Watermarking with Retrieval System

Ee-Chien Chang Sujoy Roy

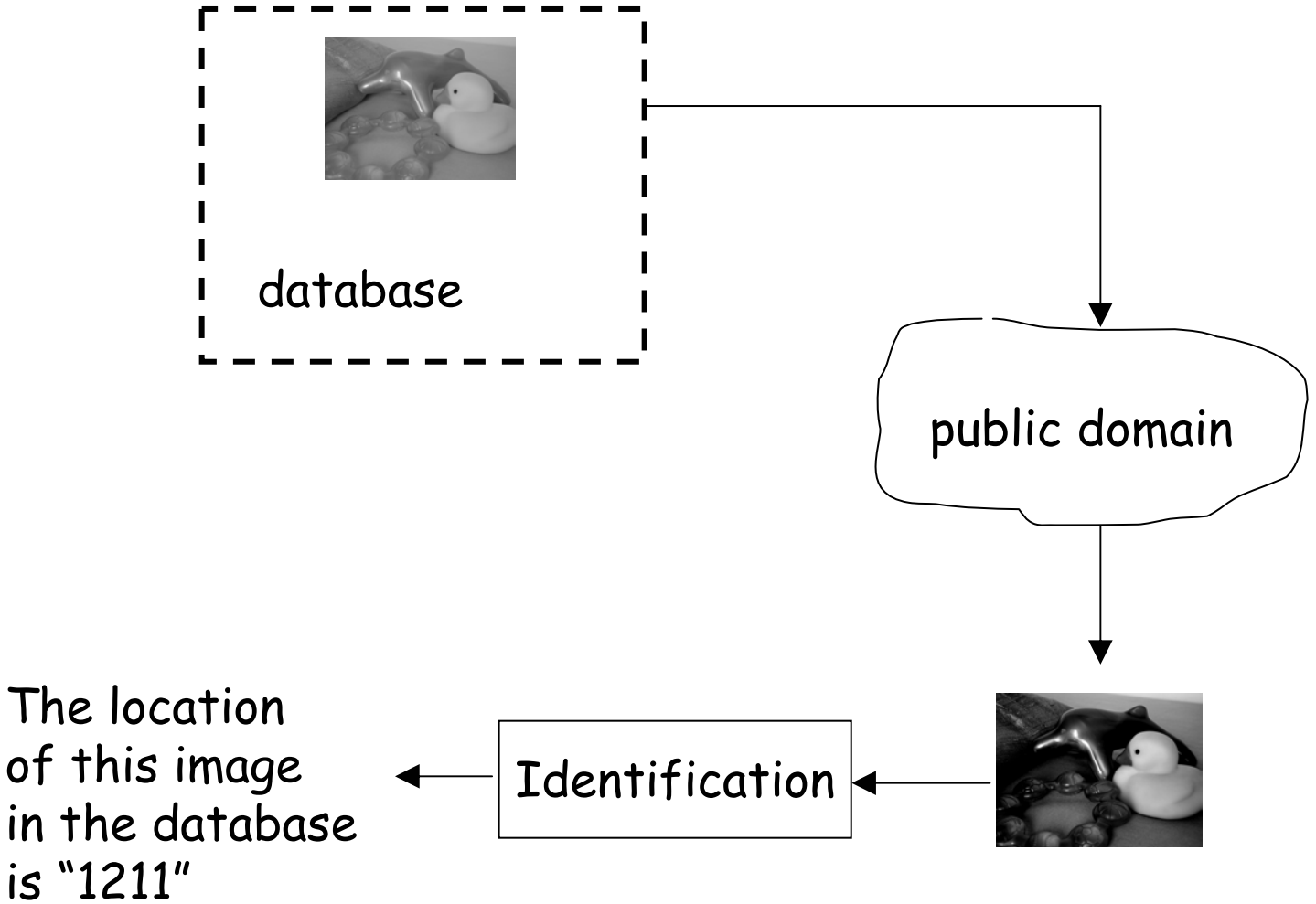
School Of Computing
National University of Singapore

{changec,sujoyroy}@comp.nus.edu.sg
www.comp.nus.edu.sg/~changec

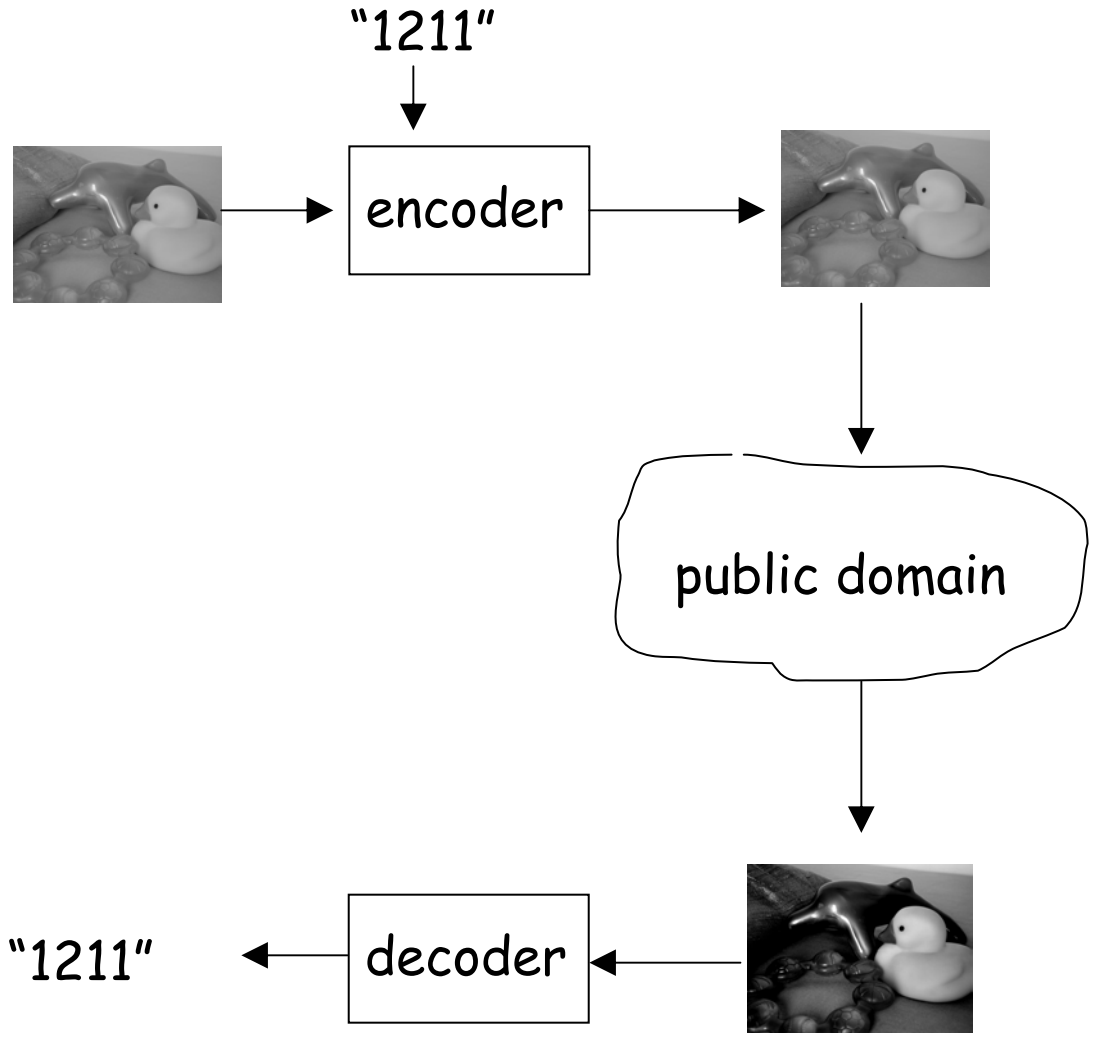
Outlines of this talk:

- Describe the identification/tracking problem
- Can be addressed by 1) Watermarking, or
  2) Retrieval systems.
- Propose a tradeoff of 1) and 2).

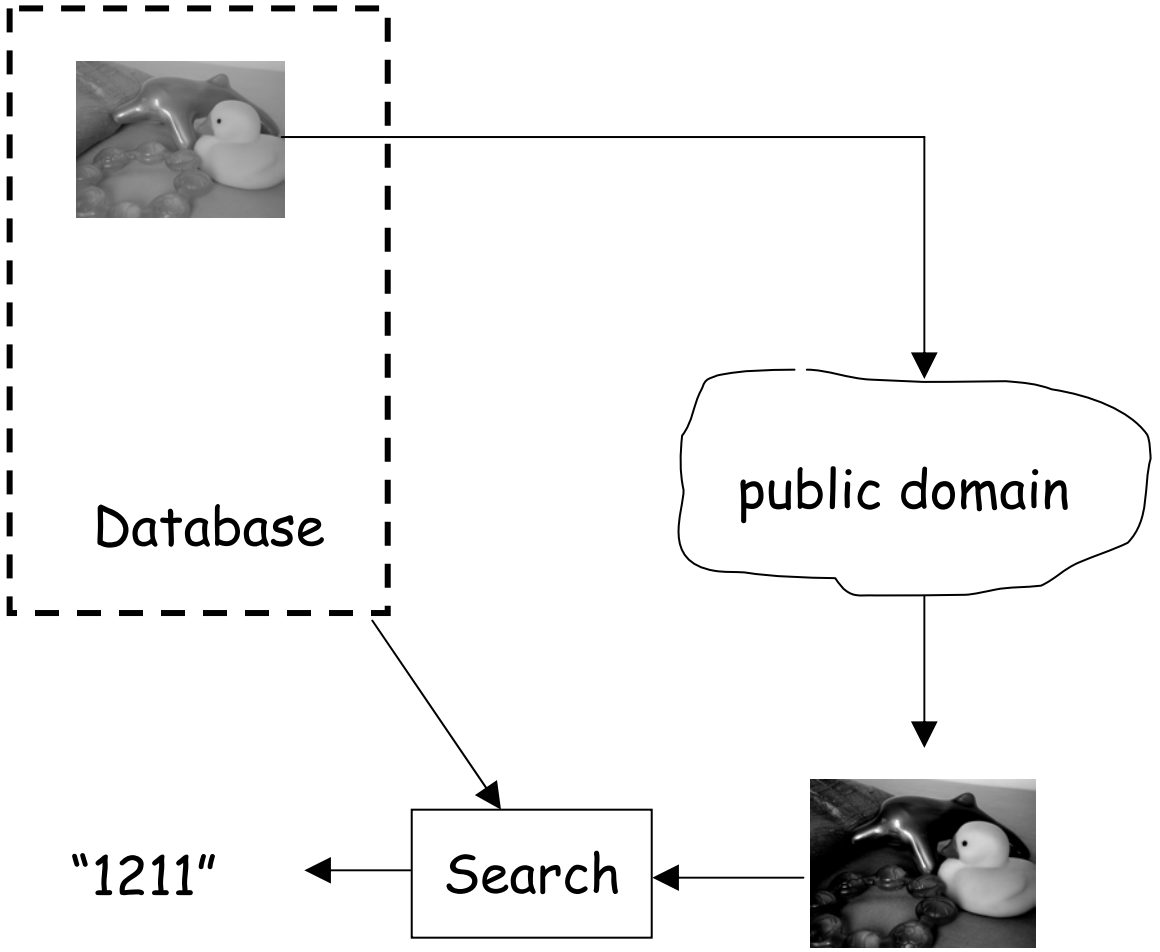
Identification/tracking



Identification/tracking by **Watermarking**



Identification/tracking by **Retrieval**



Research Goal:

