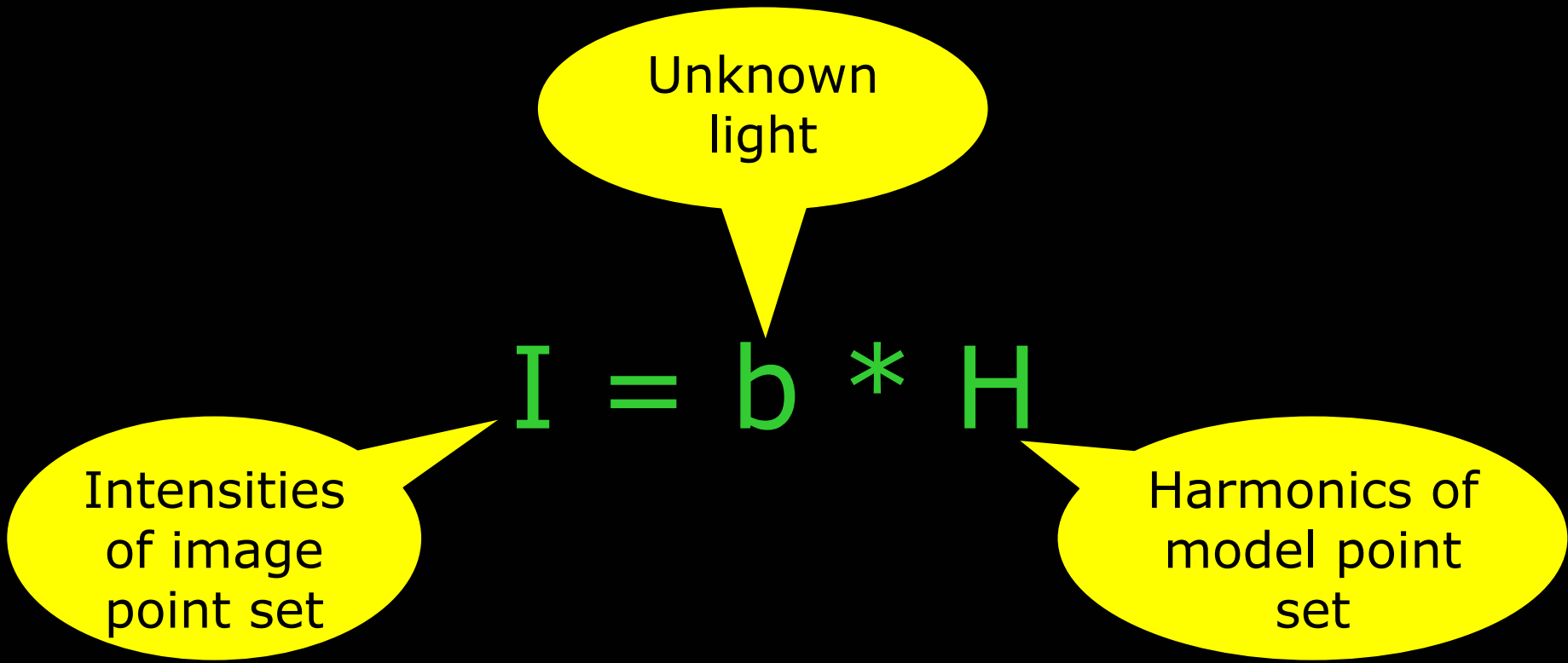
$\rho(n_x^2 - n_y^2)$

$\rho n_x n_y$

$\rho n_x n_z$

$\rho n_y n_z$

Representation by harmonics



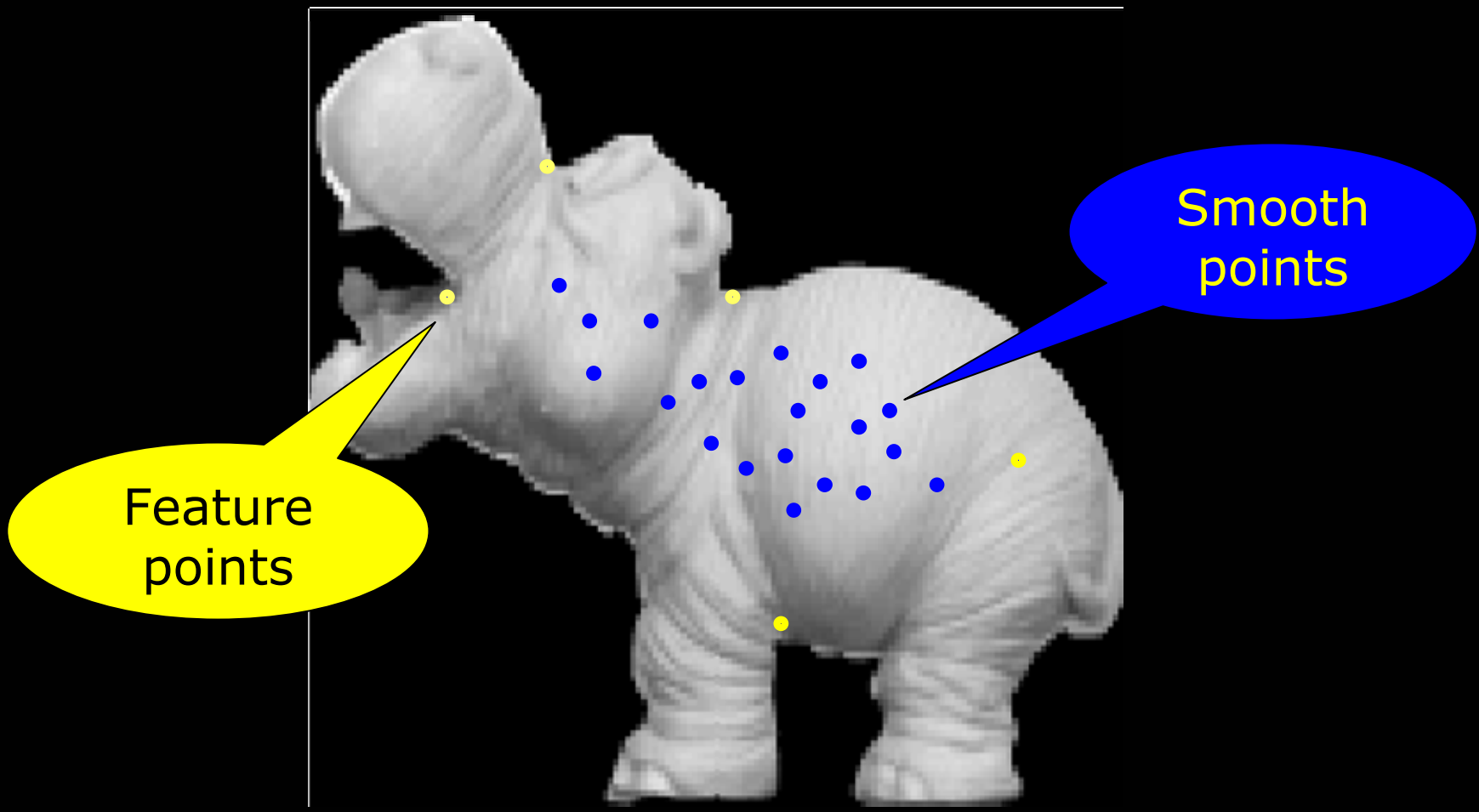
The consistency measure

$I = b * H$ ➡ $\dfrac{\| I - HH^{+}I \|^{2}}{\| I \|^{2} + \varepsilon}$

For *corresponding* image and model sets this is *minimal*

Should we apply it on feature points?

“Smooth points”

