Bag of Words

Jeon Hyun Ho

Introduction

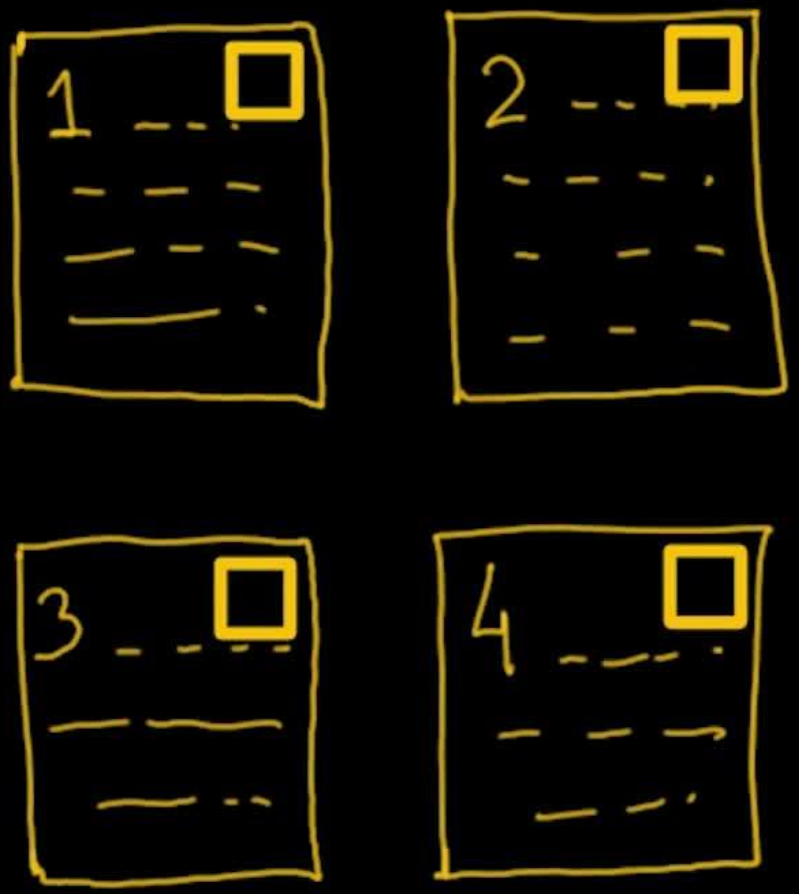
Of all the sensory impressions proceeding to the brain, the visual experiences are the dominant ones. Our perception of the world around us is based essentially on the messages that re… from our eyes. For a l… that the retinal … by point to … cerebra… speak, … was pr… Hubel … behind … the brain … complicated … the visual imp… the various cell layers … Hubel and Wiesel have been able … demonstrate that the *message abo… image falling on the retina undergoe… step-wise analysis in a system of nerv… stored in columns. In this system each … has its specific function and is responsibl… for a specific detail in the pattern of the retinal image.*

sensory, brain, visual, perception, retinal, cerebral cortex, eye, cell, optical nerve, image Hubel, Wiesel

China is forecasting a trade surplus of $90bn (£51bn) to $100bn this year, a threefold increase on 2004's $32bn. The Commerce Ministry said the surplus would be created by a predicted 30% … ports to $750bn, com… imports to … to further a… ued that Ch… a deliber… agrees … yuan is … governo… also nee… demand so… country. Chin… yuan against the … and permitted it to trade within a narr… nd, but the US wants the yuan to be a… to trade freely. However, Beijing has ma… clear that it will take its time and tread carefully before allowing the yuan to ris… further in value.

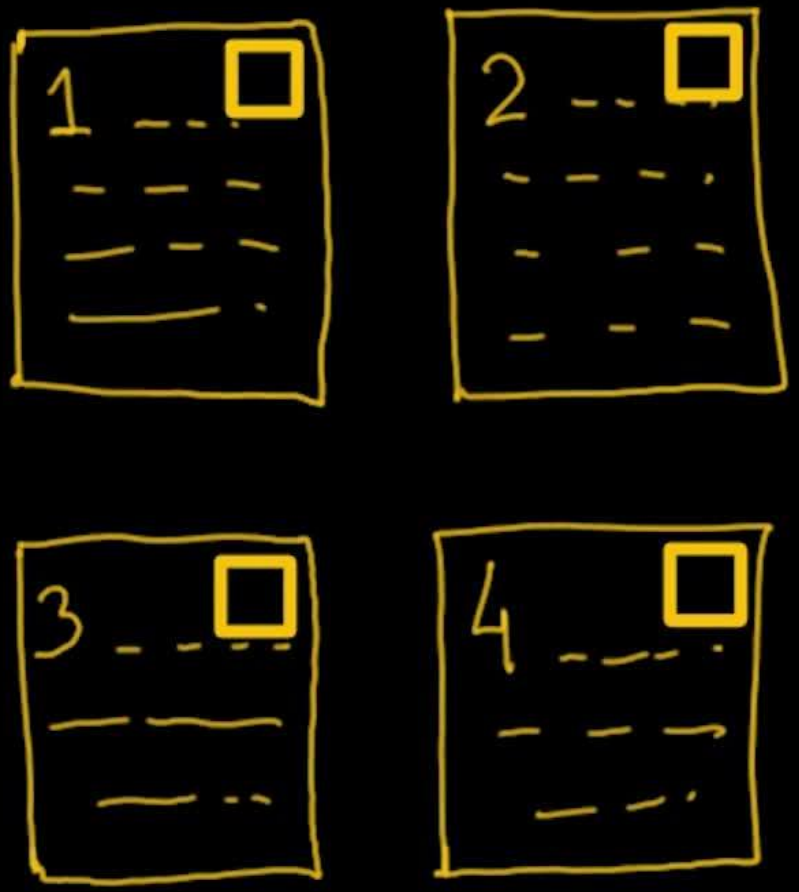
China, trade, surplus, commerce, exports, imports, US, yuan, bank, domestic, foreign, increase, trade, value

Find Likely Page



ant 1,4	gong 2,4
bread 2	hammer 2,3
bus 4	leaf 1,2,3
chair 2,4	map 2,3
chisel 3,4	net 2
desk 1	pepper 1,2
fly 1,3	shoe 1,3

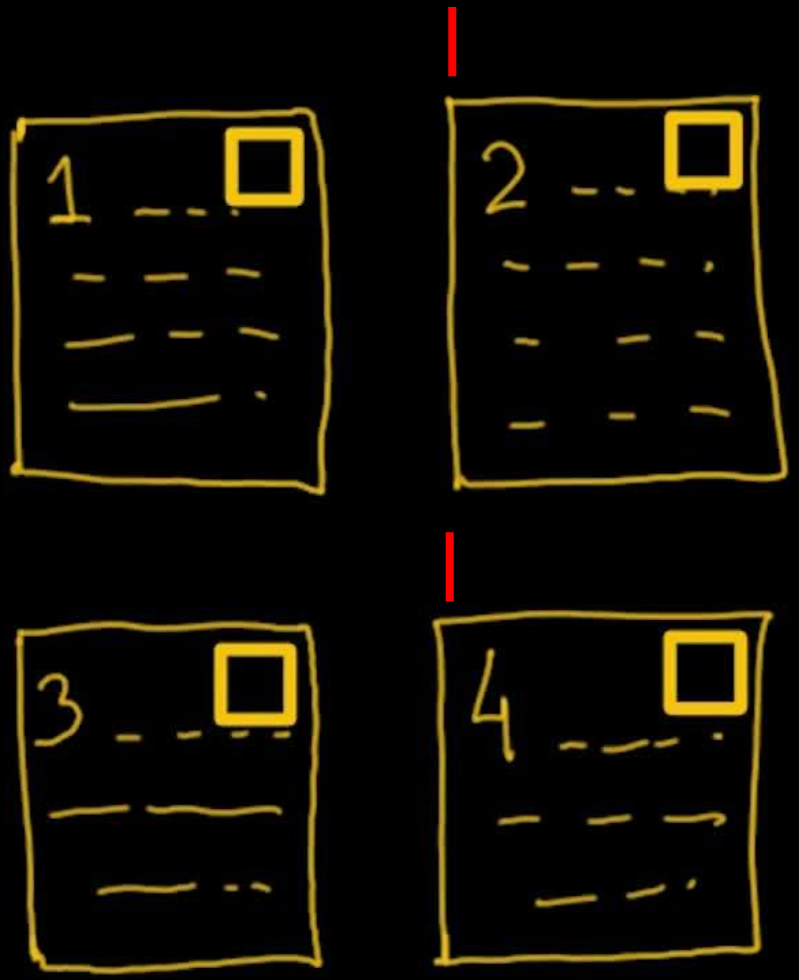
Find Likely Page



ant 1,4	gong 2,4
bread 2	hammer 2,3
bus 4	leaf 1,2,3
chair 2,4	map 2,3
chisel 3,4	net 2
desk 1	pepper 1,2
fly 1,3	shoe 1,3