- Retrieval systems: Slow searching (dimensionality curse),
  might causes ambiguity,
  no distortion.

- Watermarking: Fast decoding,
  resolve ambiguity,
  distortion.

- We propose to study tradeoff of these two extremes.

Image Space

Only watermarking...

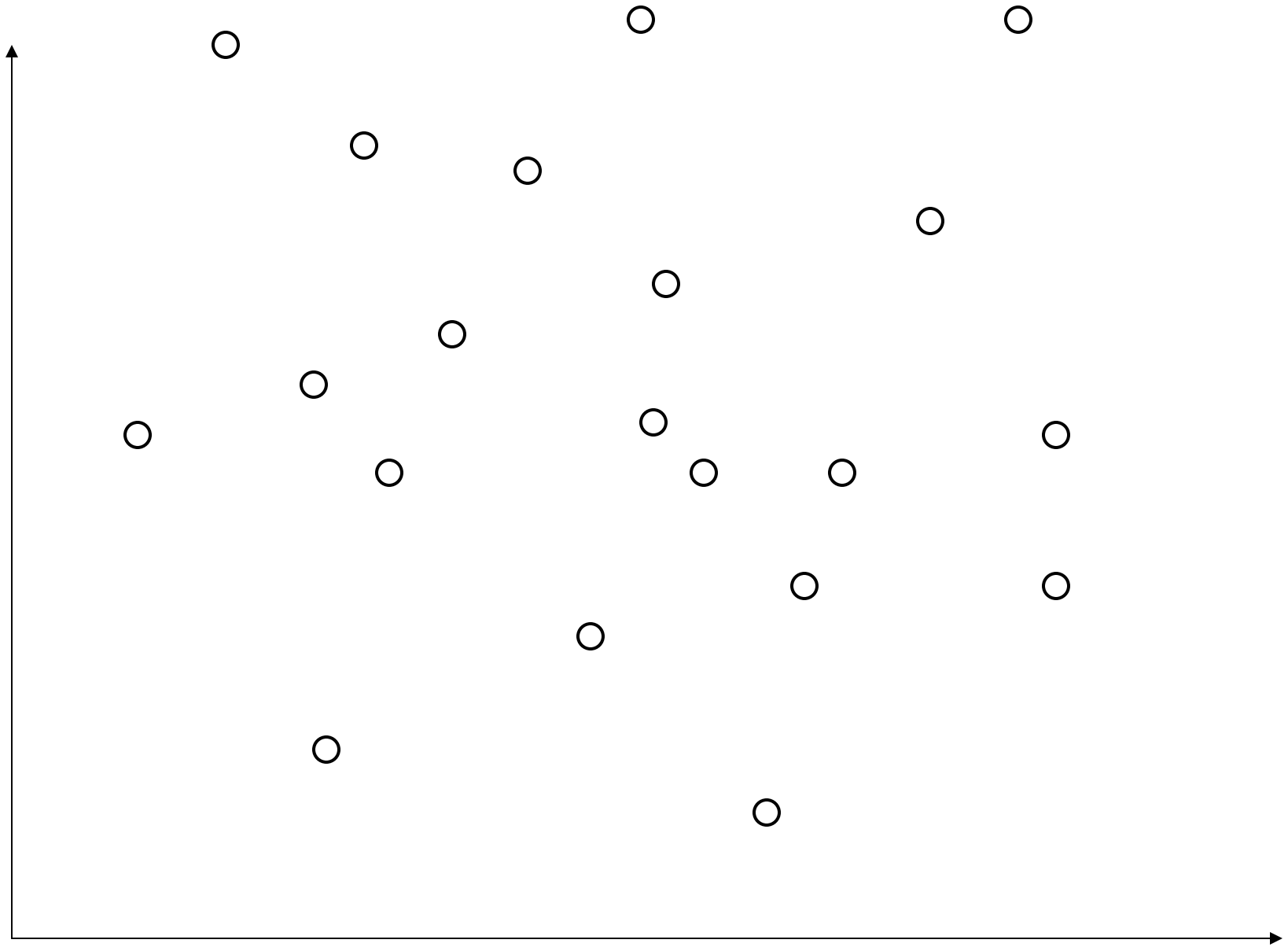
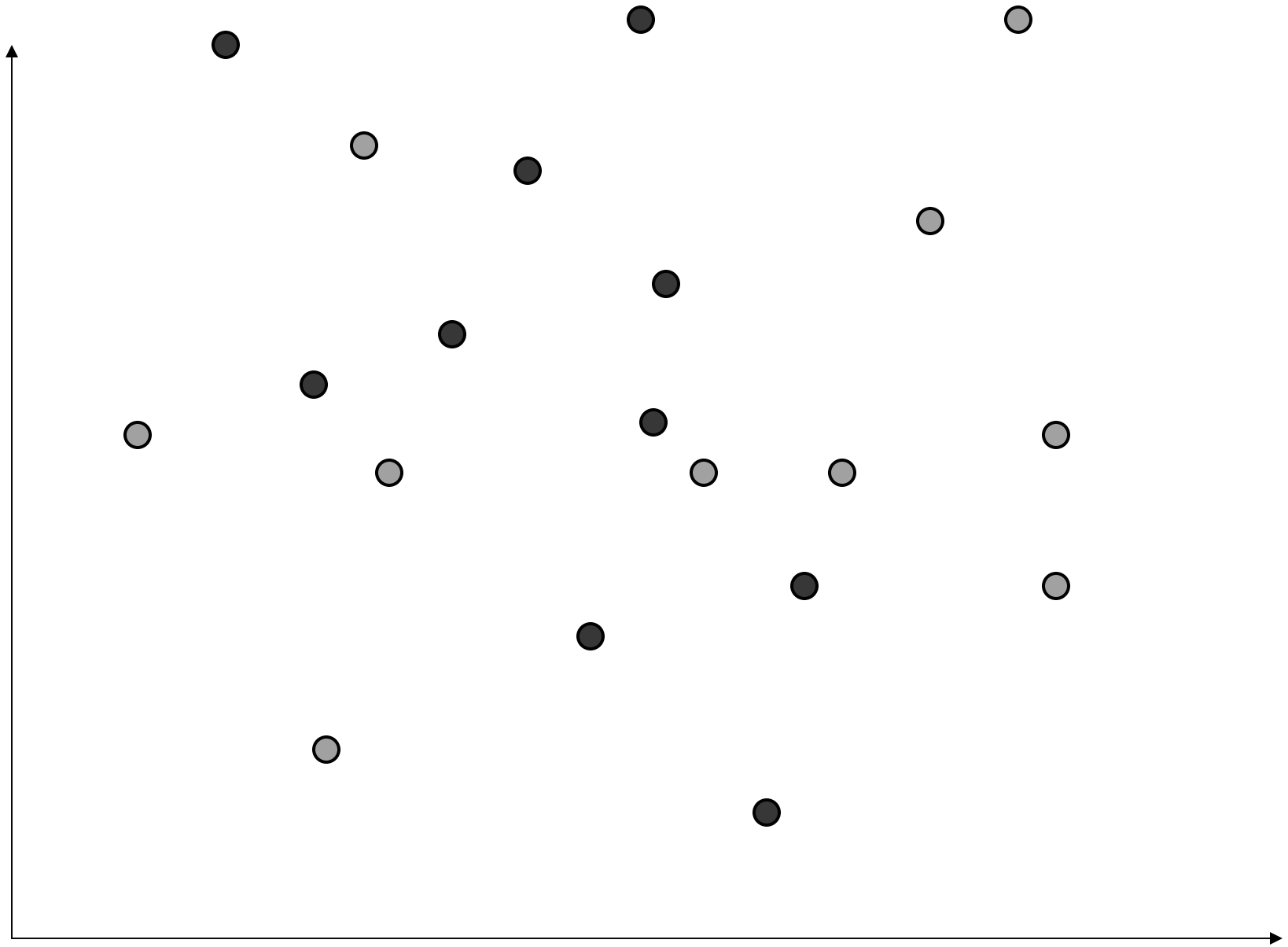


Image Space

Only watermarking...