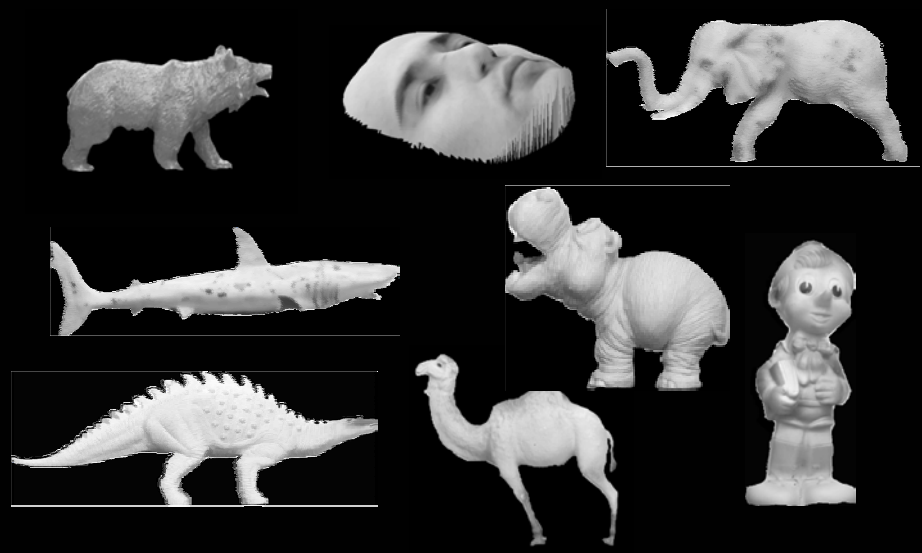


Voting

- Sets of points that pass the lighting test vote for their respective model
- All models receive scores:
- Score = fraction of image sets for which the model appears min
- Once model is selected its corresponding subsets used to determine its pose and lighting