Test: gong fly pepper ant shoe

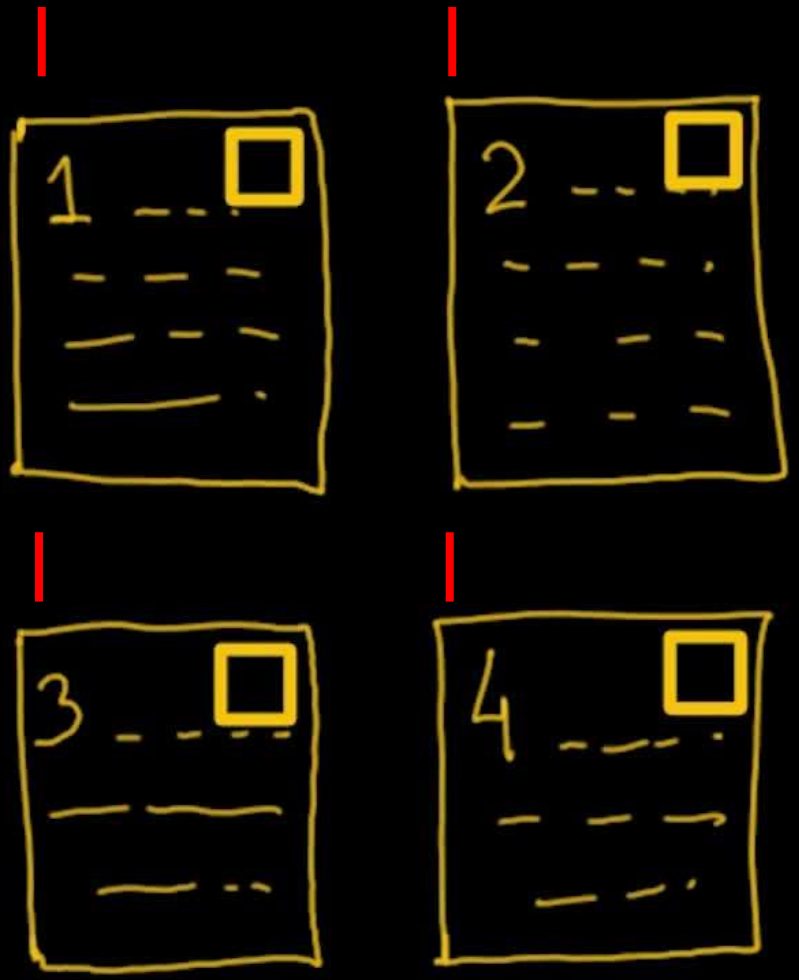
Find Likely Page



ant 1,4	gong 2,4
bread 2	hammer 2,3
bus 4	leaf 1,2,3
chair 2,4	map 2,3
chisel 3,4	net 2
desk 1	pepper 1,2
fly 1,3	shoe 1,3

Test: gong fly pepper ant shoe

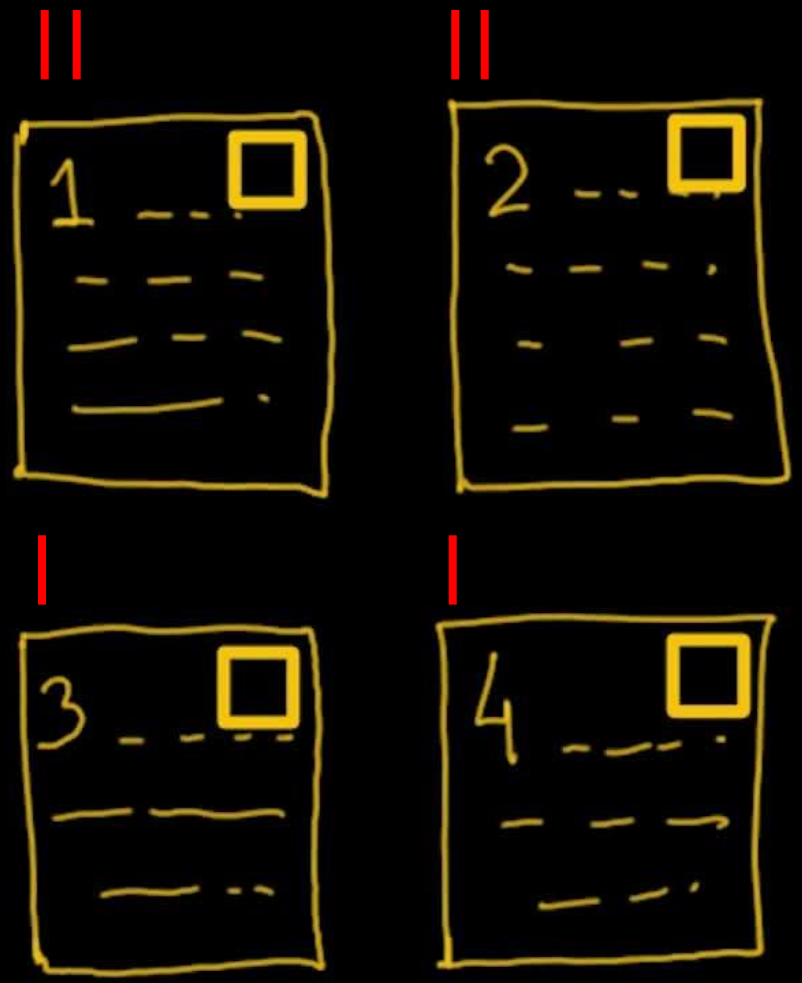
Find Likely Page



ant 1,4	gong 2,4
bread 2	hammer 2,3
bus 4	leaf 1,2,3
chair 2,4	map 2,3
chisel 3,4	net 2
desk 1	pepper 1,2
fly 1,3	shoe 1,3

Test: gong fly pepper ant shoe

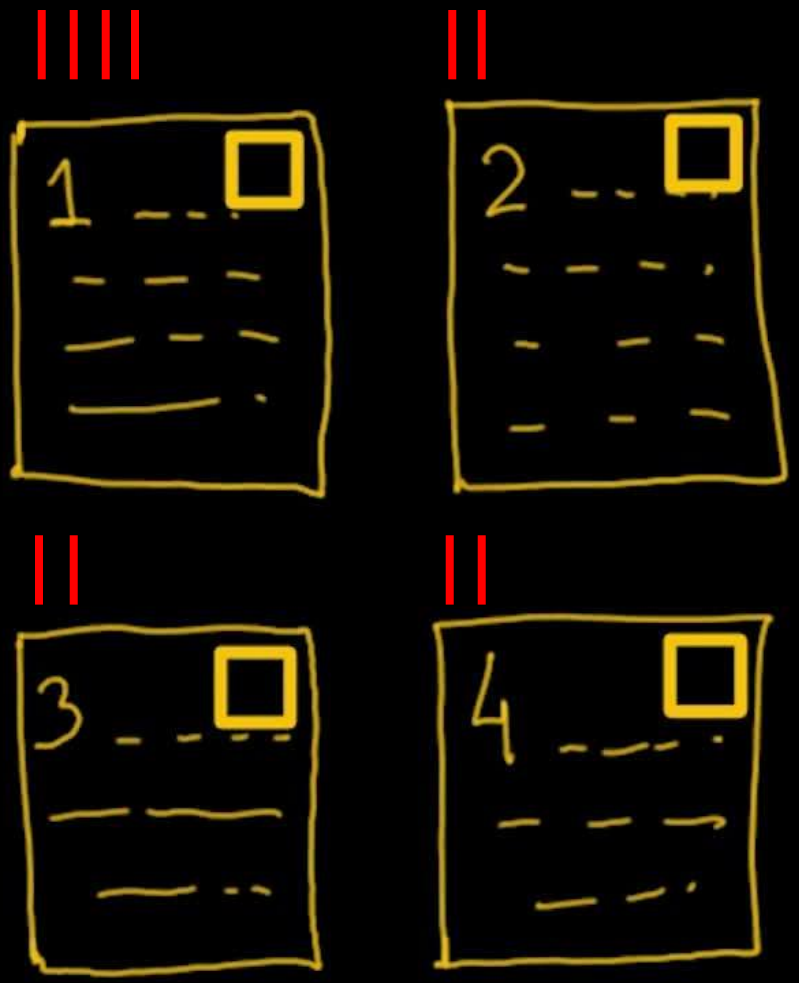
Find Likely Page



ant 1,4	gong 2,4
bread 2	hammer 2,3
bus 4	leaf 1,2,3
chair 2,4	map 2,3
chisel 3,4	net 2
desk 1	pepper 1,2
fly 1,3	shoe 1,3

