

Facoltà di Scienze MM. FF. NN.

Università di Verona

A.A. 2014-15

# **Teoria e Tecniche del Riconoscimento**

*Estrazione delle feature: Bag of words*



# Part 1: Bag-of-words models

by Li Fei-Fei (Princeton)

# Related works

- Early “bag of words” models: mostly texture recognition
  - Cula & Dana, 2001; Leung & Malik 2001; Mori, Belongie & Malik, 2001; Schmid 2001; Varma & Zisserman, 2002, 2003; Lazebnik, Schmid & Ponce, 2003;
- Hierarchical Bayesian models for documents (pLSA, LDA, etc.)
  - Hoffman 1999; Blei, Ng & Jordan, 2004; Teh, Jordan, Beal & Blei, 2004
- Object categorization
  - Csurka, Bray, Dance & Fan, 2004; Sivic, Russell, Efros, Freeman & Zisserman, 2005; Sudderth, Torralba, Freeman & Willsky, 2005;
- Natural scene categorization
  - Vogel & Schiele, 2004; Fei-Fei & Perona, 2005; Bosch, Zisserman & Munoz, 2006

**Object**



**Bag of 'words'**



# Analogy to documents

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 reach our eyes.

For a long time, the visual image was considered as a movie scene. It is now known that the image is processed in a more complex way following the path to the various centers of the cortex, Hubel and Wiesel have demonstrated that the message about the image falling on the retina undergoes a fine analysis in a system of nerve cells stored in columns. In this system each cell has its specific function and is responsible 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**

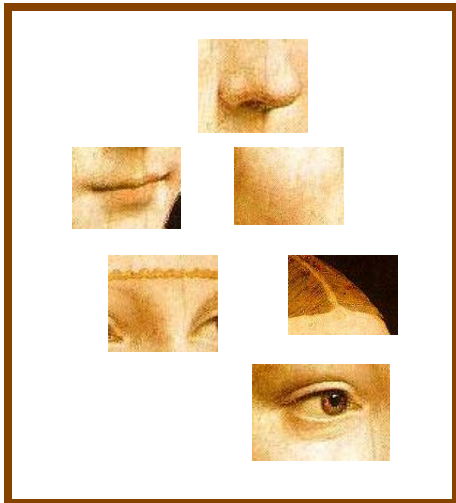
China is forecasting a trade surplus of \$90bn (£51bn) to \$100bn this year, a threefold increase on 2004's \$32bn. The Commerce Ministry said the surplus would be created by a predicted 30% increase in exports to \$750bn, compared with \$560bn in 2004.

The increase in exports will annoy the US. China's government has deliberately agreed to keep the yuan's value against the dollar at a level that also needs to be maintained to meet the demand so that the country can continue to grow. China has also announced that it will allow the yuan against the dollar to rise and permitted it to trade within a narrow band but the US wants the yuan to be allowed to rise freely. However, Beijing has made it clear that it will take its time and tread carefully before allowing the yuan to rise further in value.