Experiments

- Real 3d objects acquired using laser scanner



- Feature points collected *automatically* using *Harris* corner detector

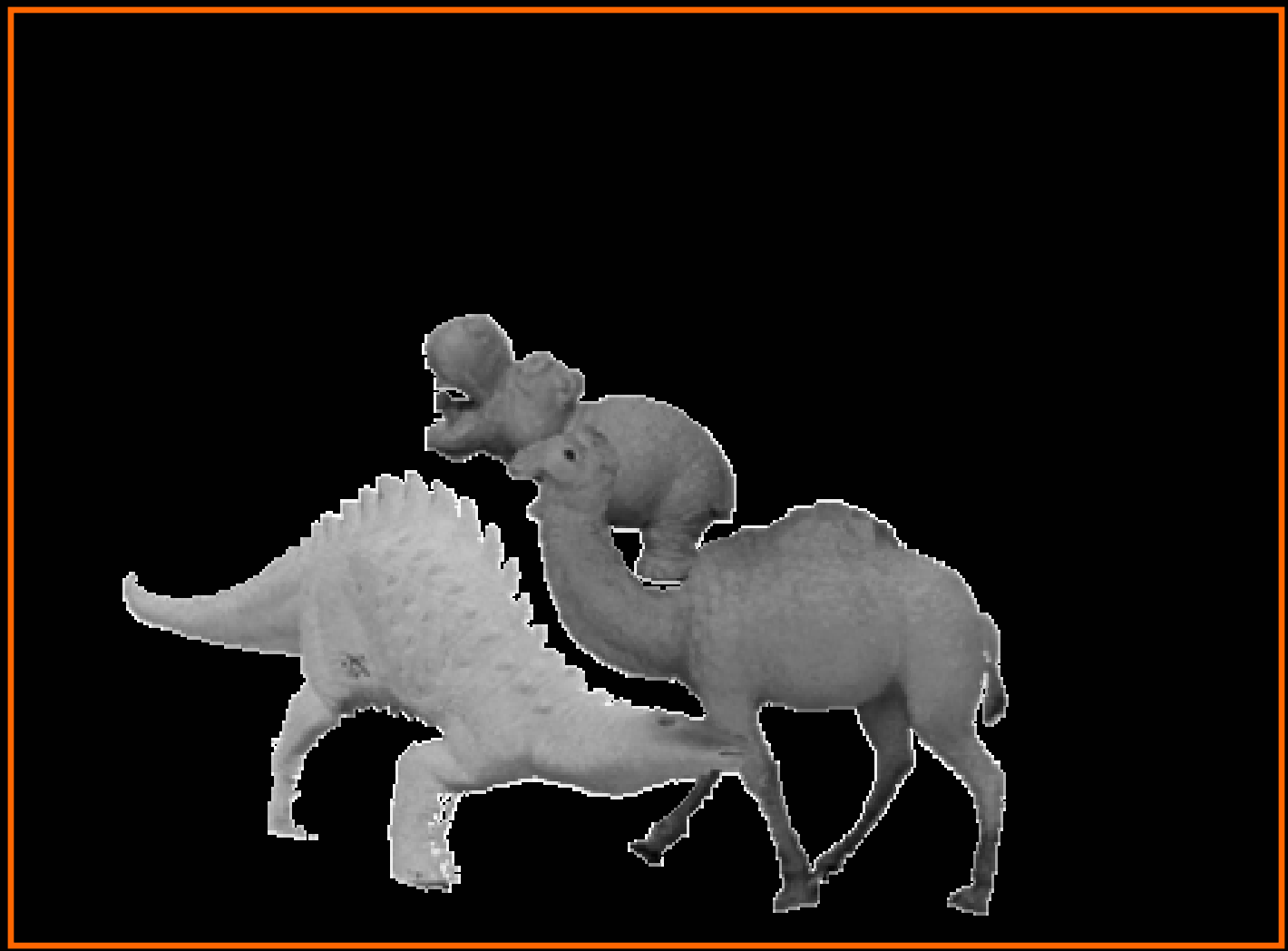
Results

dino
shark
bear
hippo
pinokio
elephant
camel