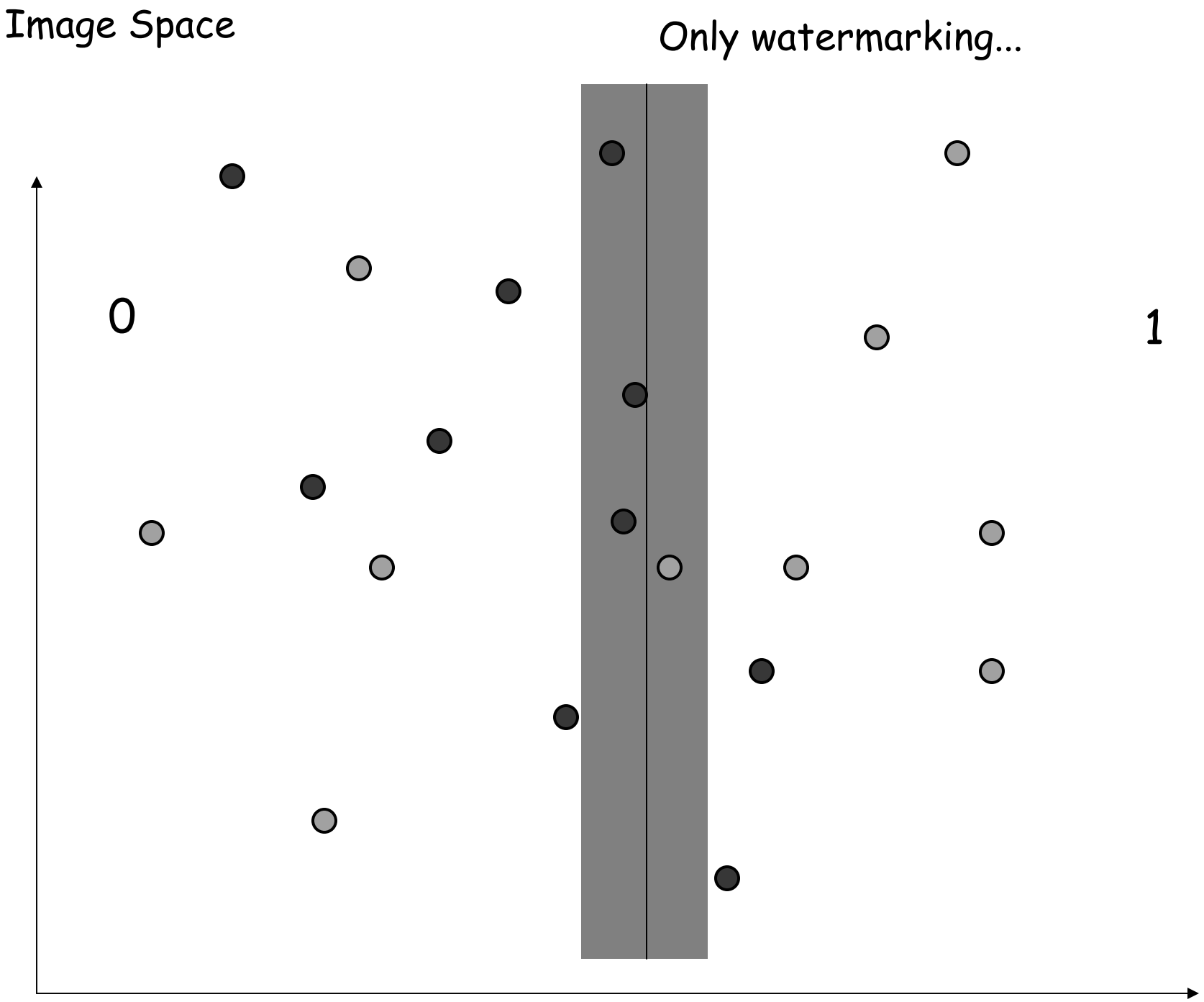


Image Space
Only watermarking...
0
1

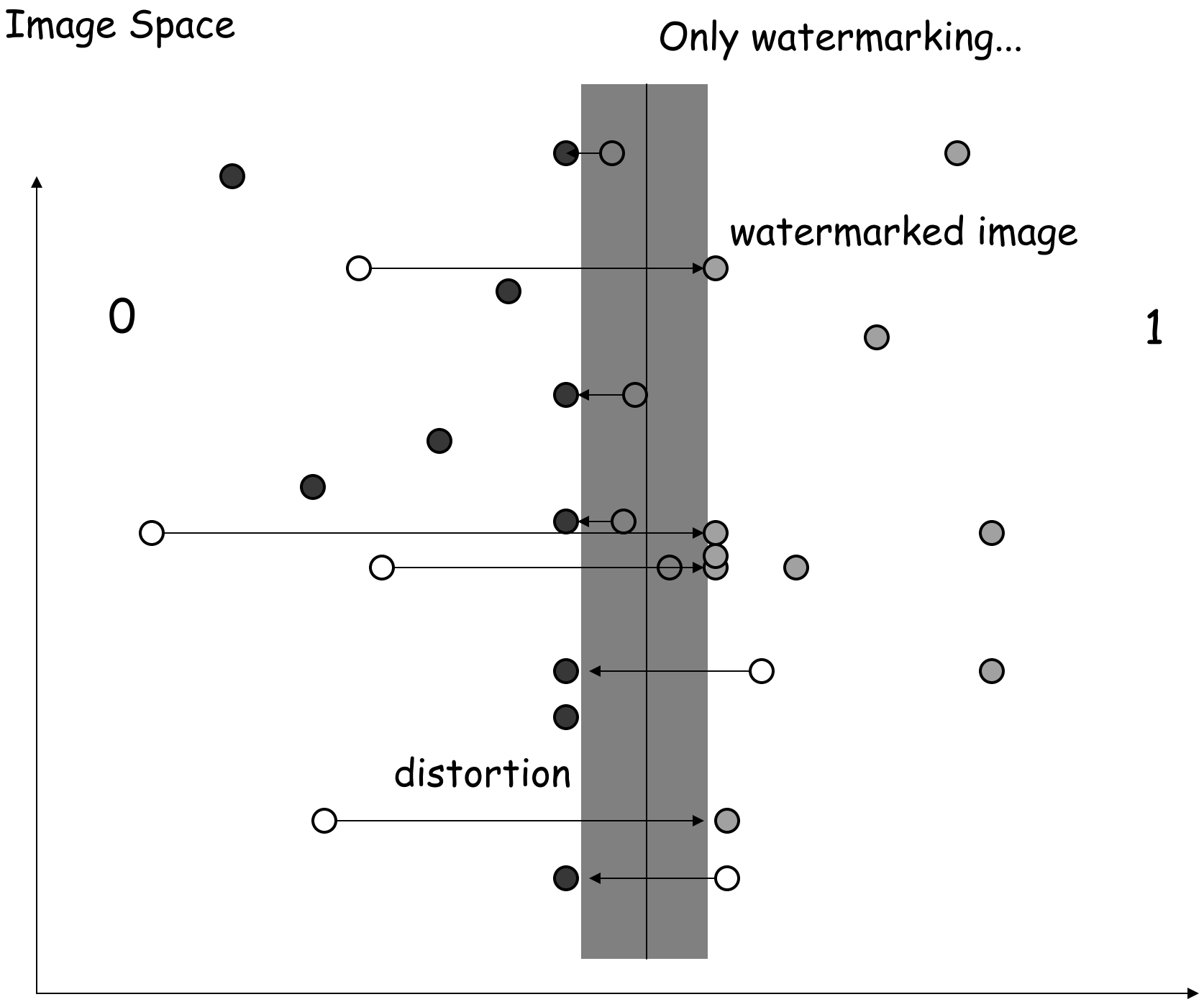


Image Space
Only watermarking...
watermarked image
0
1
distortion

Image Space

Only by Retrieval...