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

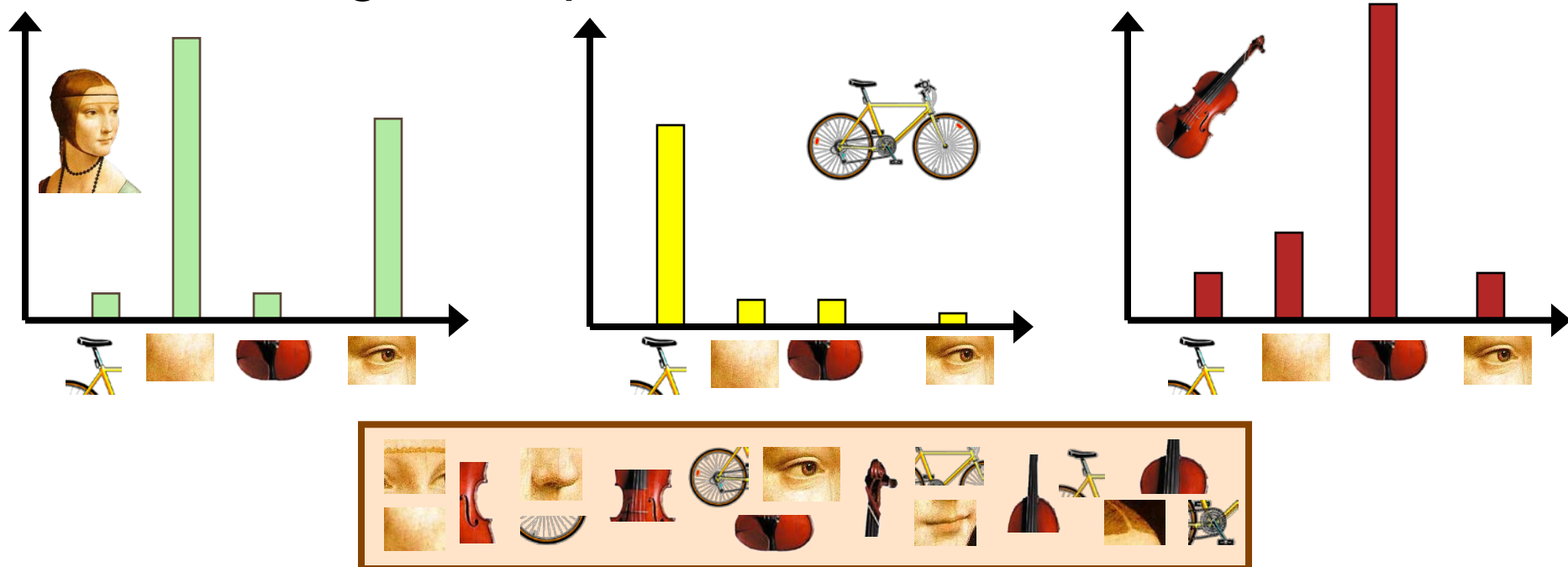
# A clarification: definition of “BoW”

- Looser definition
  - Independent features

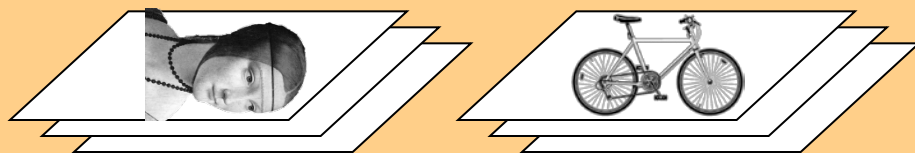


# A clarification: definition of “BoW”

- Looser definition
  - Independent features
- Stricter definition
  - Independent features
  - histogram representation



# learning



feature detection  
& representation

**codewords dictionary**

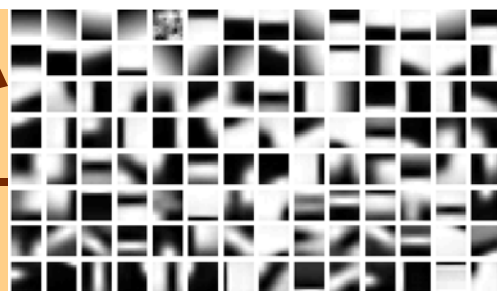
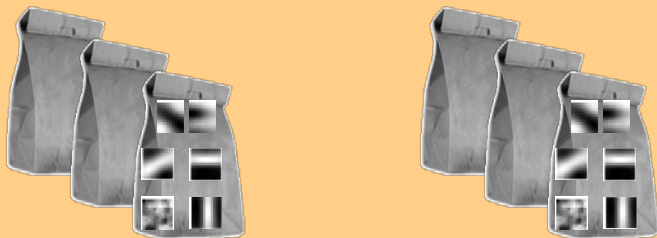