face

Results

dino
shark
bear
hippo
pinokio
elephant
camel
face

Results



Results – Indoor scene

dino
shark
bear
hippo
pinokio
elephant
camel
face

Results – Outdoor scene

dino
shark
bear
hippo

pinokio
elephant
camel
face

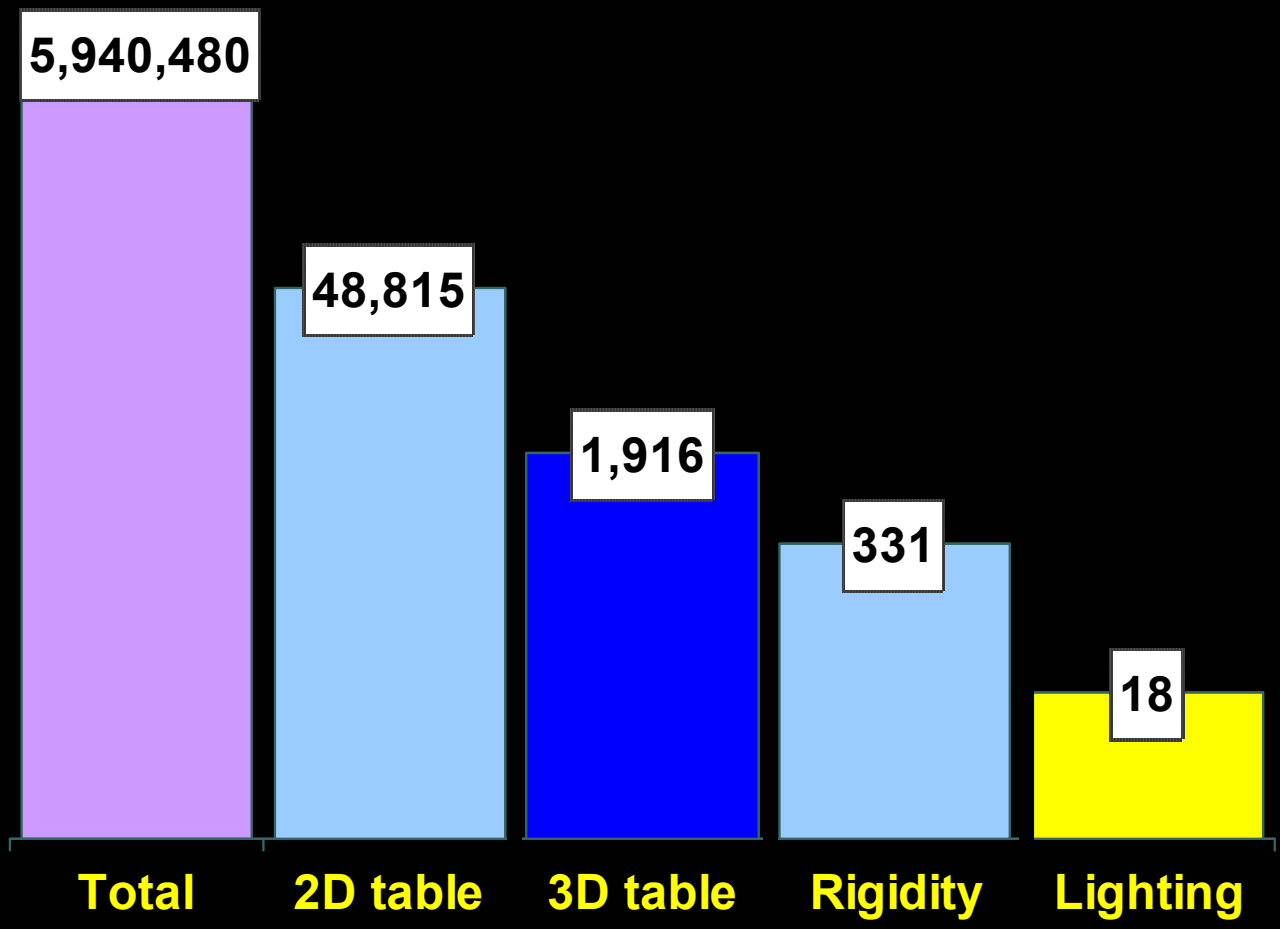
Results – Night Scene

dino
shark
bear
hippo
pinokio
elephant
camel
face

Results – Night Scene 2

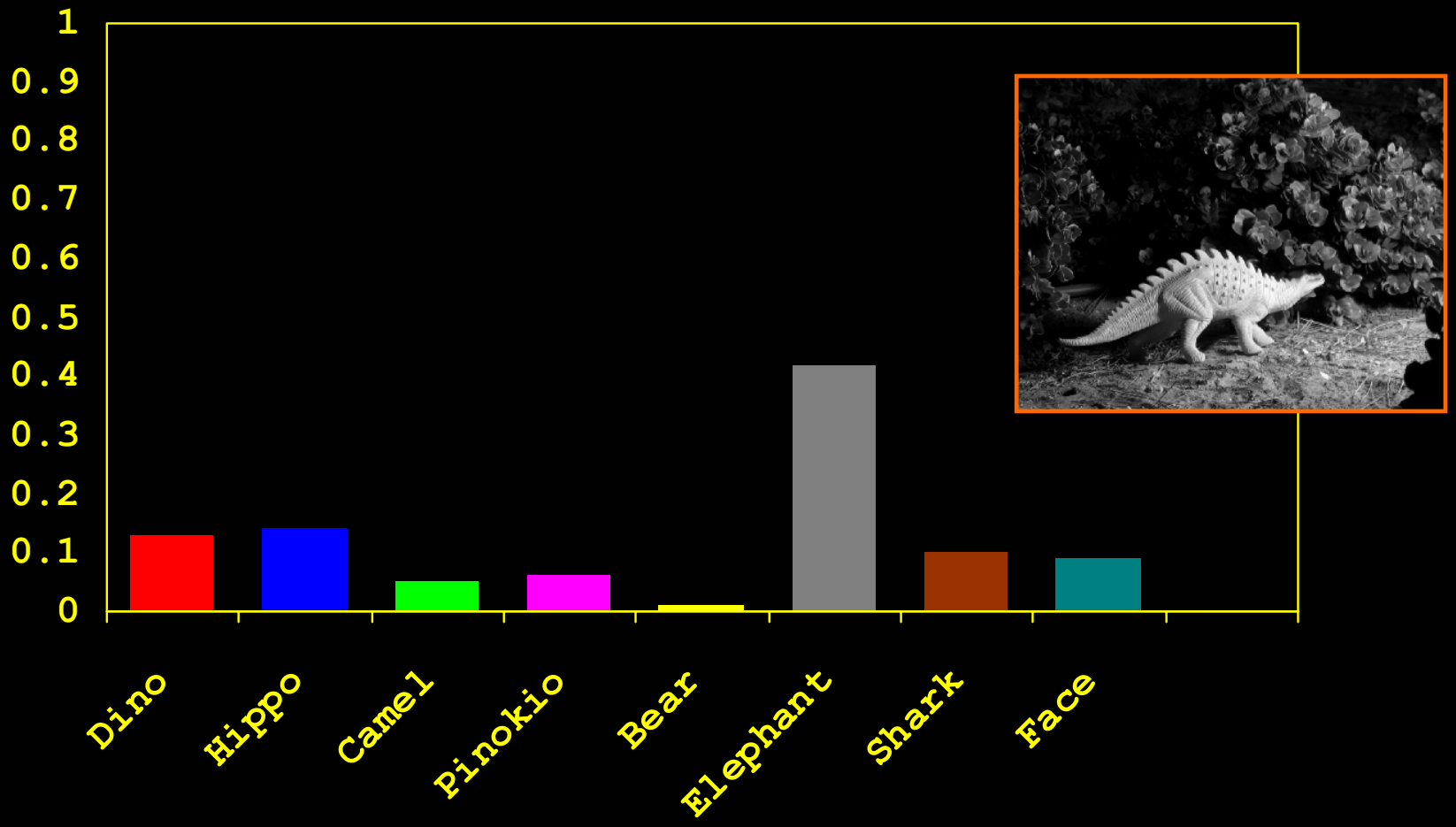
dino
shark
bear
hippo
pinokio
elephant
camel
face