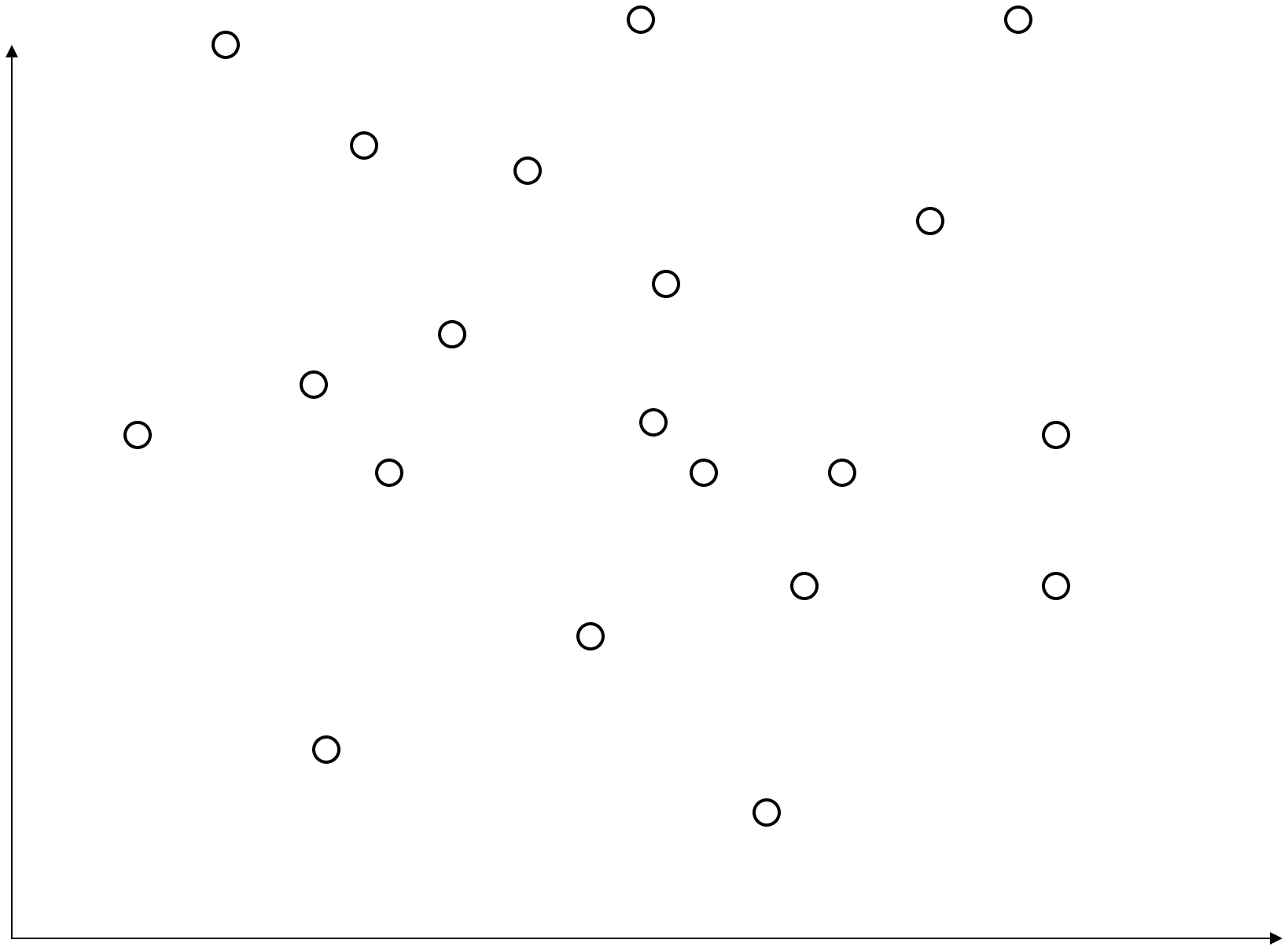
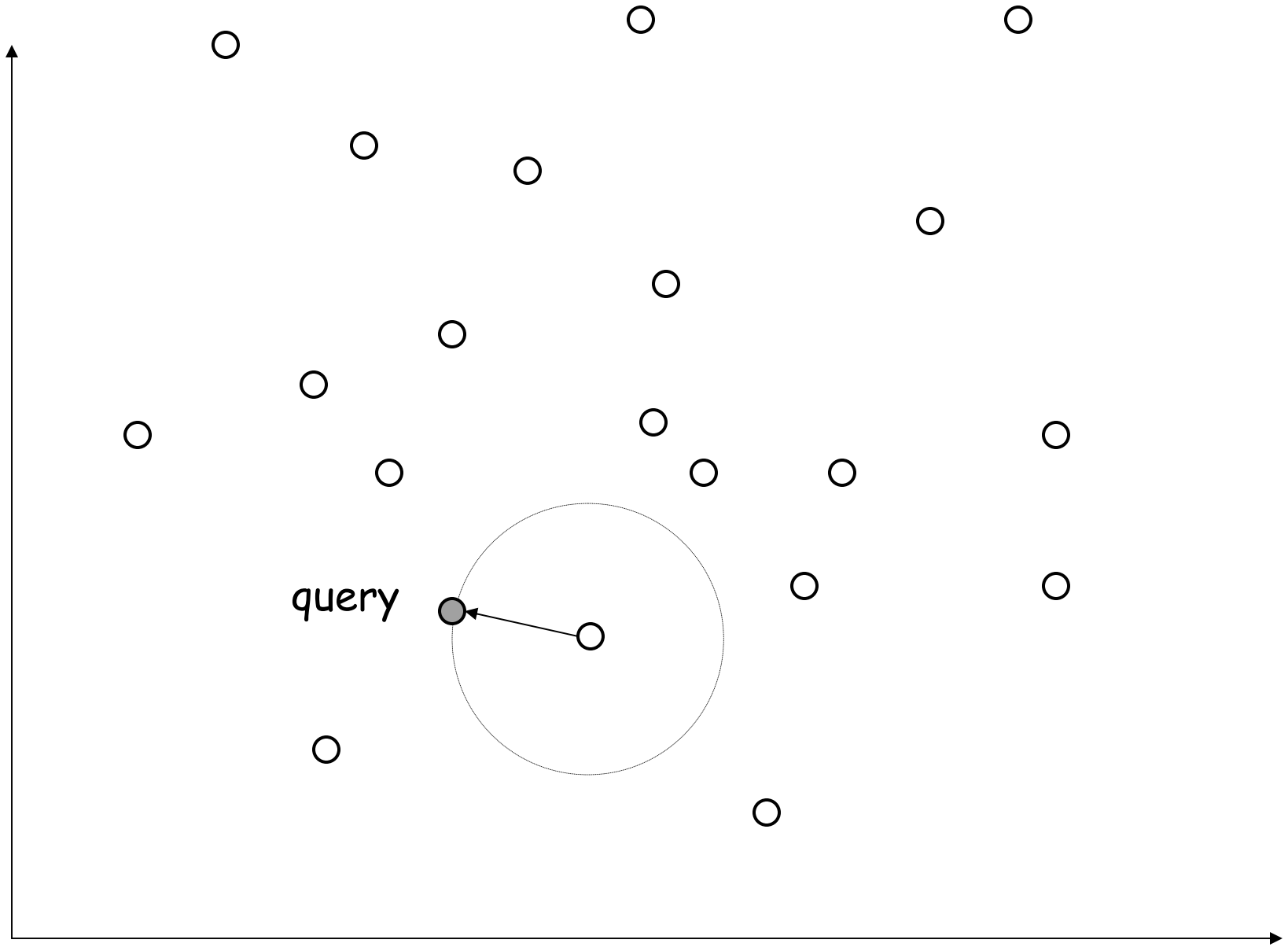


Image Space

Only by Retrieval...



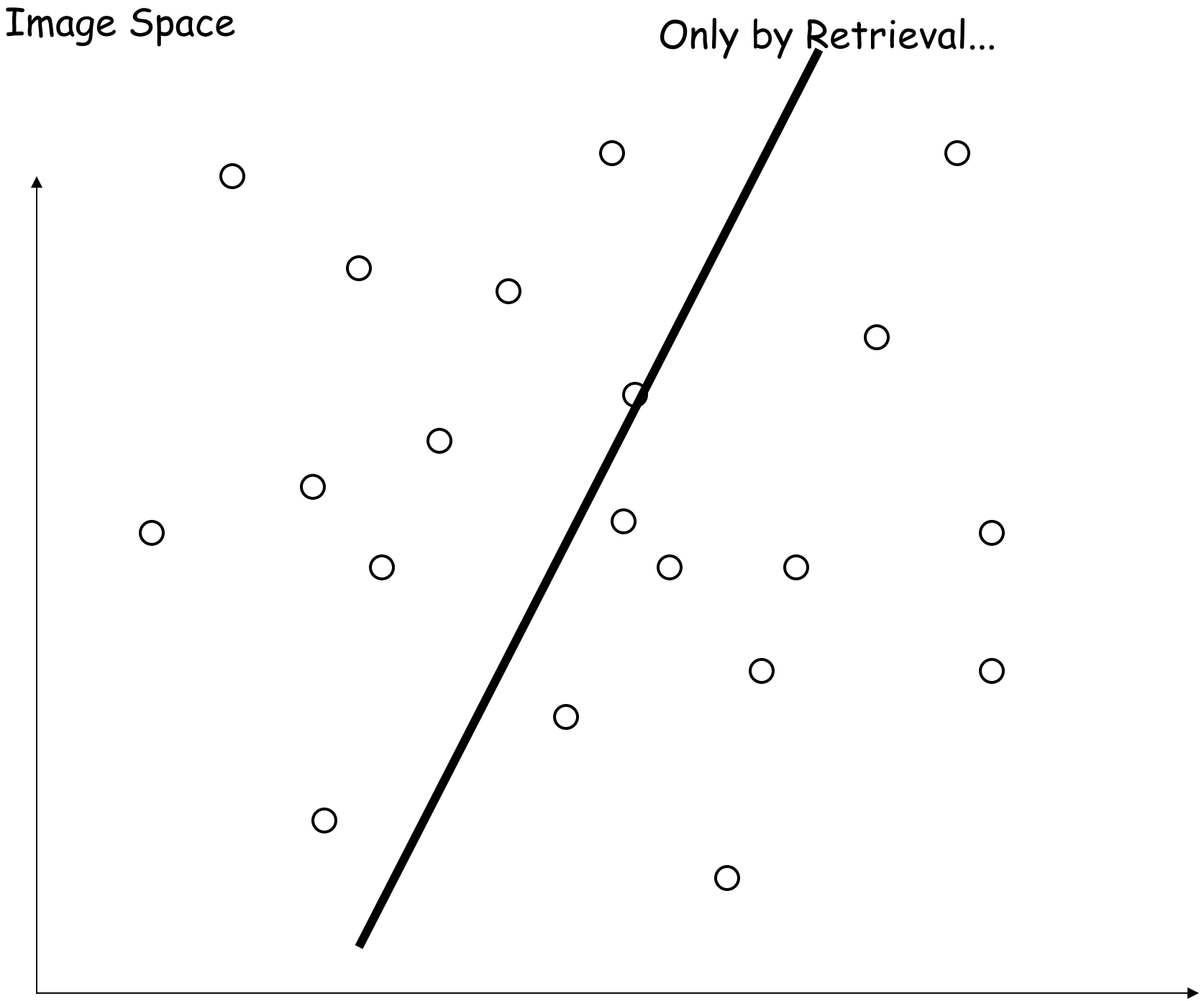


Image Space
Only by Retrieval...

