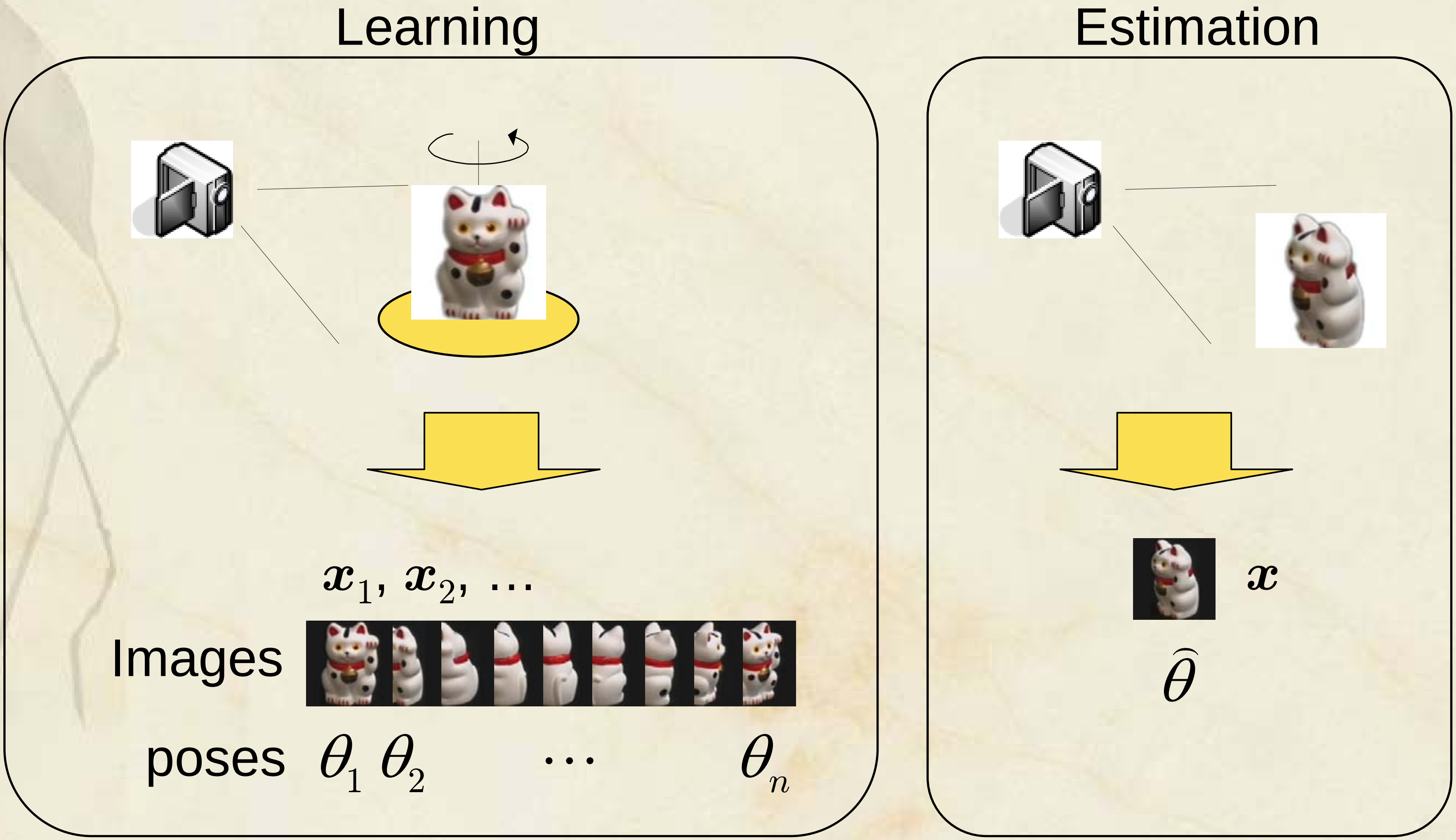
Experimental study on performance of view-based pose estimation

Toru Tamaki HIROSHIMA UNIVERSITY
Hiroyuki Okugawa HIROSHIMA UNIVERSITY
Toshiyuki Amano NAIST
Kazufumi Kaneda HIROSHIMA UNIVERSITY