Test: gong fly pepper ant shoe

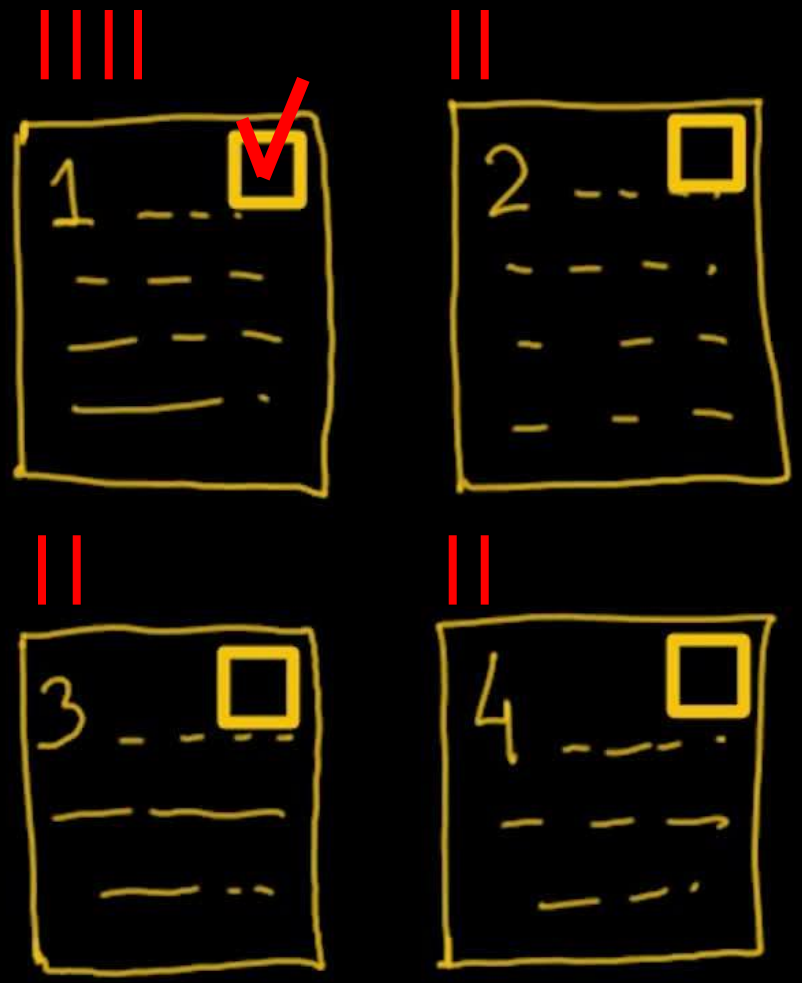
Find Likely Page



ant 1,4	gong 2,4
bread 2	hammer 2,3
bus 4	leaf 1,2,3
chair 2,4	map 2,3
chisel 3,4	net 2
desk 1	pepper 1,2
fly 1,3	shoe 1,3

Test: gong fly pepper ant shoe

Find Likely Page

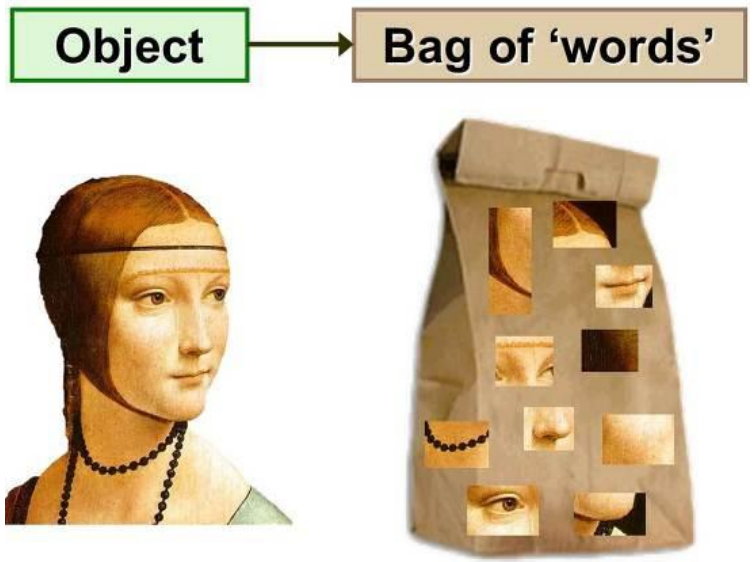


ant 1,4	gong 2,4
bread 2	hammer 2,3
bus 4	leaf 1,2,3
chair 2,4	map 2,3
chisel 3,4	net 2
desk 1	pepper 1,2
fly 1,3	shoe 1,3

Test: gong fly pepper ant shoe

Bag of Words

- **Bag of Words 기법**
  - BoW는 원래 문서를 자동으로 분류하기 위한 방법.
  - CV에서는 Image classification에 사용하며 scene의 인식을 위해도 사용함.



Bag of Words

- **Bag of Words 동작 순서**

  1) Feature Extraction (e.g., SIFT)

  2) Clustering (e.g., K-means clustering)