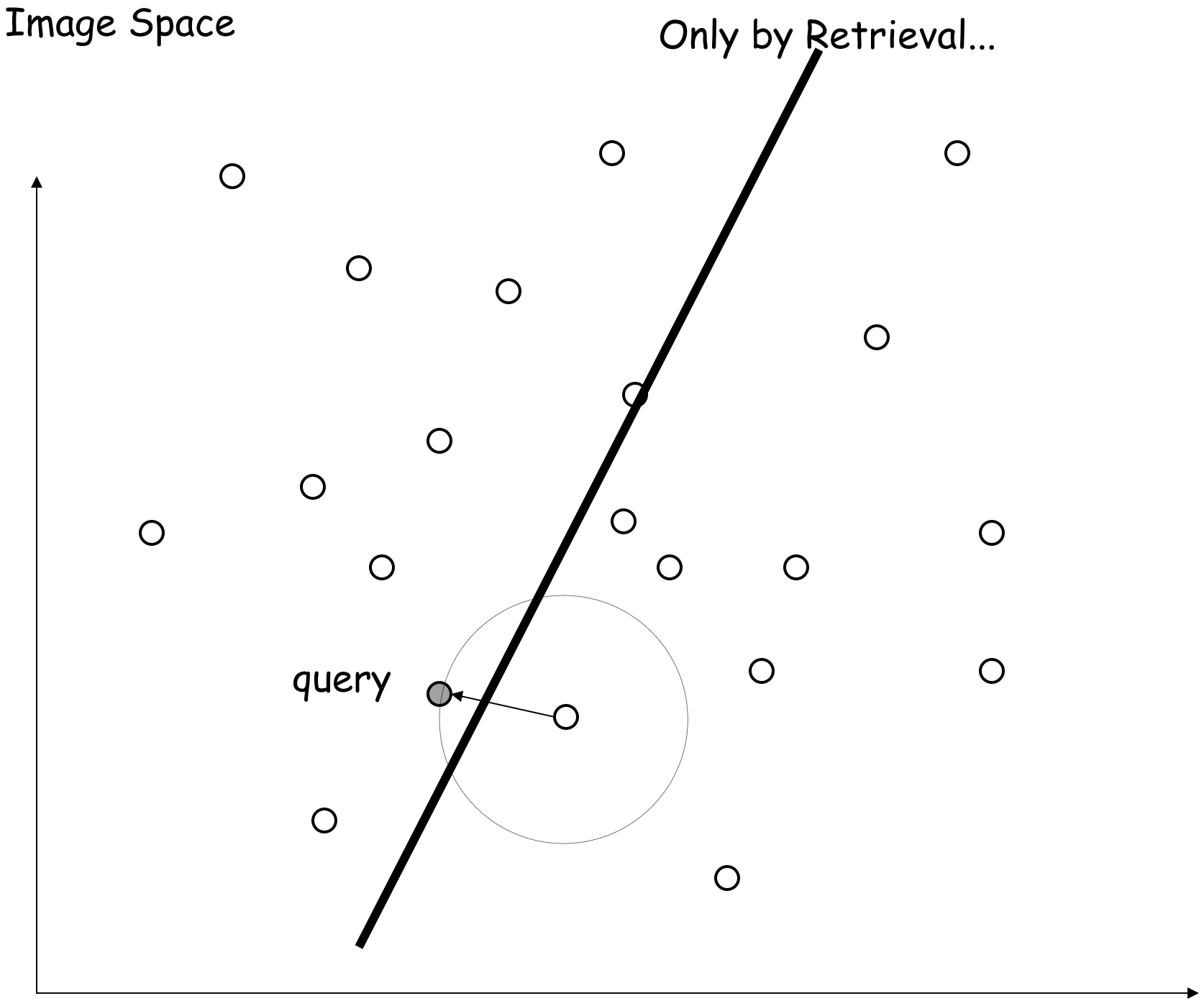
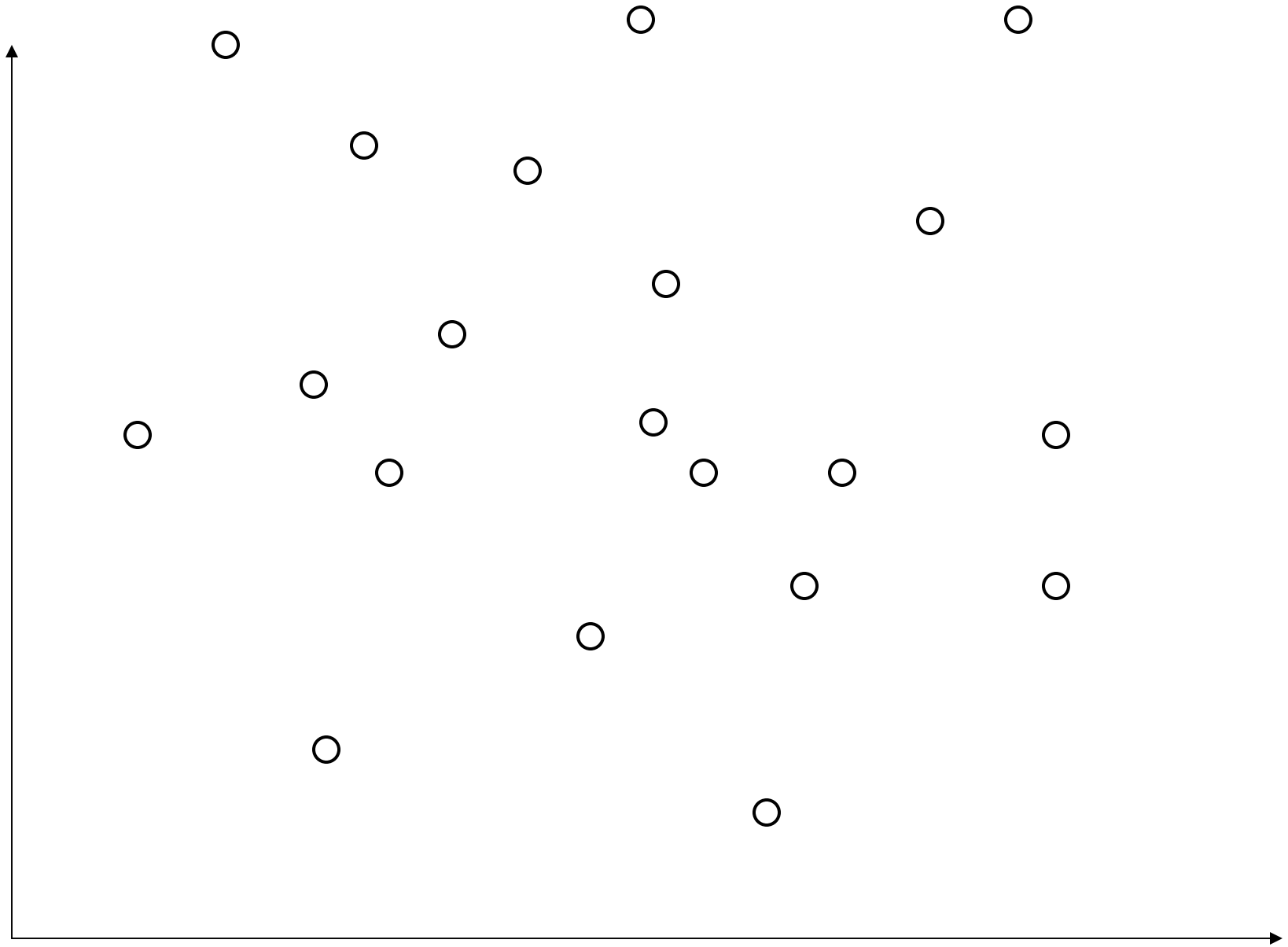