View-based pose estimation



Learning relations

- Learning set
  - $\{\theta_j, \boldsymbol{x}_j\}$
    $(i=1, 2, \dots, n)$
- Relations
  - Nonlinear $\theta_j = f(\boldsymbol{x}_j)$
  - Linear $\theta_j = F\boldsymbol{x}_j$
- Estimation
  - Nonlinear $\theta = f(\boldsymbol{x})$
  - Linear $\theta = F\boldsymbol{x}$

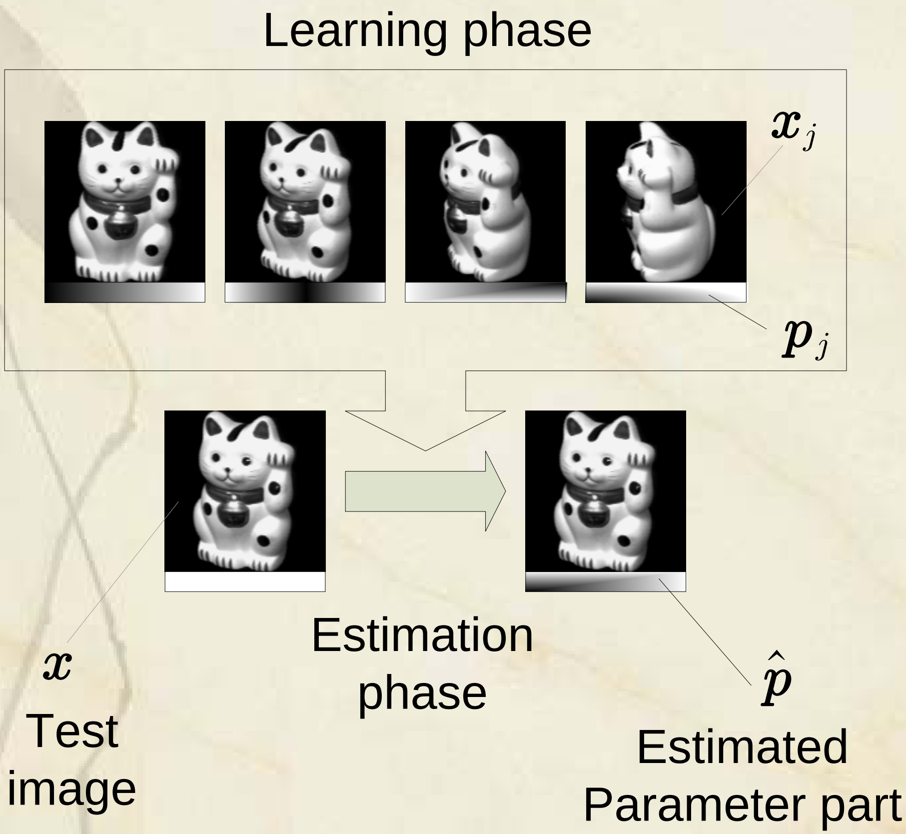
- Nonlinear methods
  - Parametric Eigenspace method
    - (Murase, 1995)
  - Kernels
    - (Melzer, 2003)
    - (Ando, 2005)
  - Manifold learning

Learning relations

- Learning set
  - $\{\theta_j, \boldsymbol{x}_j\}$
    $(i=1, 2, \ldots , n)$
- Relations
  - Nonlinear $\theta_j = f(\boldsymbol{x}_j)$
  - Linear $\theta_j = F\boldsymbol{x}_j$
- Estimation
  - Nonlinear $\theta = f(\boldsymbol{x})$
  - Linear $\theta = F\boldsymbol{x}$

- Linear methods
  - Linear regression
    - (Okatani, 2000)
  - Cyclic permutation
    - (Tamaki, 2007)
  - EbC
    - (Amano, 2006/2007)