  3) Codebook Generation

  4) Image Representation

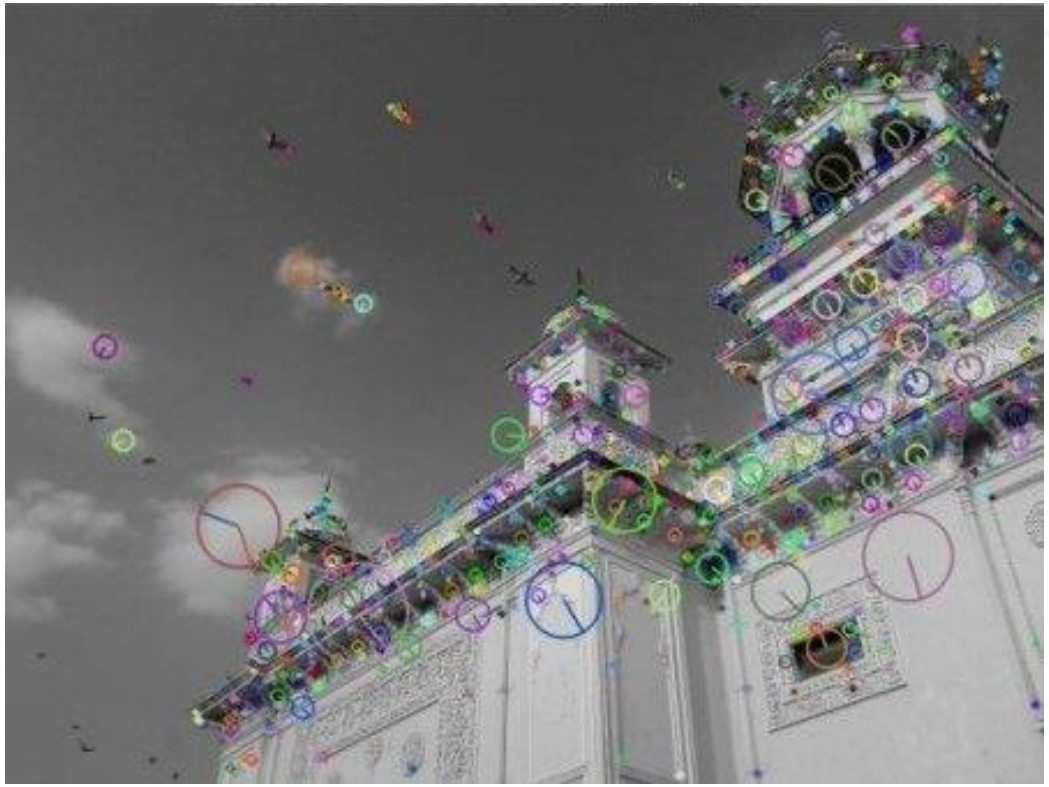
  5) Learning and Recognition

Bag of Words

- **Feature Extraction - SIFT**

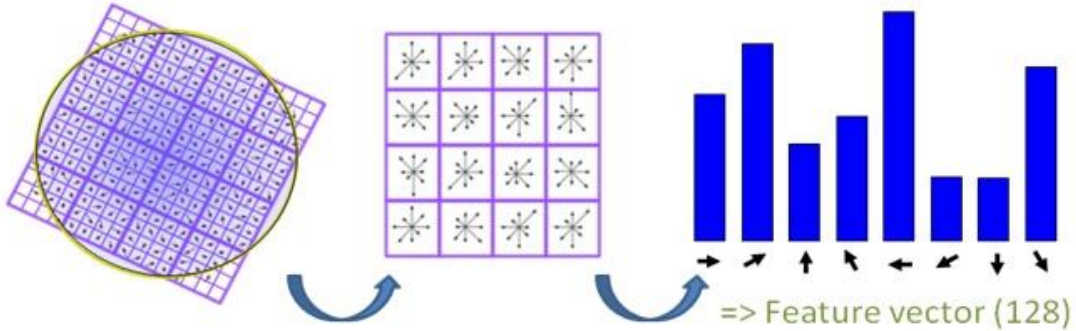
  1) Feature or Patch detection

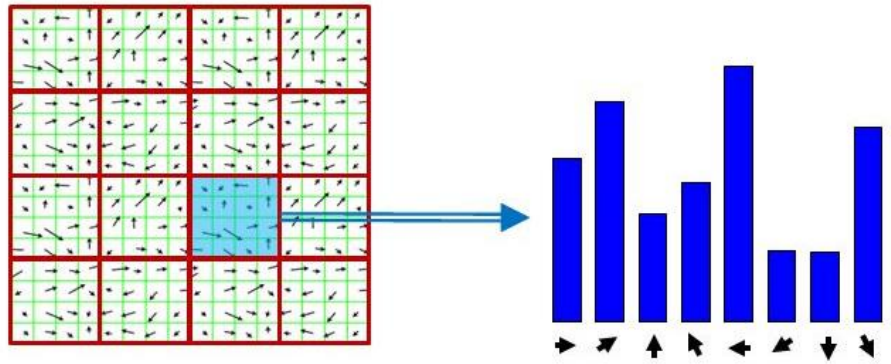
  2) Create descriptor [128x1]



Bag of Words

- Feature Extraction - SIFT

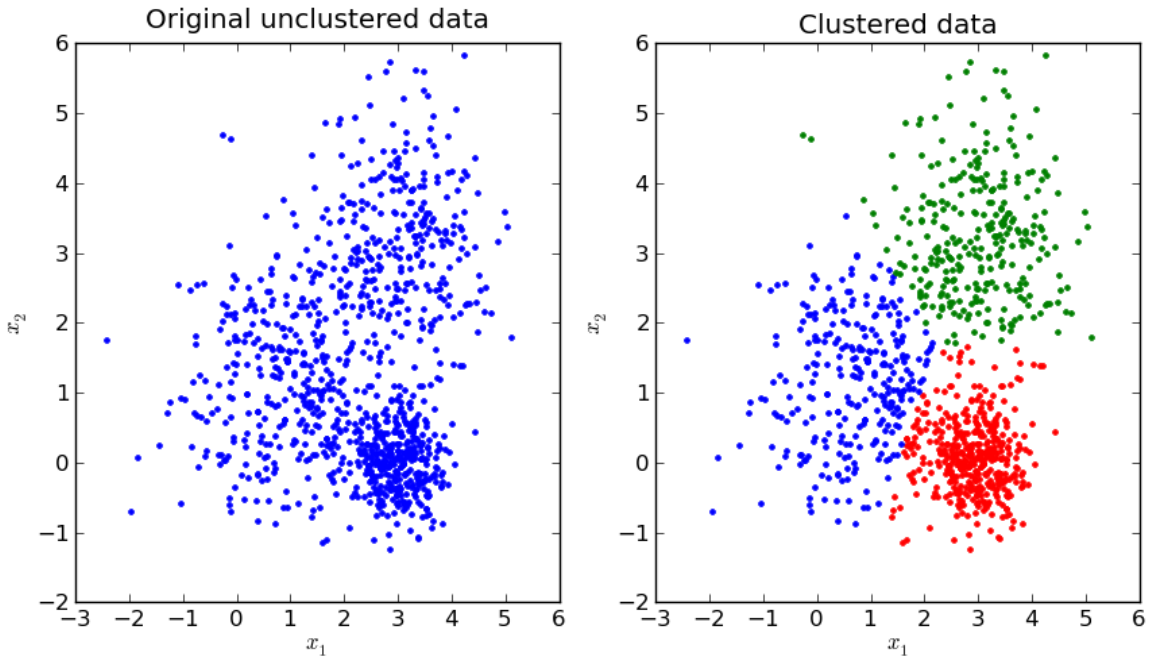




The result: 128 dimensions feature vector.