Image Space
Only by Retrieval...
query

Image Space

Proposed method



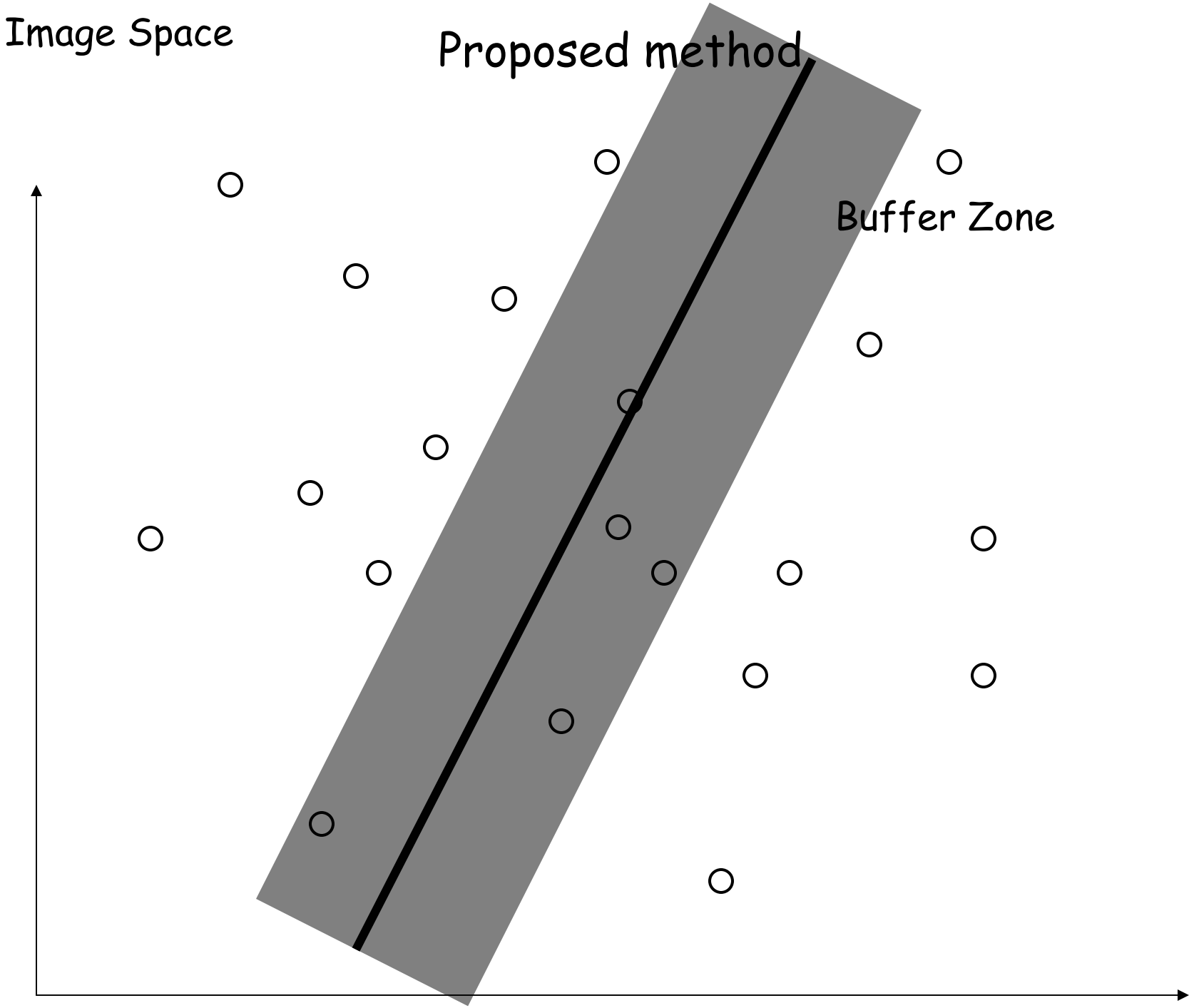


Image Space
Proposed method
Buffer Zone

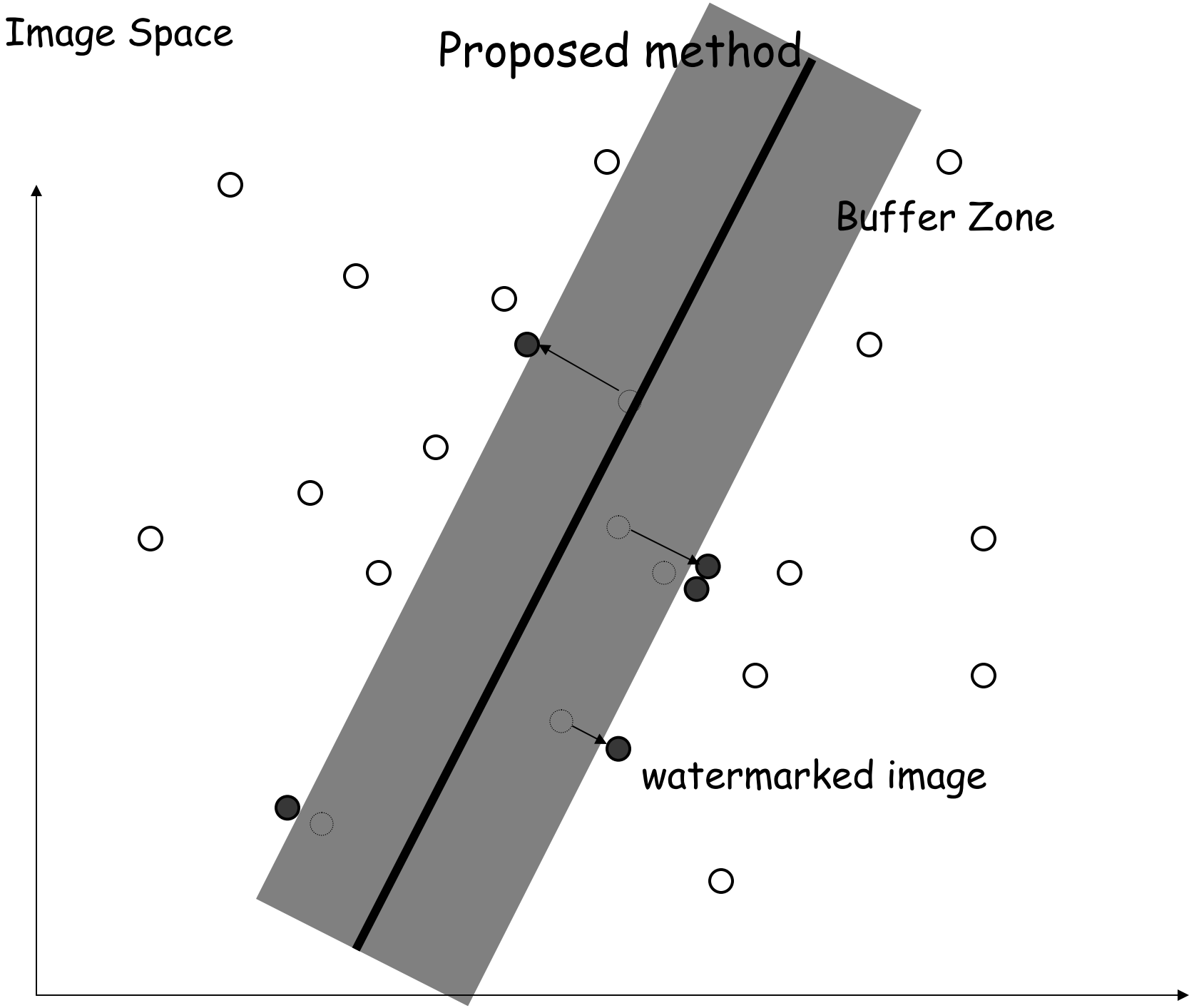
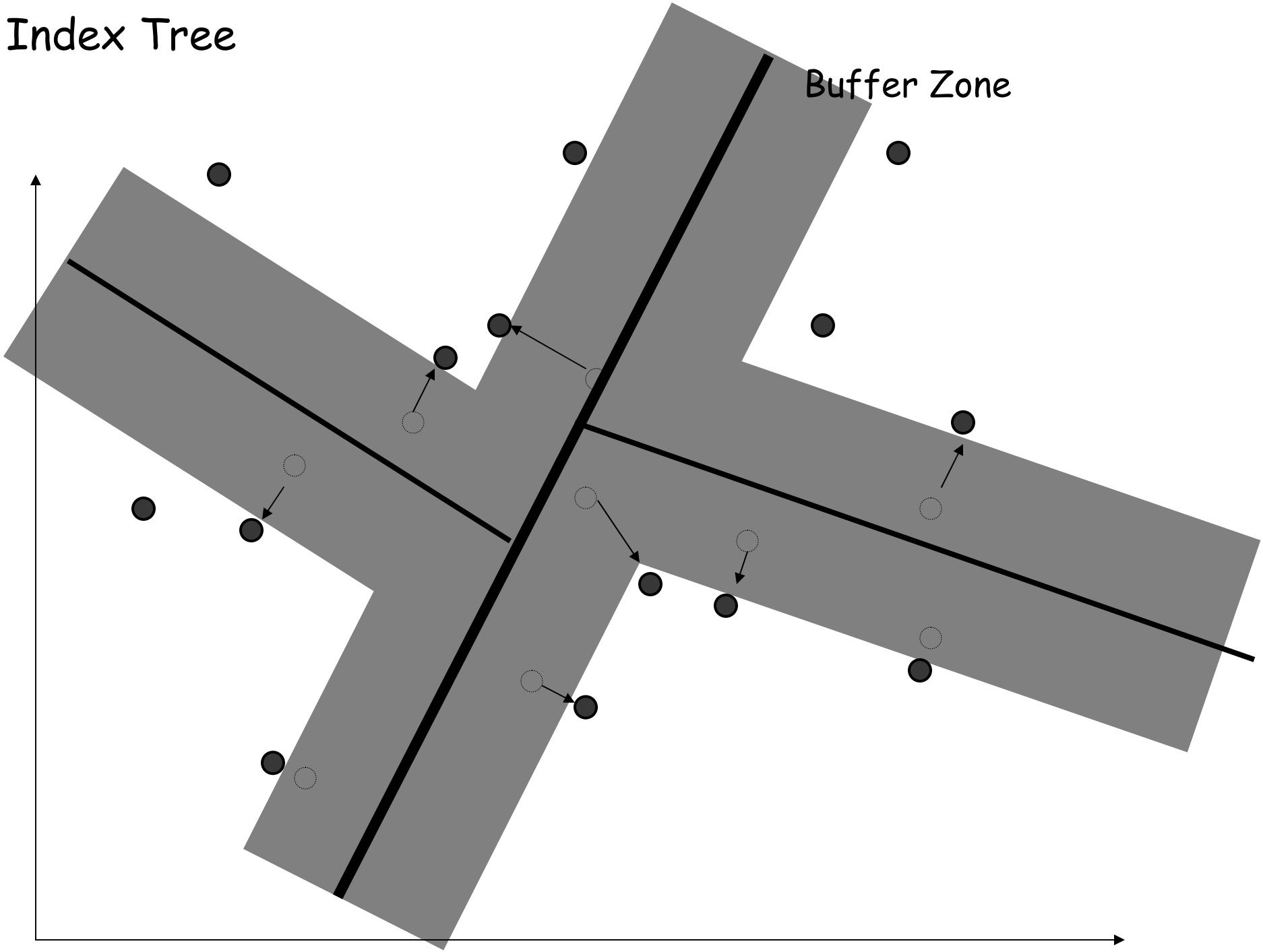
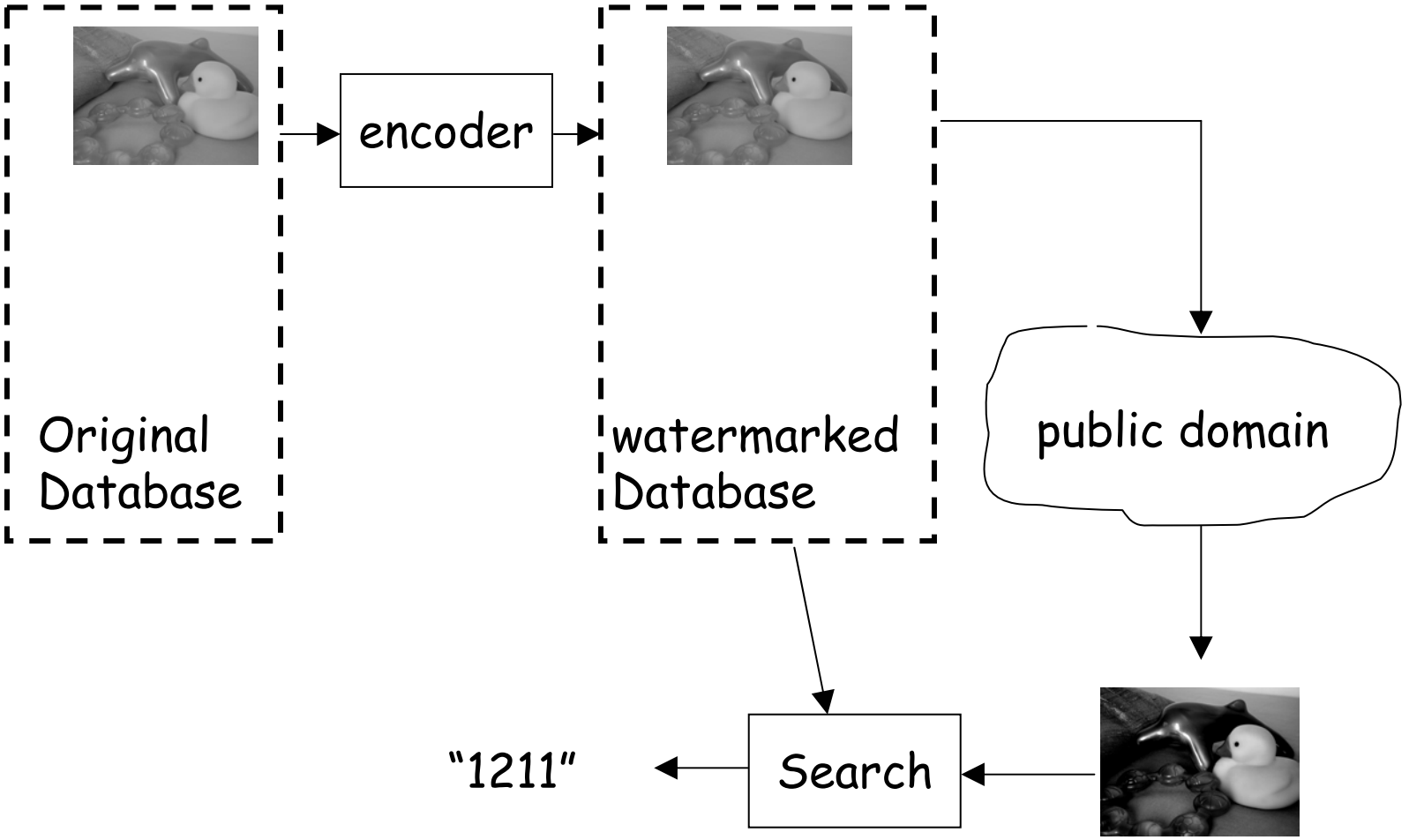