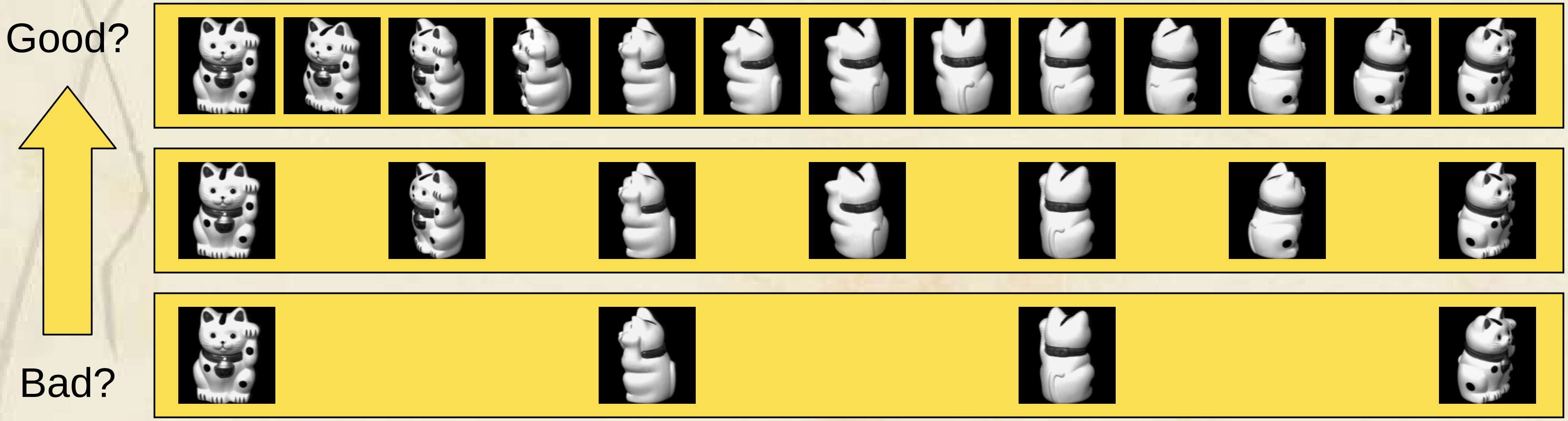
Overview of *EbC*



- EbC: "*Estimation-by-Completion*"
- Learn
  - Image part x_j
  - Parameter part p_j
  - Compute Eigenspace
- Estimate pose
  - A test image has no parameter part
  - Completed as missing image area

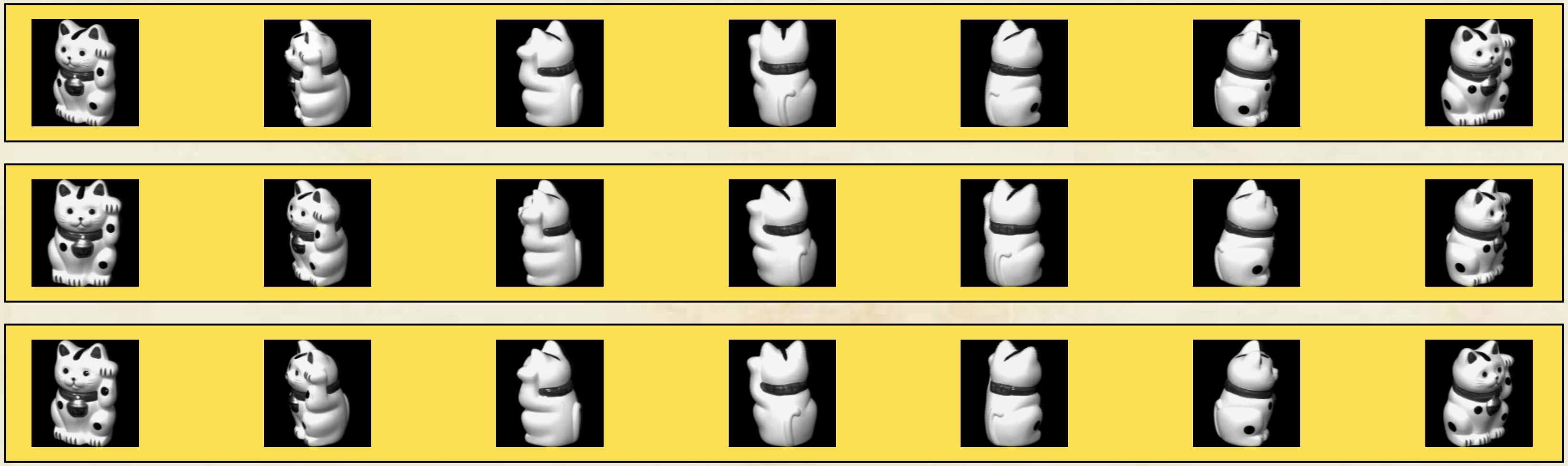
Questions to investigate

- Performance depends on the number of learning images.
  - Few images: bad estimation
  - Many images: better performance
- Is it really? How many images are enough?