Filtering out matches

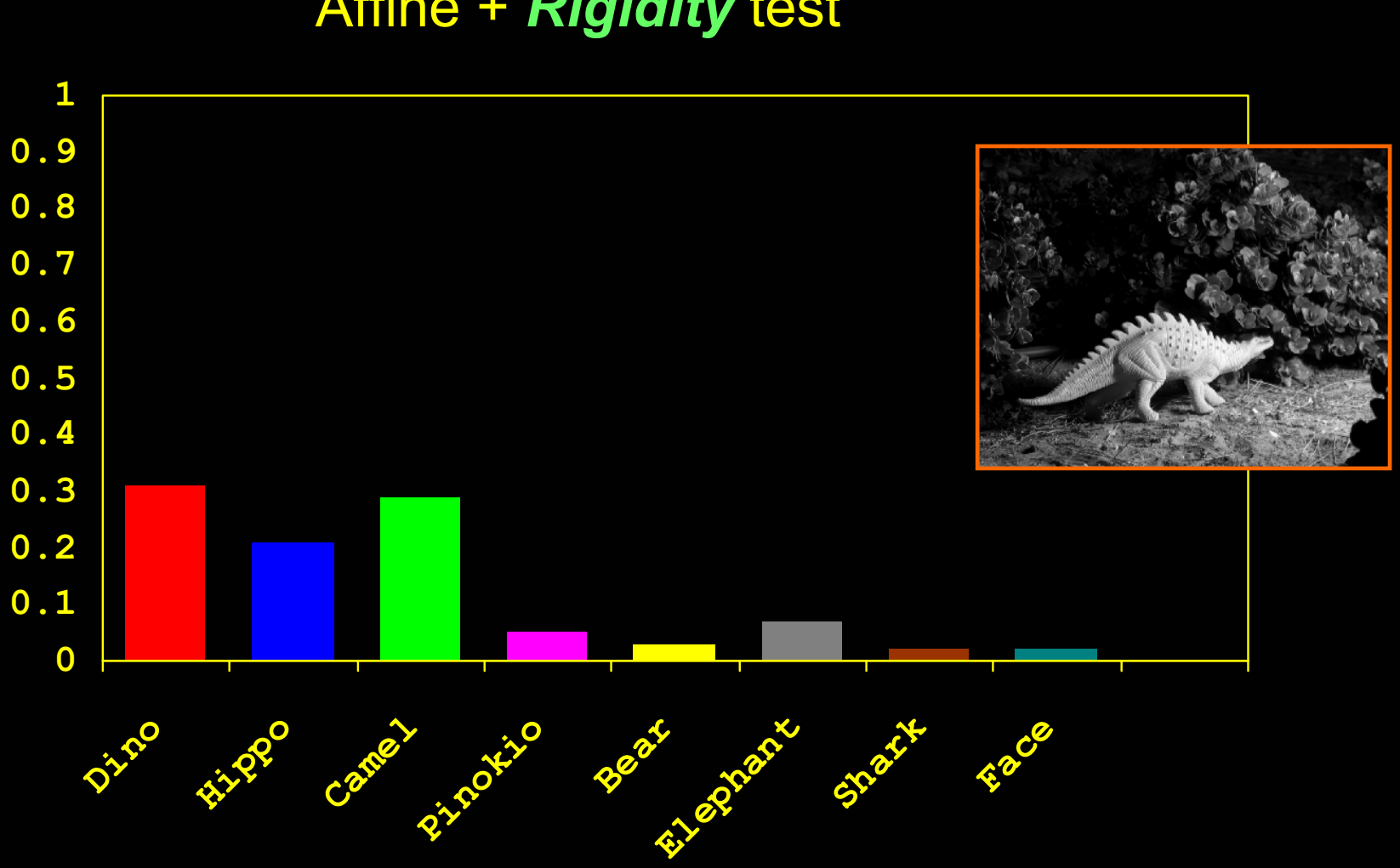


How much does lighting help?