Image Space
Proposed method
Buffer Zone
watermarked image


Index Tree
Buffer Zone

Identification/tracking by **Retrieval with WM.**



Problem Formulation:

Given $\boldsymbol{I} = \langle I_1,\ldots,I_n \rangle$, a distortion constraint, e, and robustness σ^2, we want to preprocess $\boldsymbol{I}$ to obtain the watermarked $\boldsymbol{I}' = \langle I_1',\ldots,I_n' \rangle$ and an index tree.

1. The watermarked $\boldsymbol{I}'$ satisfies the *distortion* constraint e,

$$\sum_i || \ I_i' - I_i \ ||^2 < e.$$

2. The index tree supports fast search, such that given the query I_i', we can output i efficiently.

3. The searching is *robust* in the sense that if I_i' is corrupted by AWGN with power σ^2, the output is correct with high probability.

Experiment

•A database of 2048 images.
Each image is downsampled to 64×64.

•Robustness chosen to be 2
(energy of image is normalized to 1)

• Average distortion: 0.00085
Maximum distortion: 0.010

Comparison

If only watermarking is used, with average distortion of 0.00085 and robustness 2, the theoretical maximum number of images allowed is

$$(1+0.00085/2)^{64\times 64/2} < 3.$$

capacity by watermarking = $(d/2) \log (1+ e/\sigma^2)$

[1] M.Costa, Writing on dirty paper, *IEEE Trans. Info. Theory*, 29(3), 1983.

12 images (original) from the database.

Distortion

original Image

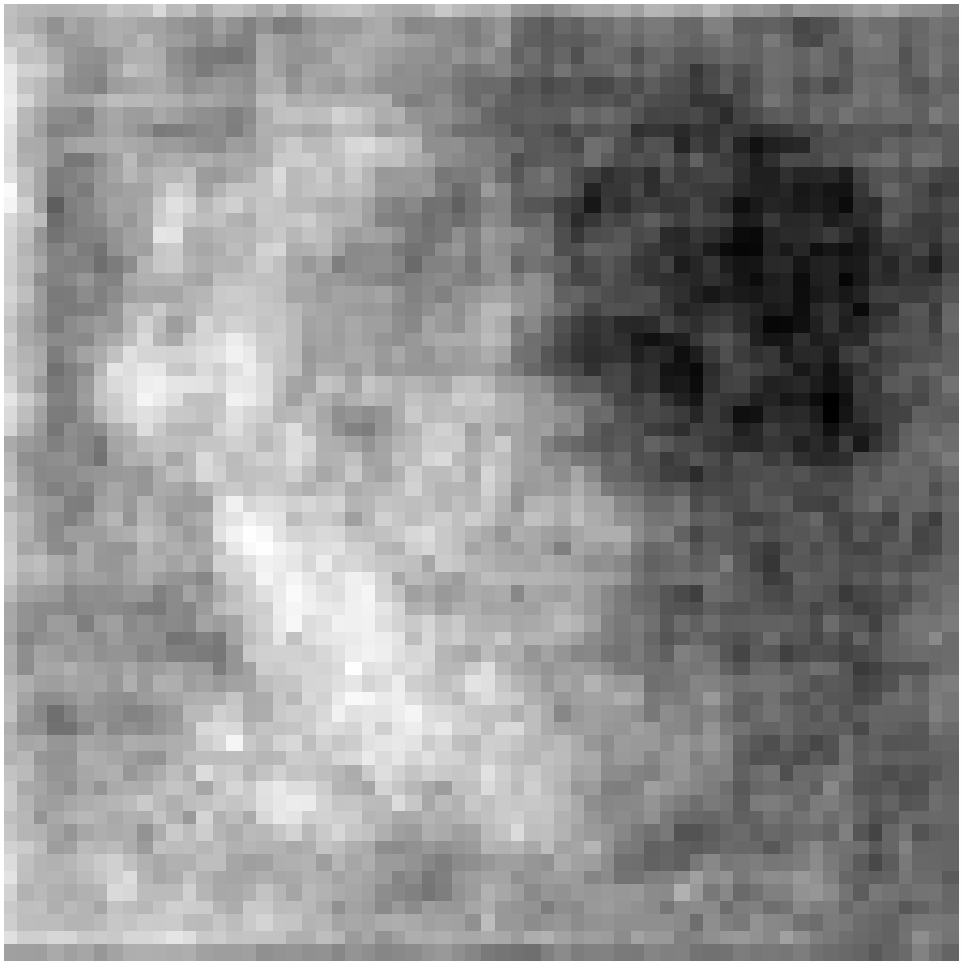
Distortion

Watermarked Image

Distortion

differences (orignal – watermarked)