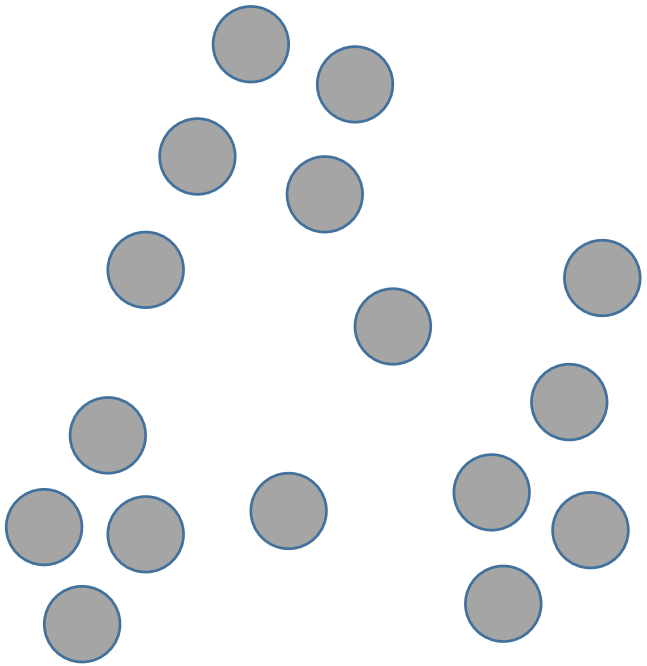
Bag of Words

- Clustering



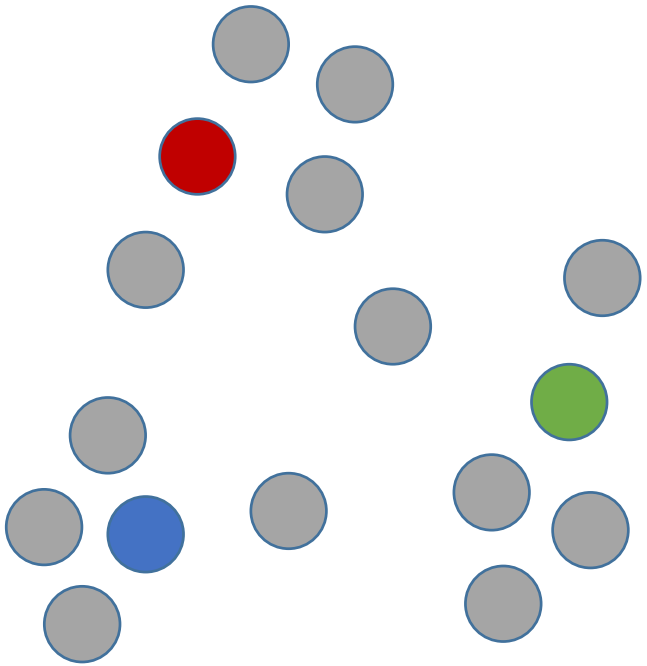
Bag of Words

- Process



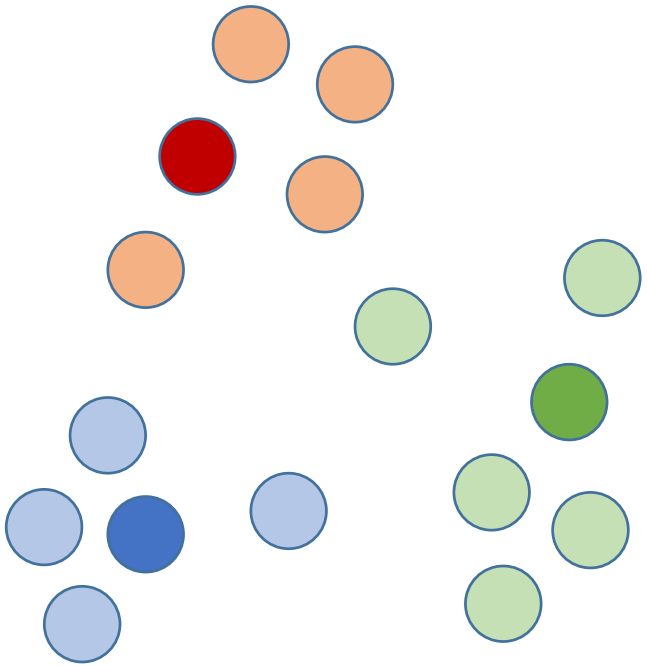
Bag of Words

- Process



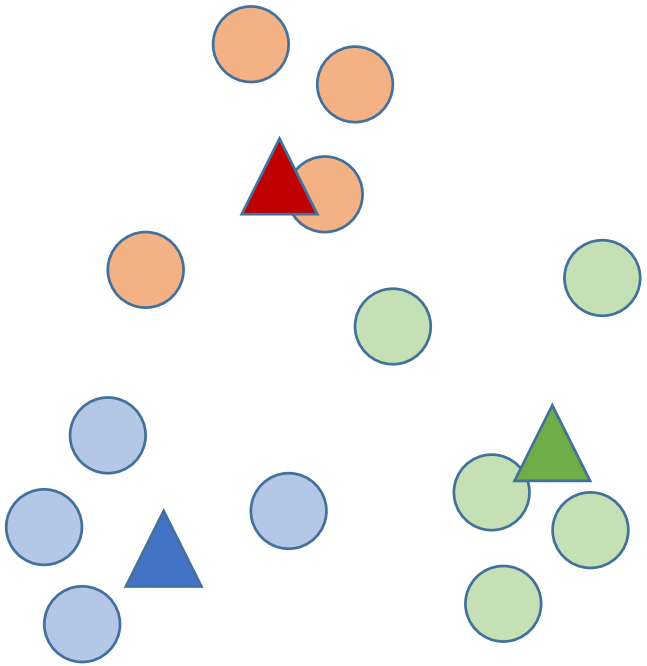
Bag of Words

- Process



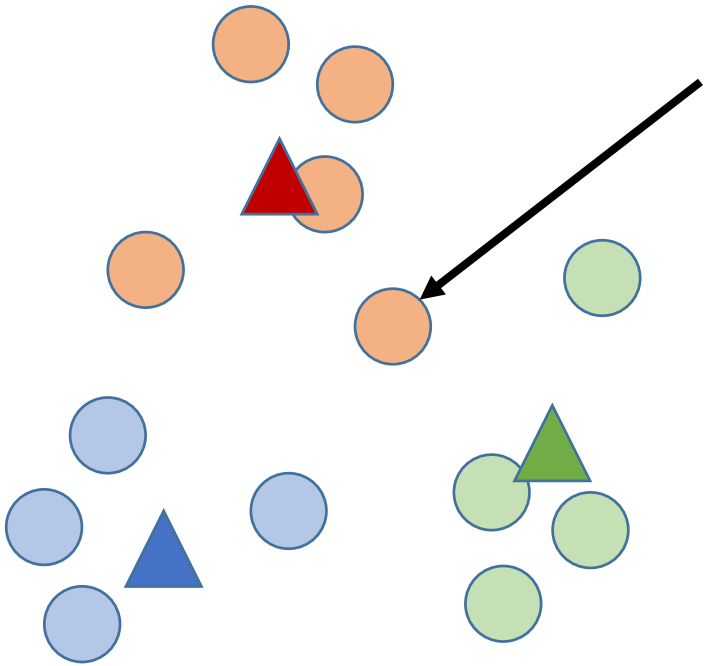
Bag of Words

- Process



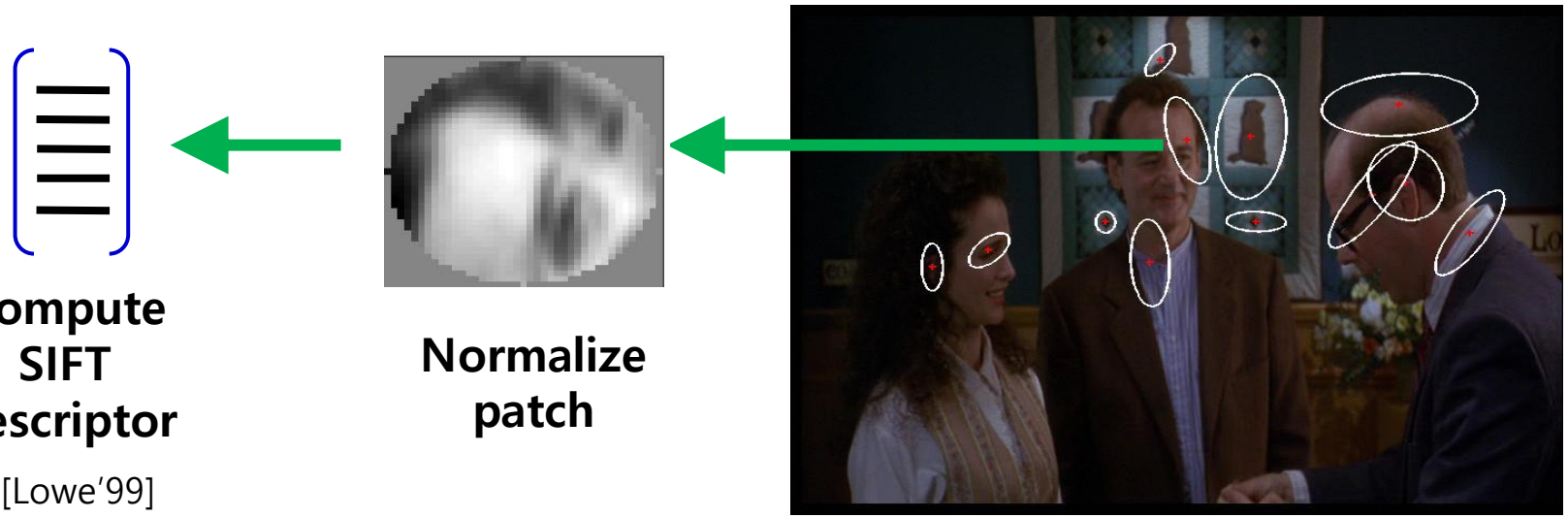
Bag of Words

- Process



Bag of Words

- K-means Clustering



Compute SIFT descriptor

[Lowe'99]