image representation



**category models  
(and/or) classifiers**

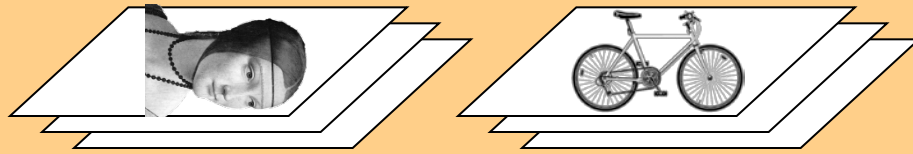
# recognition



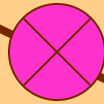
**category  
decision**



# Representation



1. feature detection  
& representation



2. **codewords dictionary**

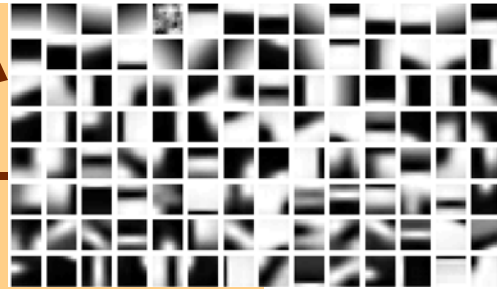
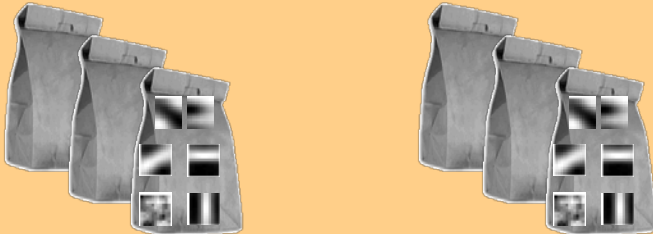
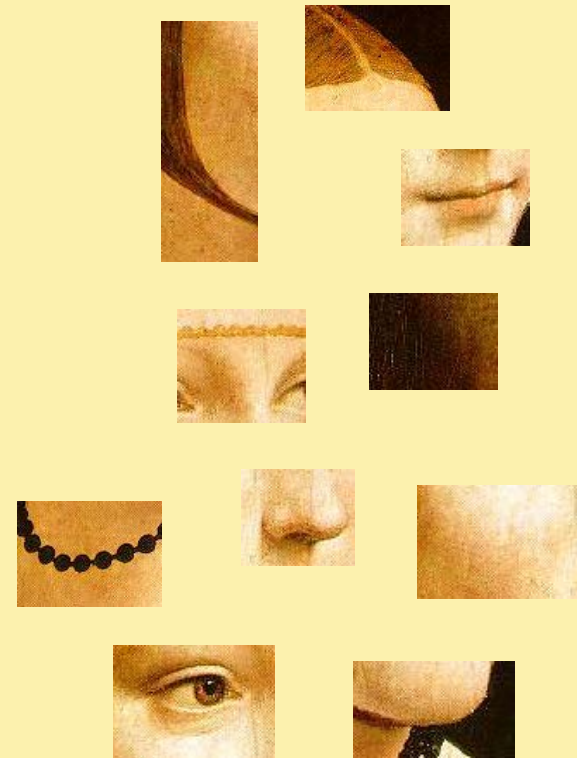


image representation

3.