Questions to investigate

- Performance depends on the number of learning images.
- What is an appropriate set of images when we fix the number of images?
  - Any set is enough?



Learning image set

Definition of a learning set :

$$S_{i,s} = \{\boldsymbol{x}_{ik+s}\}$$

$\boldsymbol{x}_\theta$: images at θ

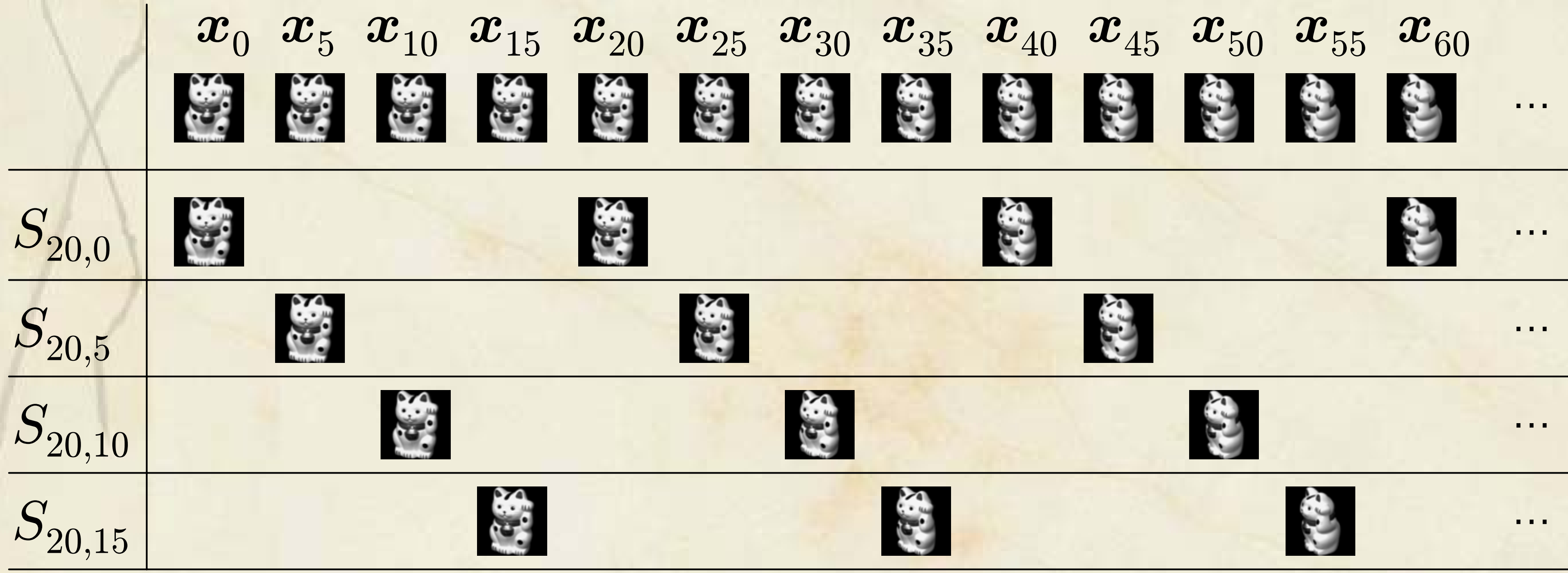
i : sample span [deg]

s : start angle [deg]

$k = 0, 1, \ldots, n_i - 1$

$n_i = 360/i$

Example :



Performance evaluation

Root mean square error (RMSE):

$$RMSE_{i,s} = \sqrt{\frac{1}{72-n_i}\sum_{x_j \notin S_{i,s}}\left(\hat{\theta}_j - \theta_j\right)^2}$$