Normalize patch

Detect patches

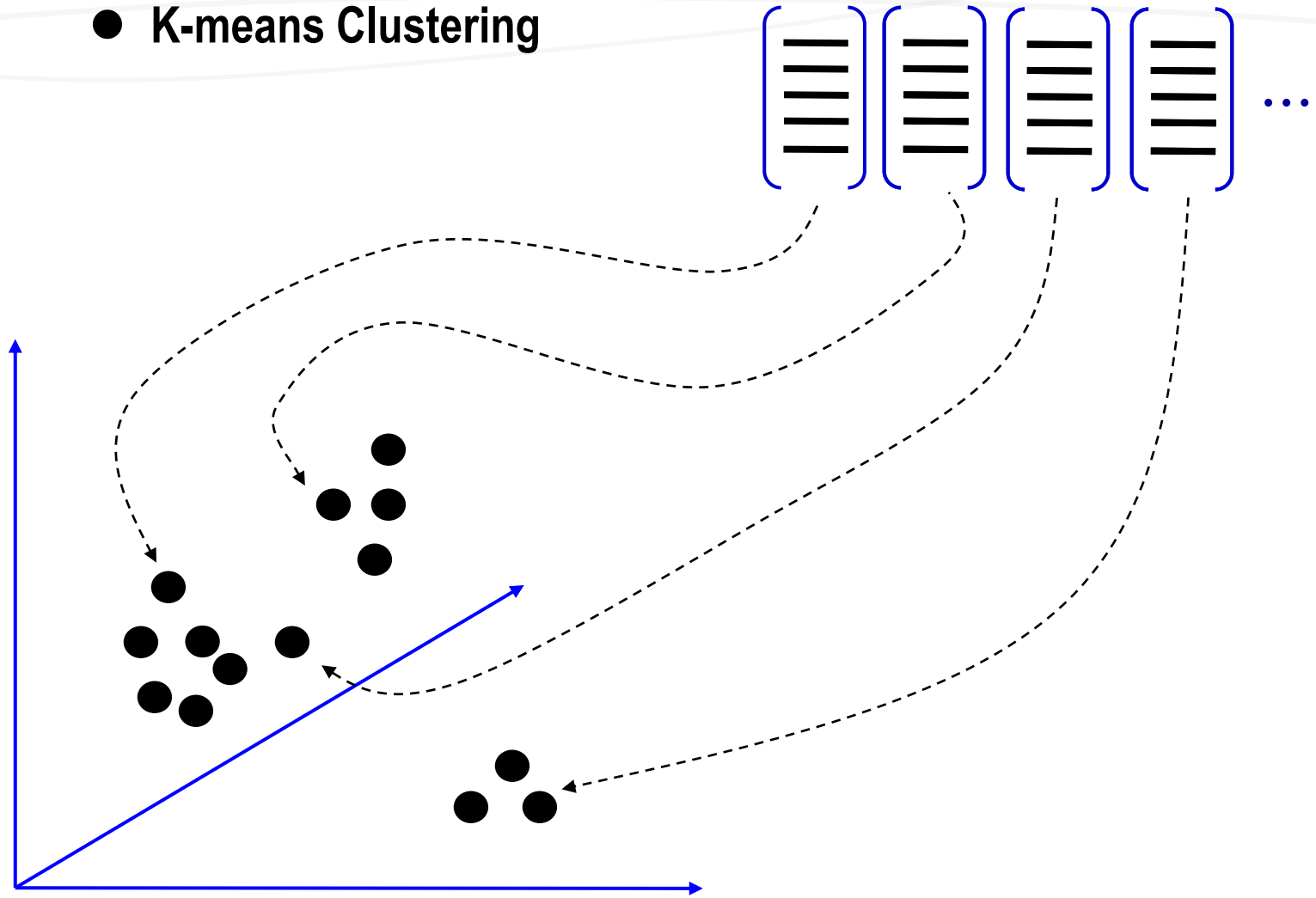
[Mikojaczyk and Schmid '02]

[Matas et al. '02]

[Sivic et al. '03]