Voting based on ***Affine model***

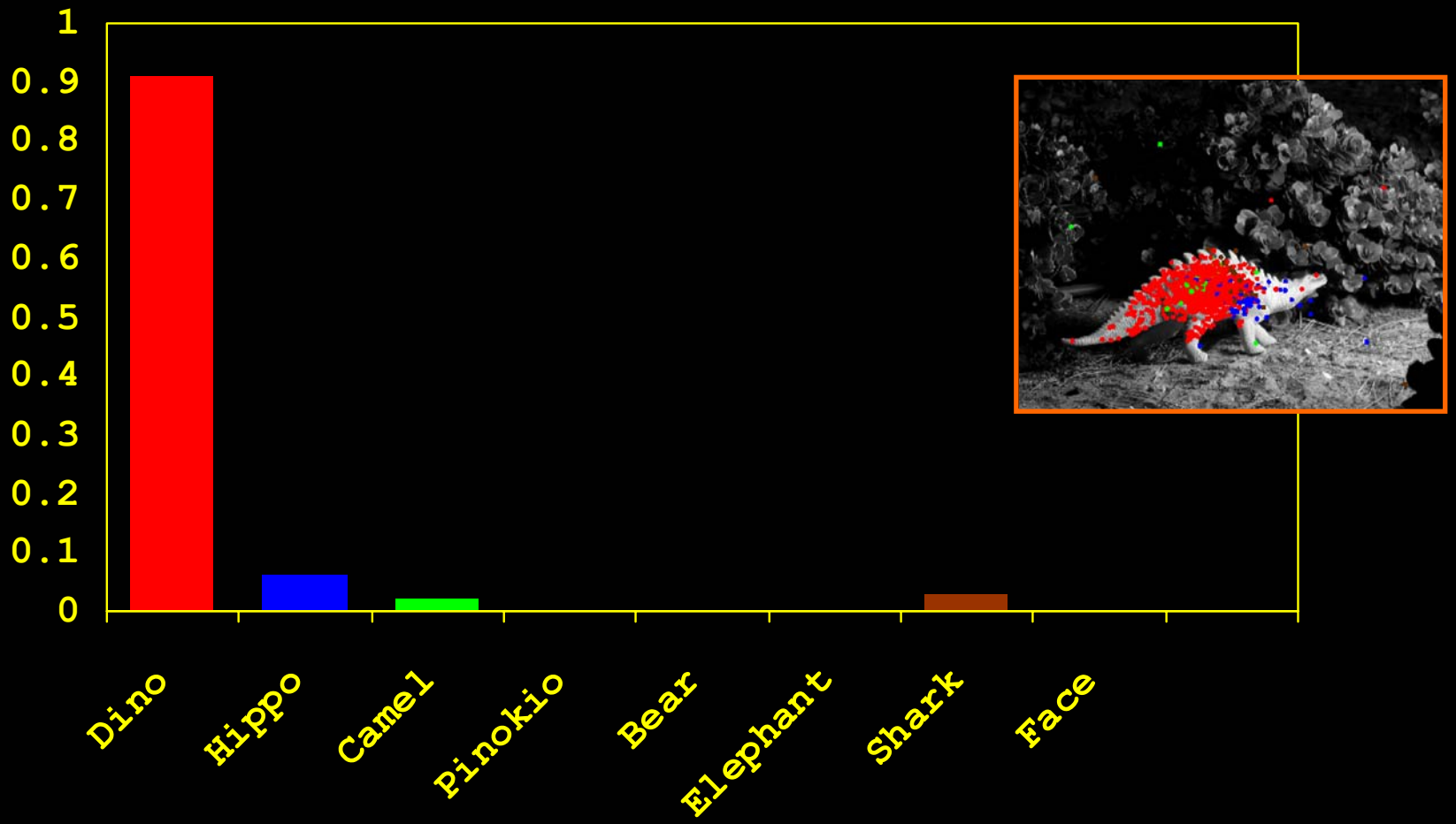


How much does lighting help?



How much does lighting help?

Affine + Rigidity + ***Lighting***



Conclusion

- Identify 3d objects in 2d scenes
  - Unknown *pose, light*
  - *Clutter, occlusions*
  - *General*, real objects
  - *Fast*, efficient
- Combination of *intensity* cues and geometry

Thank you!