$\theta:$ true angle

$\hat{\theta}:$ estimated angle

Exclude learned images

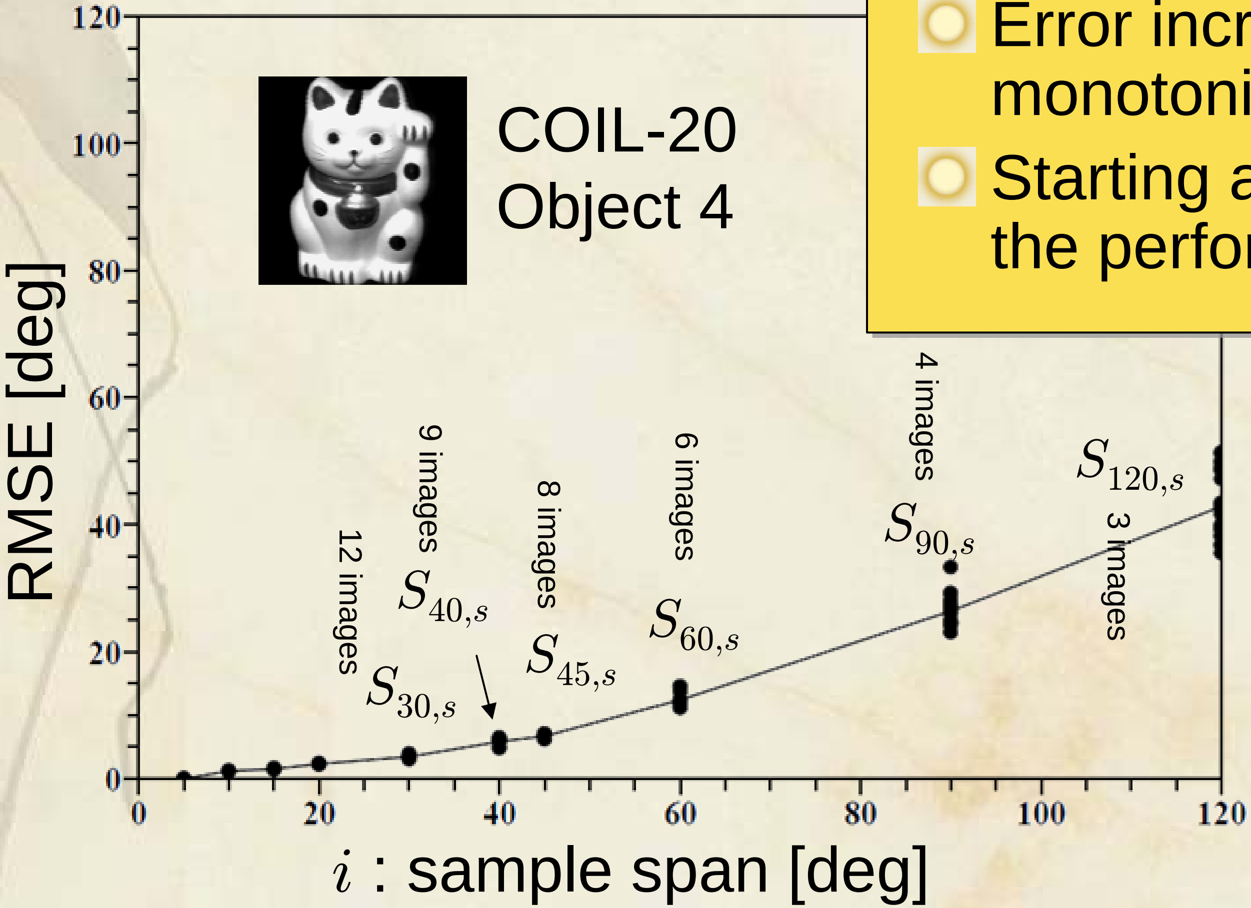
sample spans:

$$i = 5,10,15,20,30,40,45,60,90,120$$

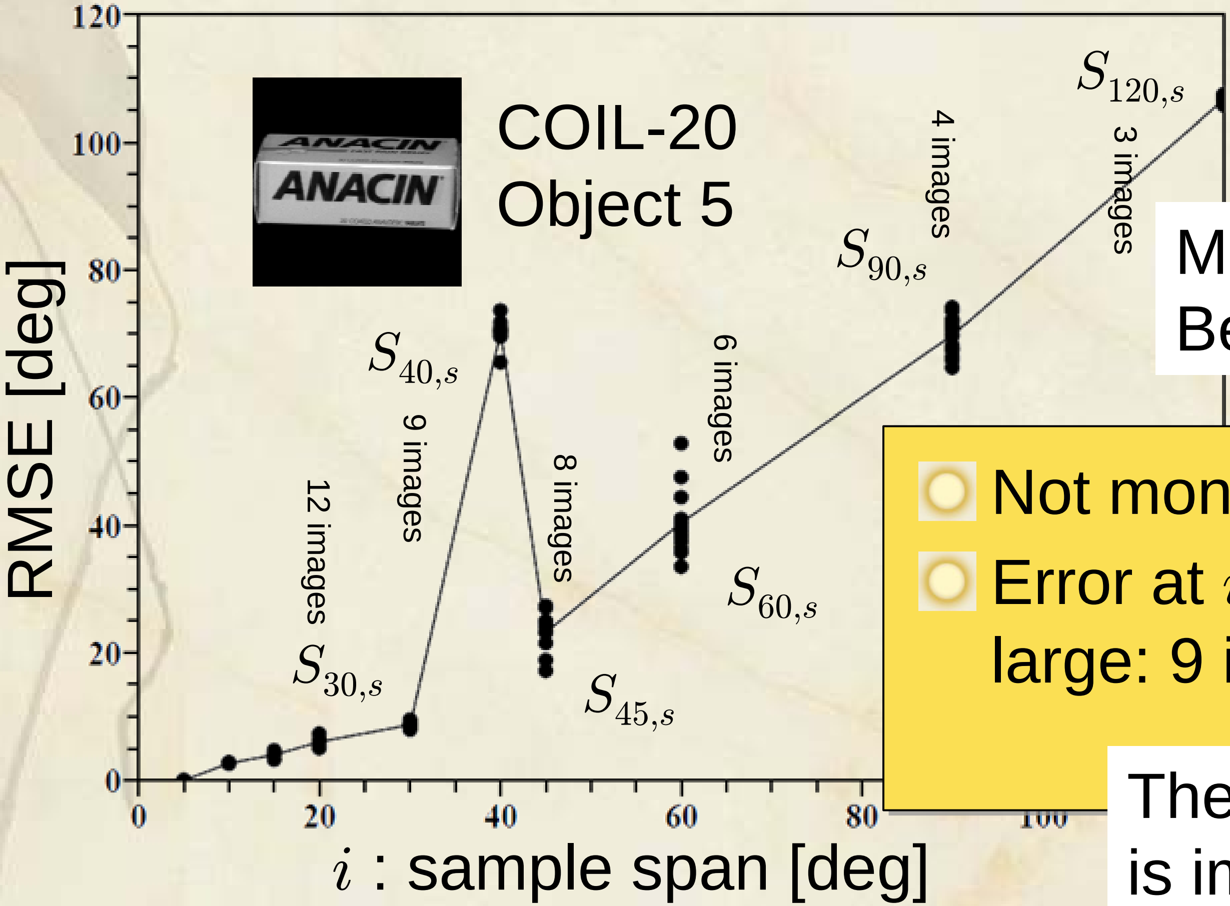
(divisors of 360 [deg])

Experimental results 1: moderate case

- Error increases monotonically
- Starting angle doesn’t affect the performance



Experimental results 2: performance dip at 40 deg.



~~More the images,~~
~~Better the performance~~

- Not monotonically
- Error at $i\!=\!40$ [deg] is very large: 9 images are learned

The number of images is important!

Examples of learning sets

$S_{60,0}$ 6 images

$S_{45,0}$ 8 images

$S_{40,0}$ 9 images

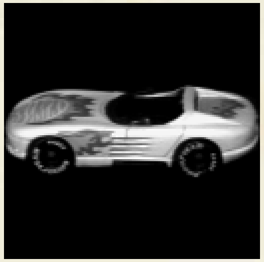
Worst !

$S_{30,0}$ 12 images

Objects that have performance dip at 40 deg.