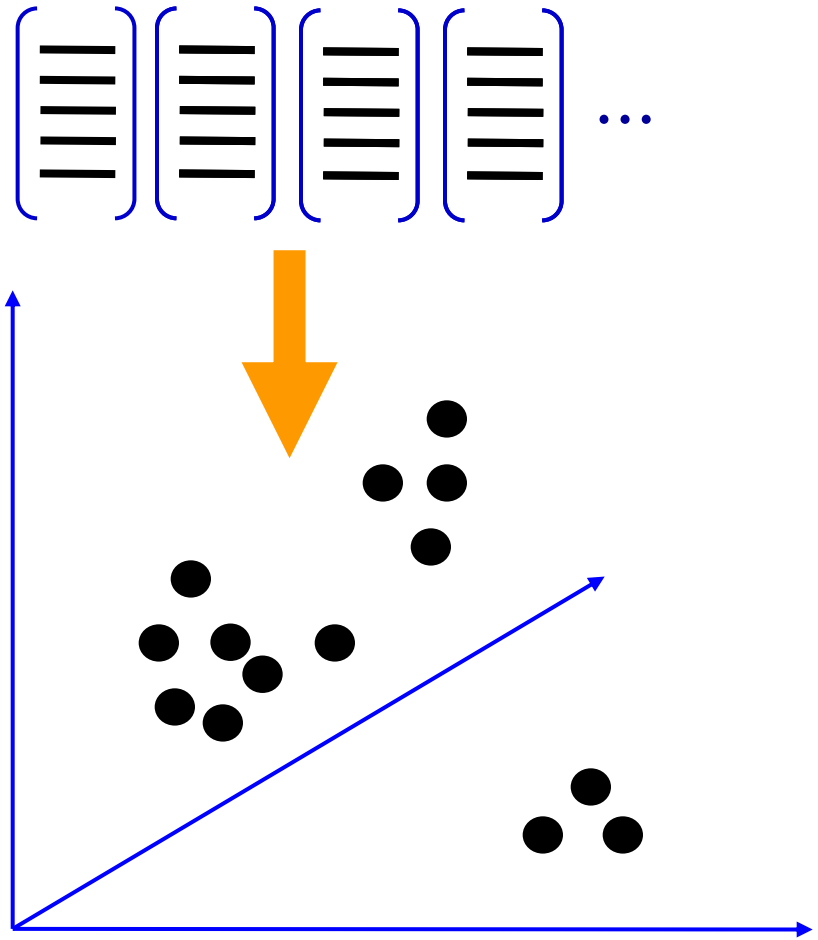
Bag of Words

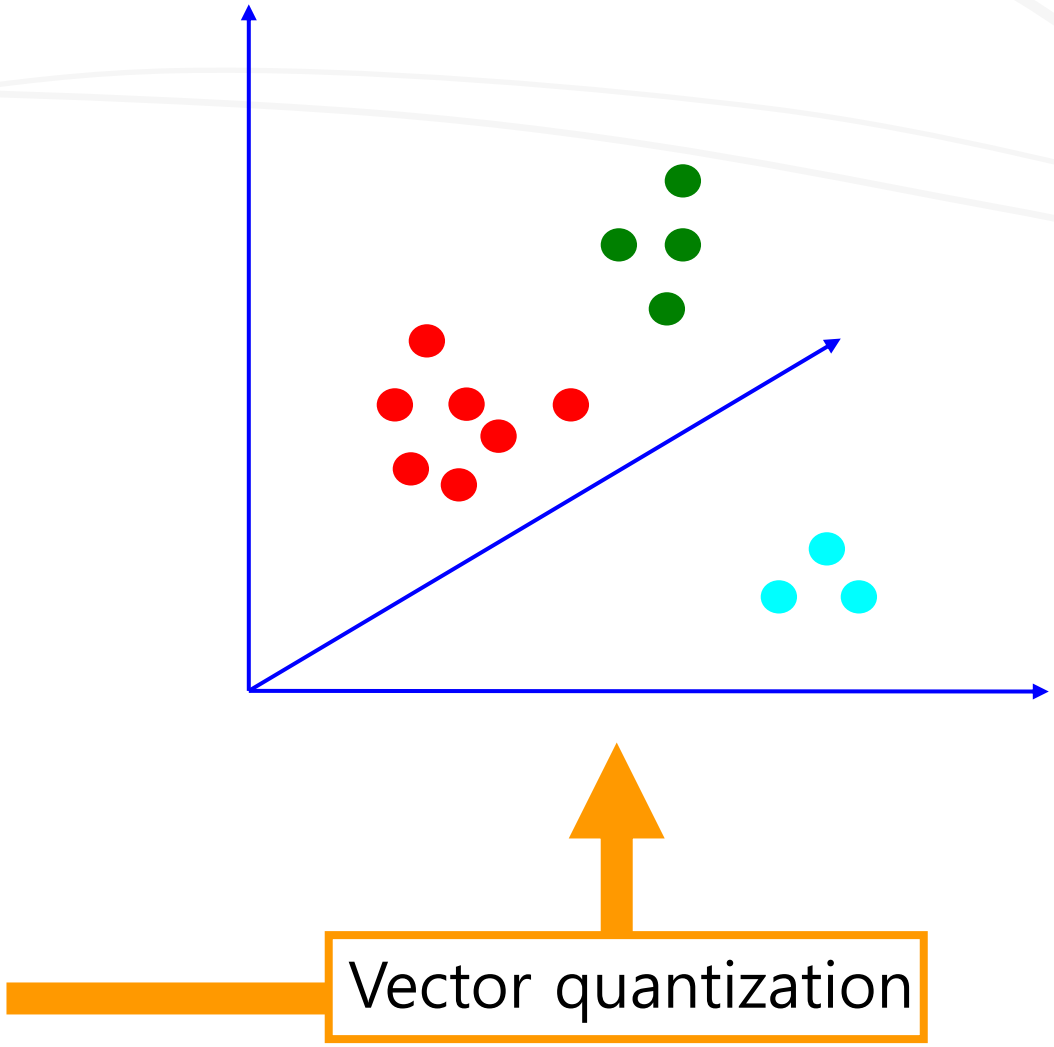
- K-means Clustering



Bag of Words

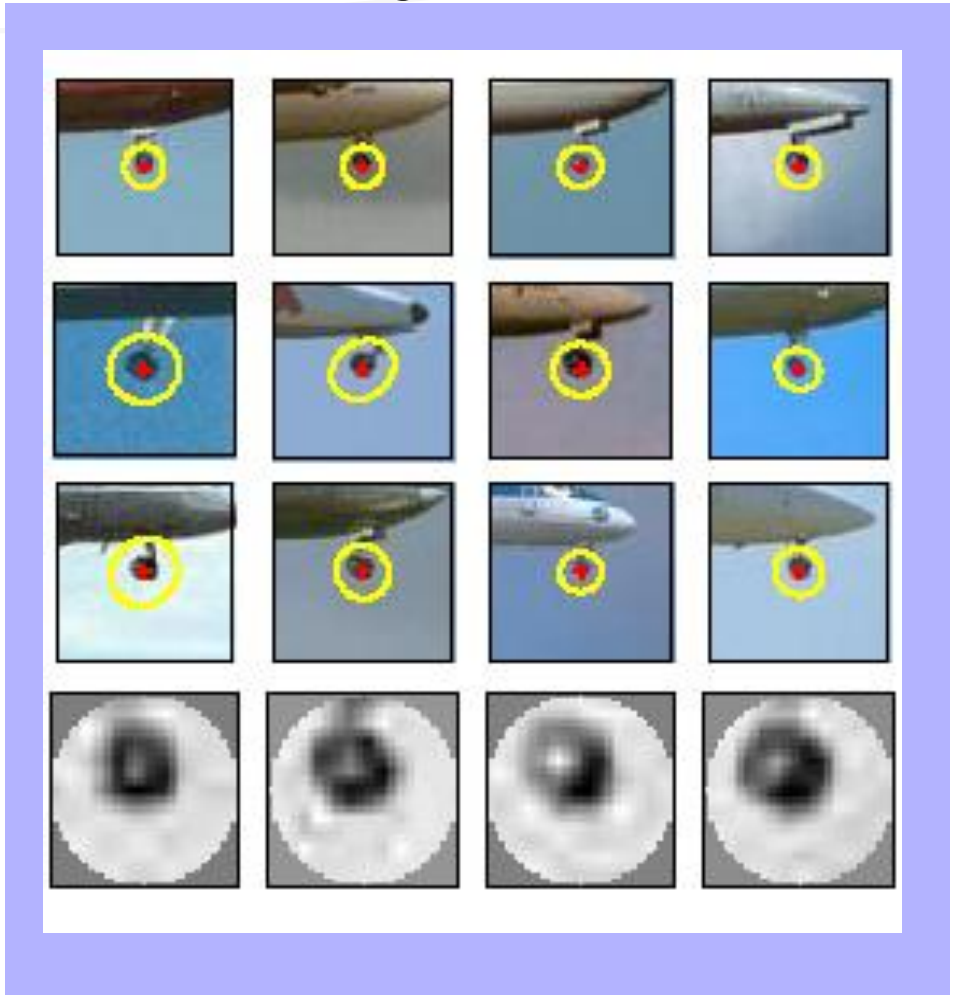
- K-means Clustering





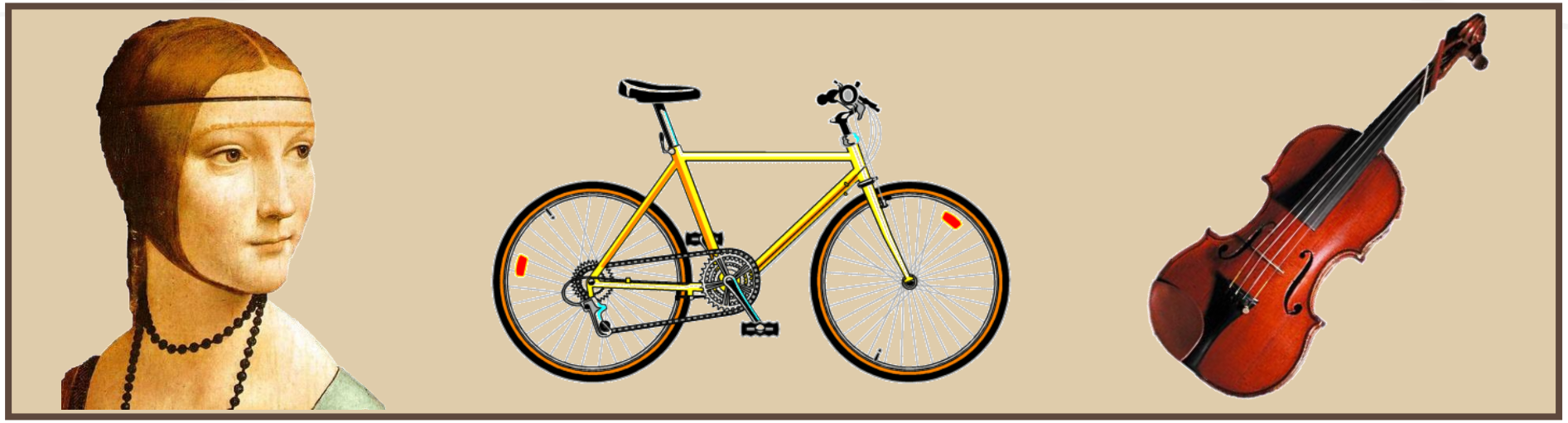
Bag of Words

- K-means Clustering



Bag of Words

- Codebook Generation



Bag

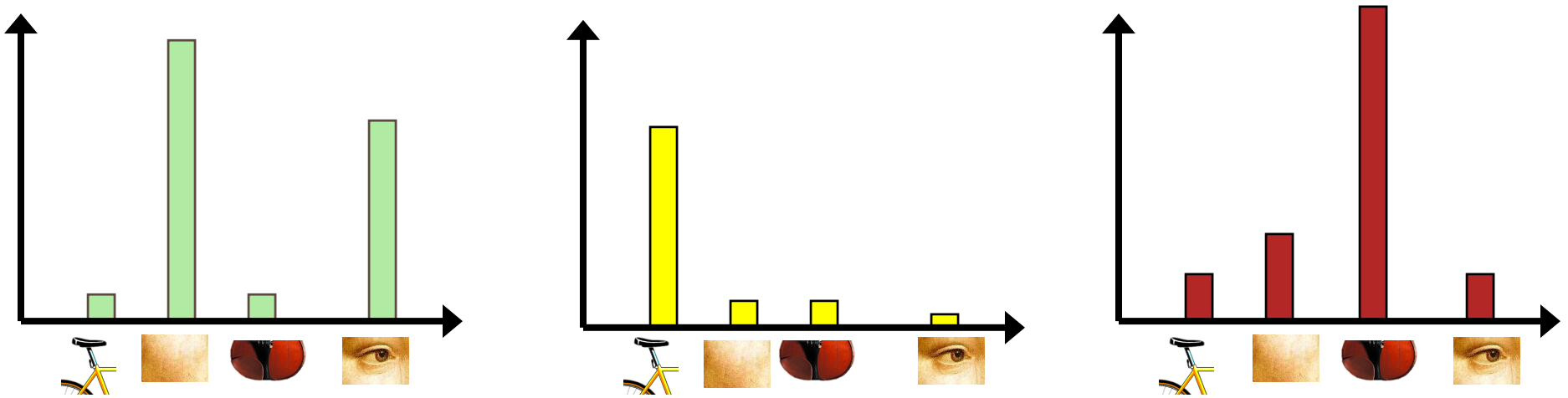
Bag of Words

- Codebook Generation

Codebook

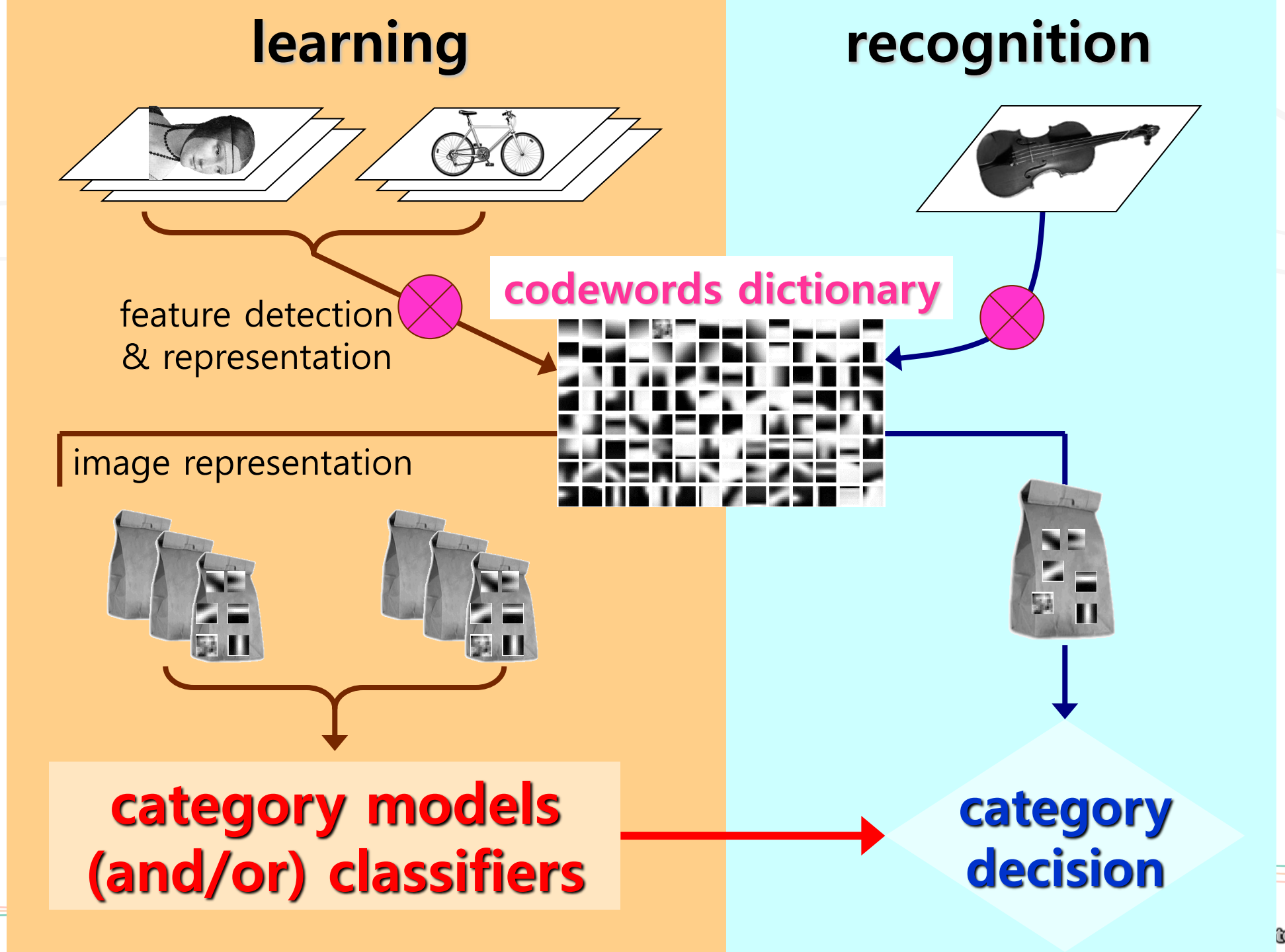
Bag of Words

- Image Representation



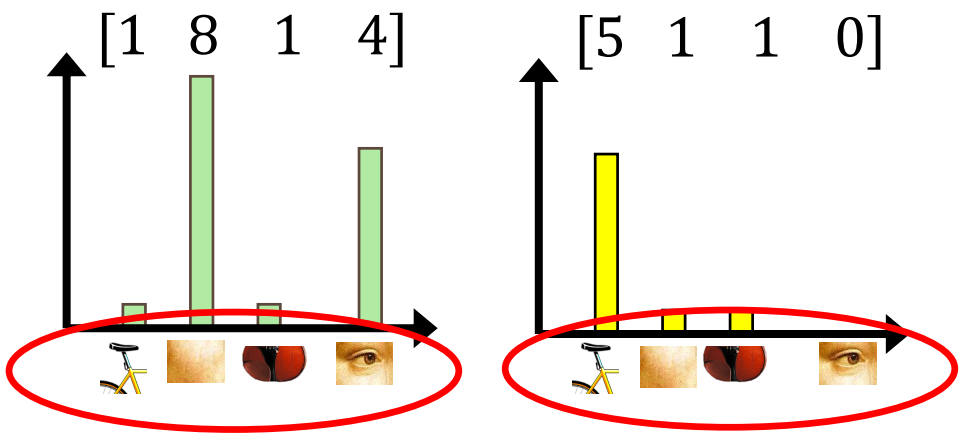
Bag of Words

- Learning and Recognition