distortion: 0.010



Distortion

original Image

Distortion

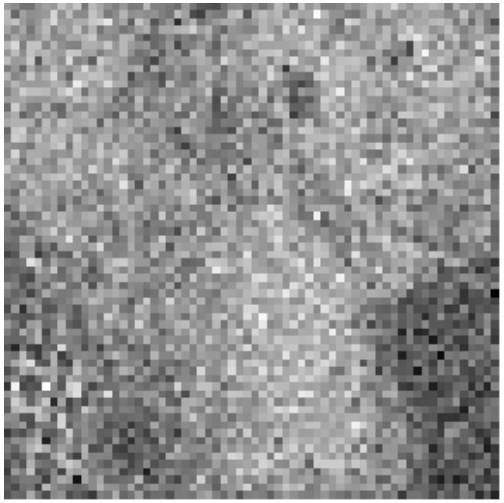
watermarked image

Distortion

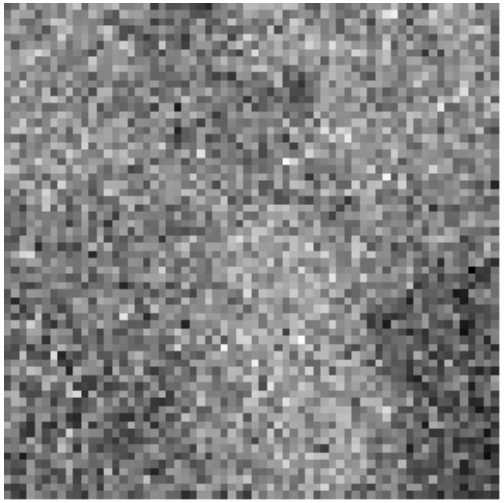
Difference

Distortion:0.0094

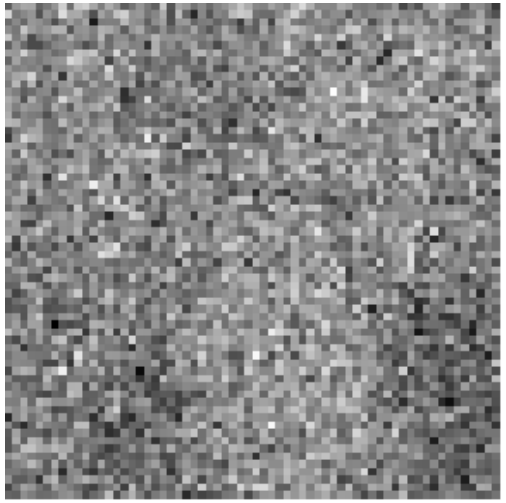
Robustness



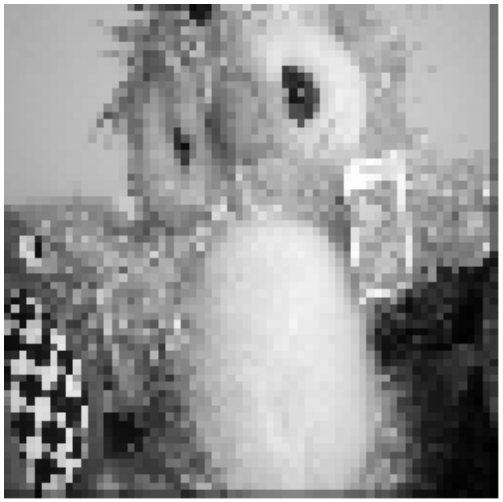
$\sigma^2 = 1$

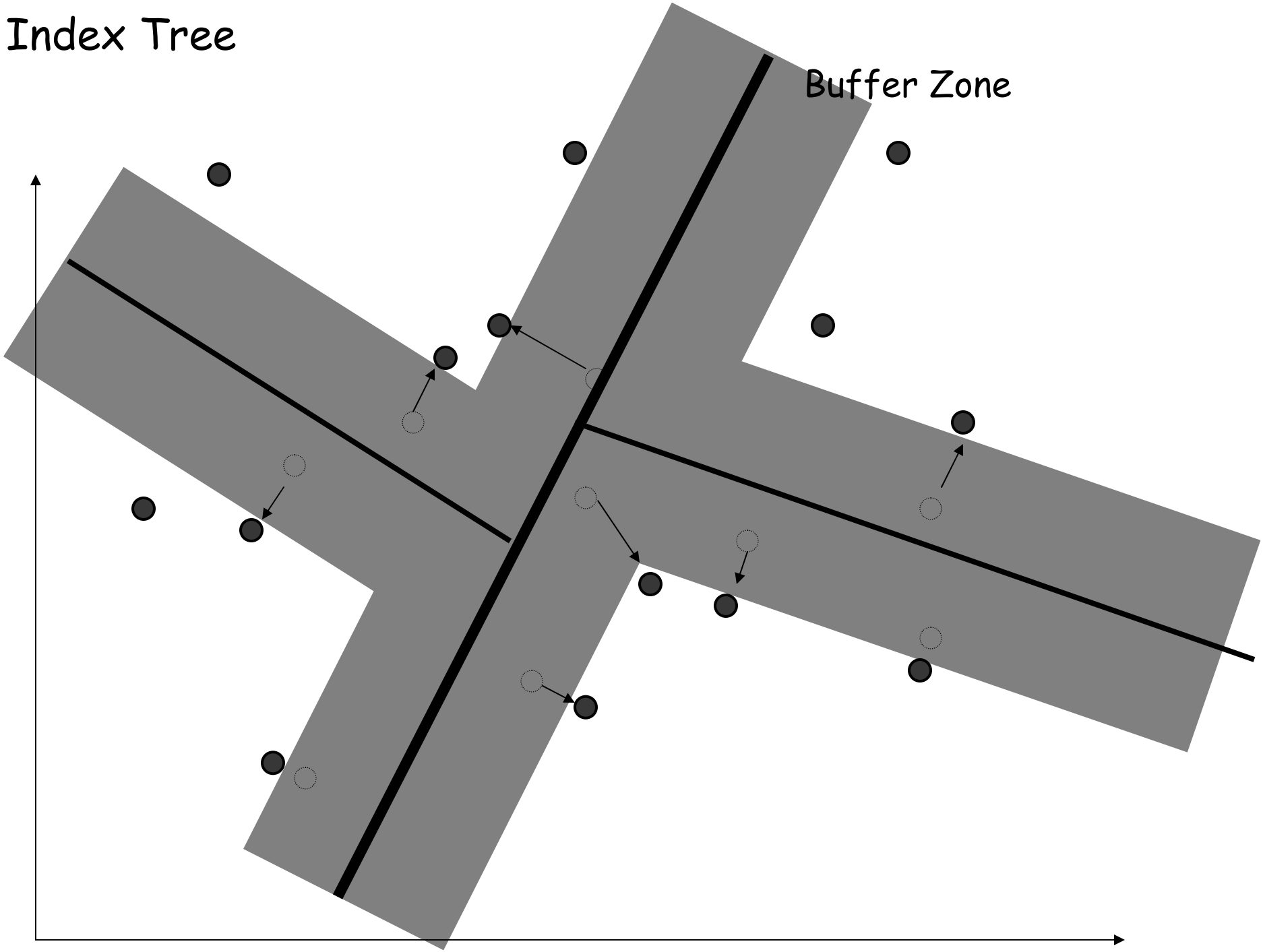


$\sigma^2 = 2$

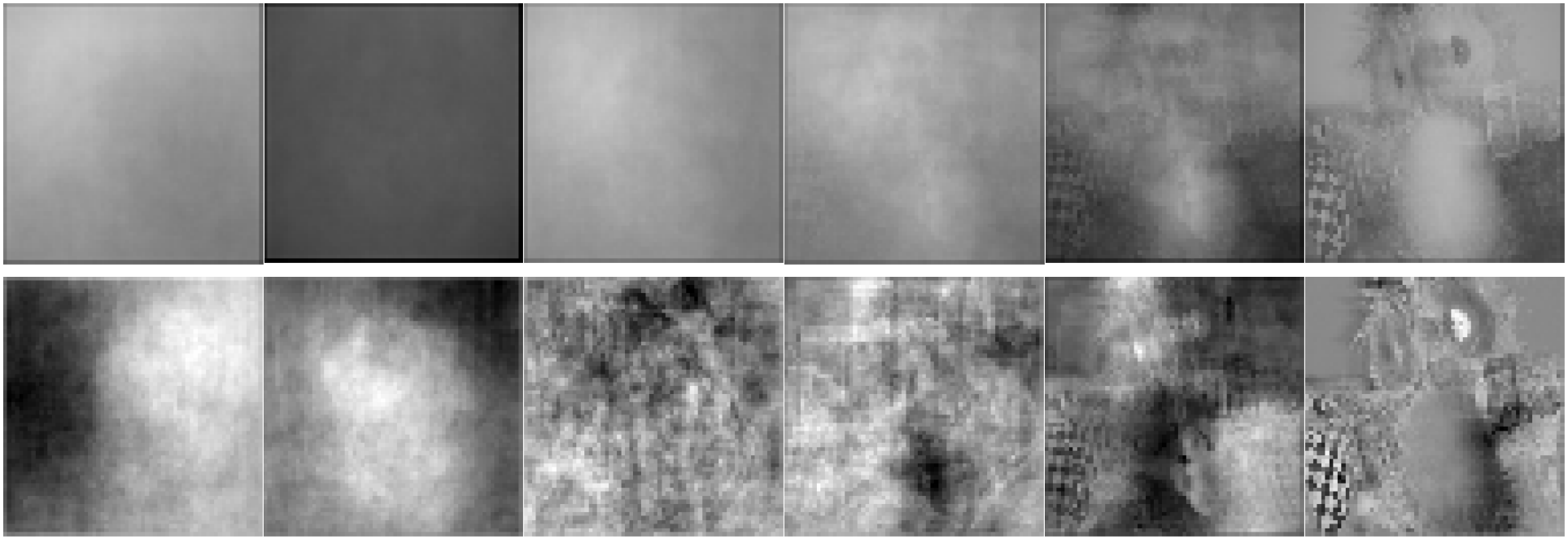


$\sigma^2 = 4$
(655 in 1000)



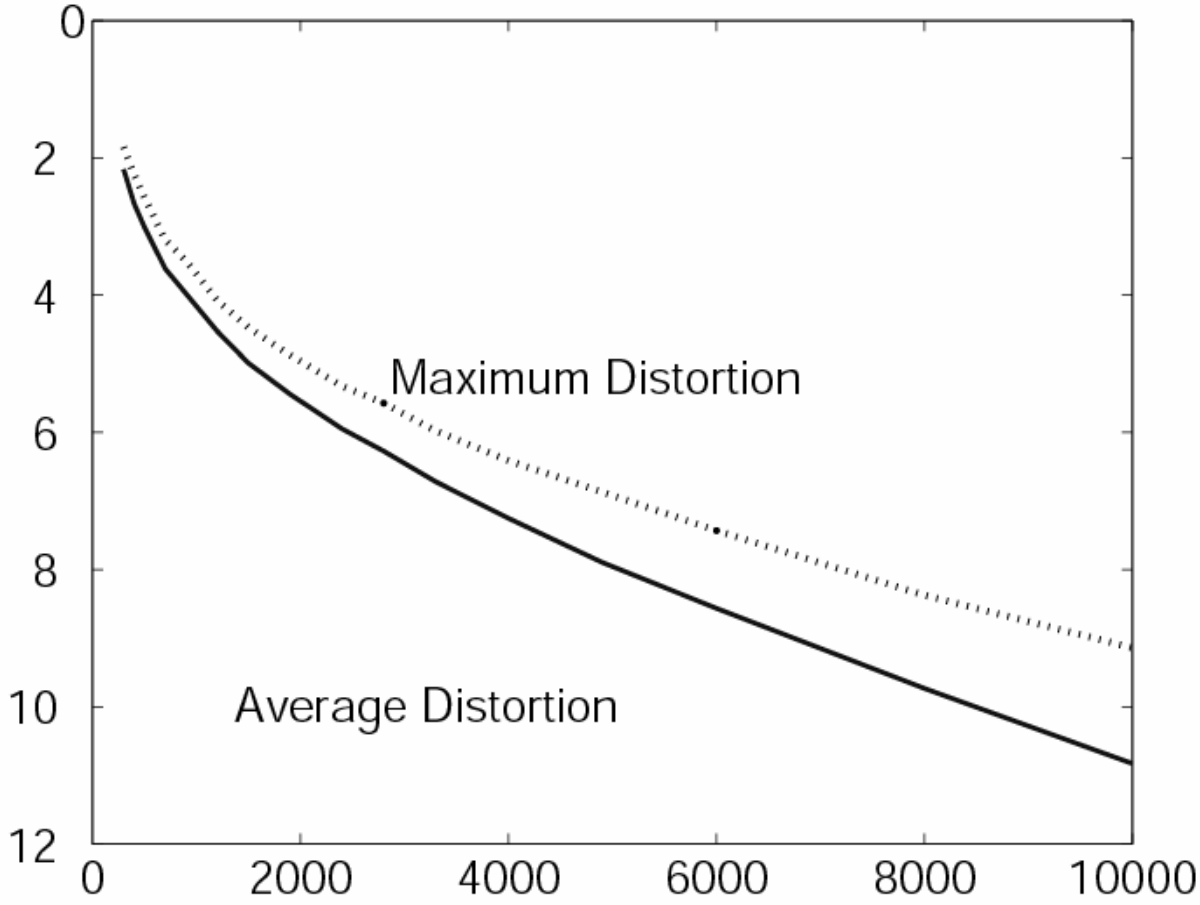

Index Tree
Buffer Zone

Index Tree



1st level

8th level


0
2
4
6
8
10
12
0
2000
4000
6000
8000
10000
Maximum Distortion
Average Distortion

Variations and Future works

- Dynamic: allow the database to grow.

- Practical implementation: Consider geometric distortion (rotation, translation,...), compression, cropping, etc.

- Improving the clustering algorithm.

- Trade-off with the size of the index tree.

Conclusion

- We introduce a variant of retrieval problem where the data-point can be distorted.
- Give an algorithm which is a combination of watermarking techniques and clustering algorithms.
- With small distortion, we can search fast.
- With knowledge of the database, and searching ability, we can improve watermarking performance.