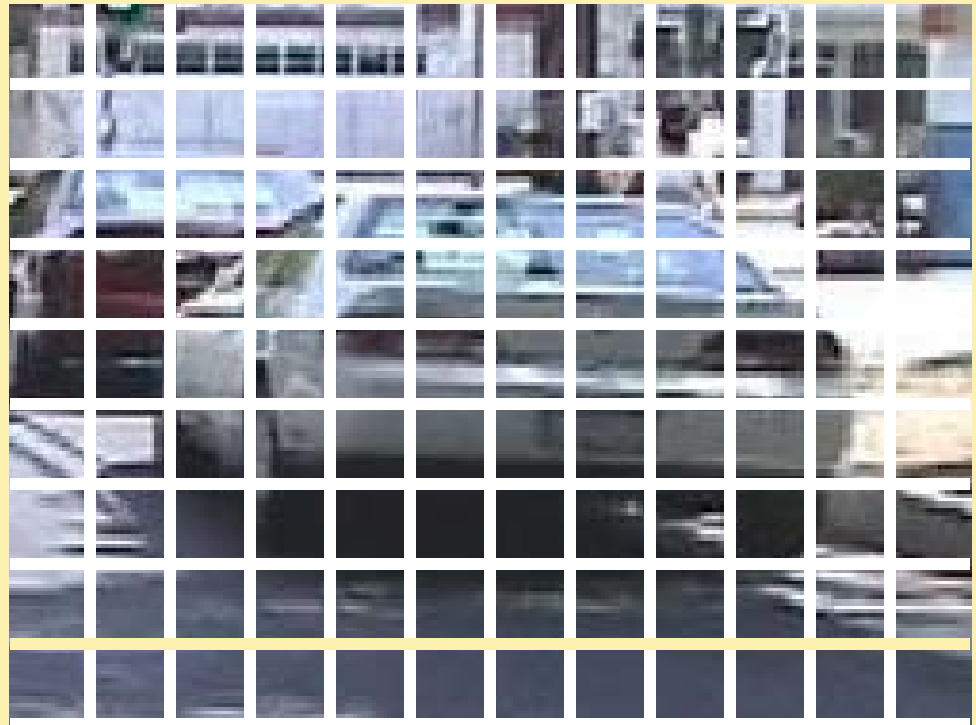


# 1.Feature detection and representation



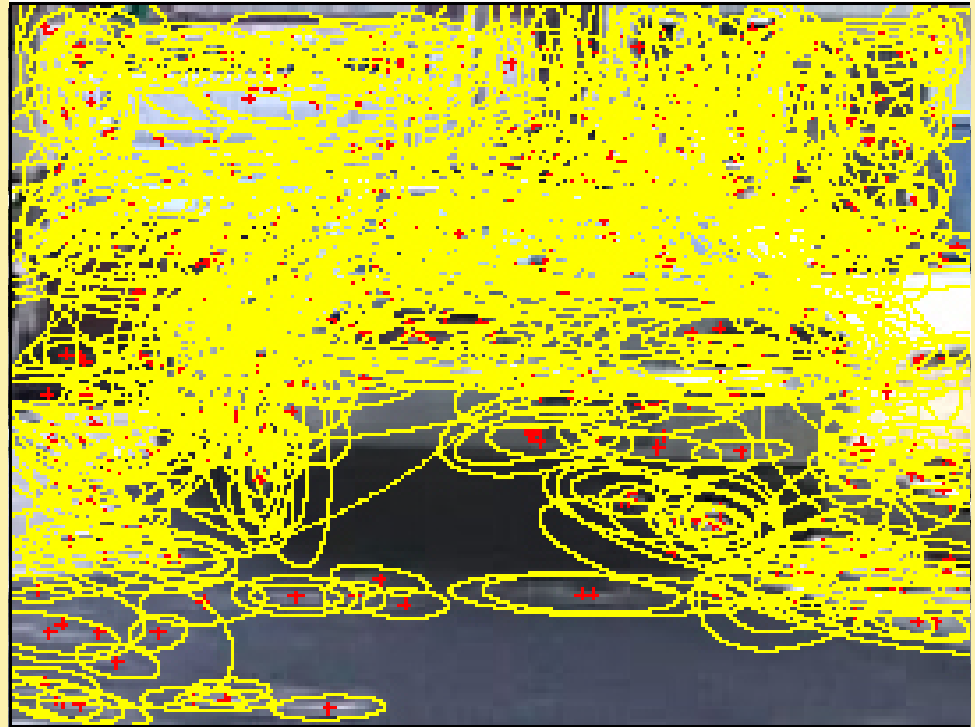
# 1.Feature detection and representation

- Regular grid
  - Vogel & Schiele, 2003
  - Fei-Fei & Perona, 2005



# 1. Feature detection and representation

- Regular grid
  - Vogel & Schiele, 2003
  - Fei-Fei & Perona, 2005
- Interest point detector
  - Csurka, et al. 2004
  - Fei-Fei & Perona, 2005
  - Sivic, et al. 2005



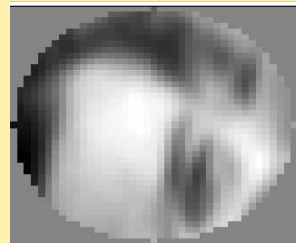
# 1.Feature detection and representation

- Regular grid
  - Vogel & Schiele, 2003
  - Fei-Fei & Perona, 2005
- Interest point detector
  - Csurka, Bray, Dance & Fan, 2004
  - Fei-Fei & Perona, 2005
  - Sivic, Russell, Efros, Freeman & Zisserman, 2005
- Other methods
  - Random sampling (Vidal-Naquet & Ullman, 2002)
  - Segmentation based patches (Barnard, Duygulu, Forsyth, de Freitas, Blei, Jordan, 2003)