learning
recognition
feature detection
& representation
codewords dictionary
image representation
category models
(and/or) classifiers
category
decision

Bag of Words

- Comparing bags of words



$\vec{d}_j$ $\vec{q}$

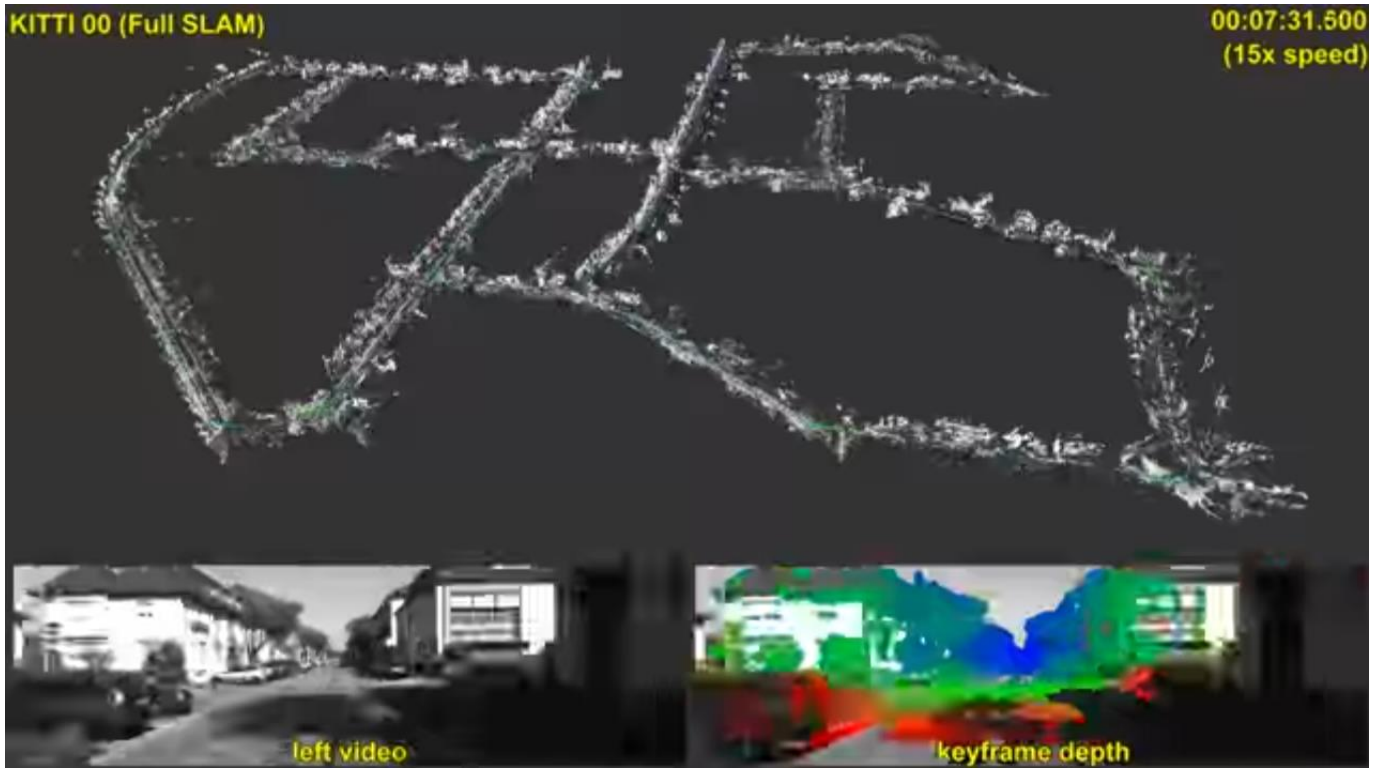
$$sim(d_j, q) = \frac{\langle d_j, q \rangle}{\left|\left|d_j\right|\right| \left|\left|q\right|\right|}$$

$$= \frac{\sum_{i=1}^{V} d_j(i) * q(i)}{\sqrt{\sum_{i=1}^{V} d_j(i)^2} * \sqrt{\sum_{i=1}^{V} q(i)^2}}$$

Result

- Loop closing (LSD-SLAM with Stereo Cameras)

https://www.youtube.com/watch?v=oJt3Ln8H03s



Q&A