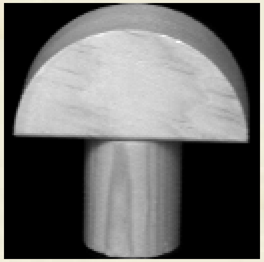
Object 5

Object 6



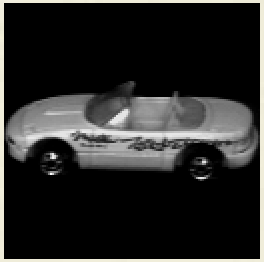
Object 9

Object 11



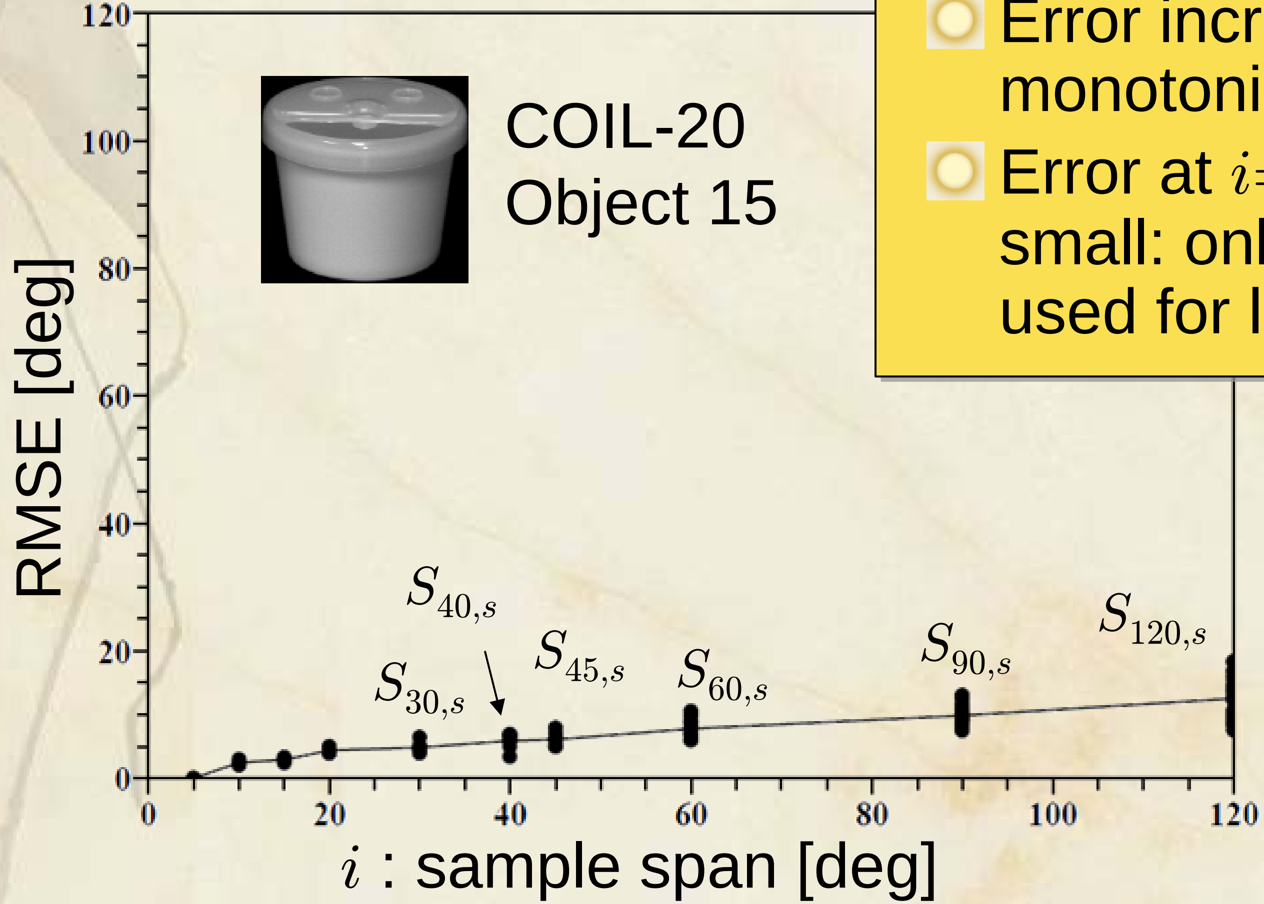
Object 14

Object 19



- What property affect the performance?
  - Future work….

Experimental results 3: keeping good performance



- Error increases monotonically
- Error at $i=120$ [deg] is so small: only 3 images are used for learning

Objects that keep good performance

COIL-20
Object 15

 $S_{120,0}$

COIL-20
Object 12

 $S_{120,0}$

COIL-20
Object 20

 $S_{120,0}$

- Round shape may affect the performance
  - Also future work…

Conclusions

- Performance evaluation of EbC
  - a view-based pose estimation
- Experimental results:
  - Some objects have the performance dip
  - Some objects keep good performance
- Future work
  - To investigate the relationship between performance and object shape