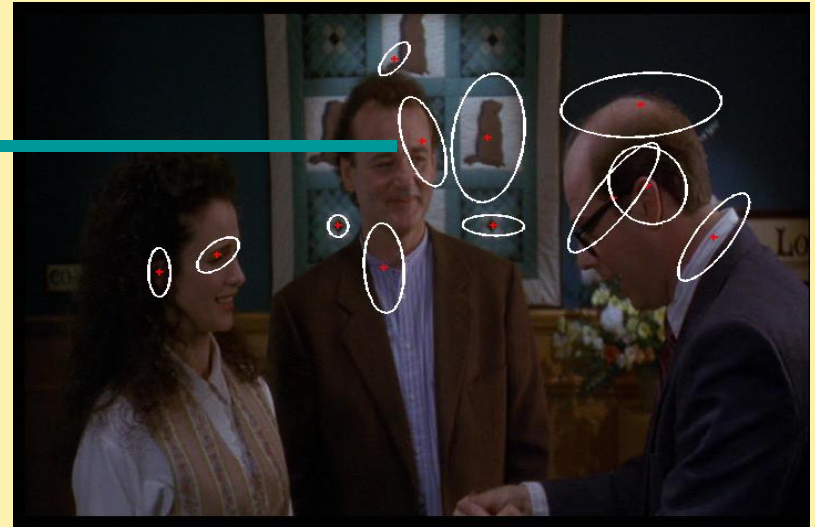
# 1. Feature detection and representation



**Compute  
SIFT  
descriptor**  
[Lowe'99]



**Normalize  
patch**



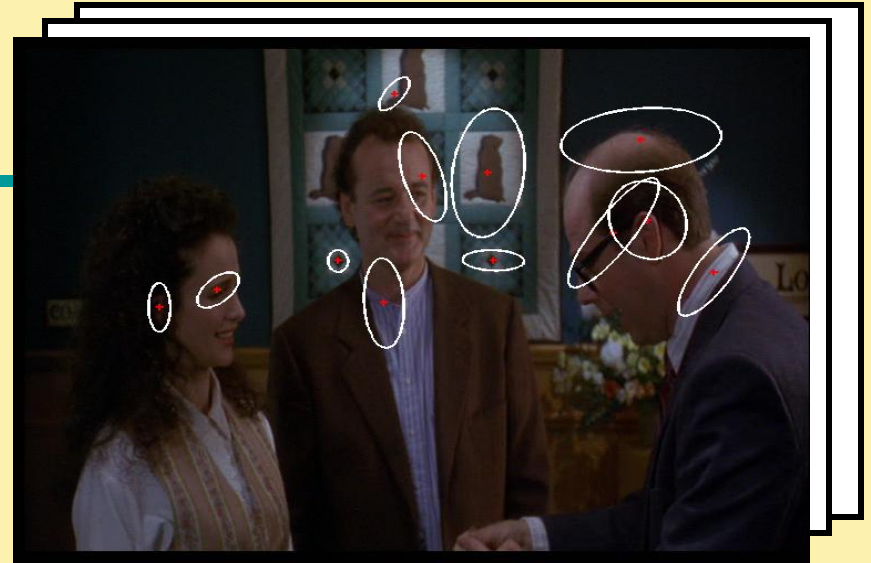
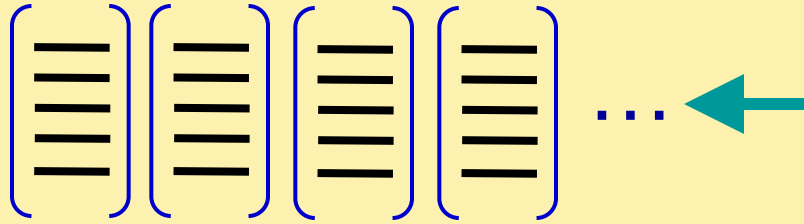
**Detect patches**

[Mikojaczyk and Schmid '02]

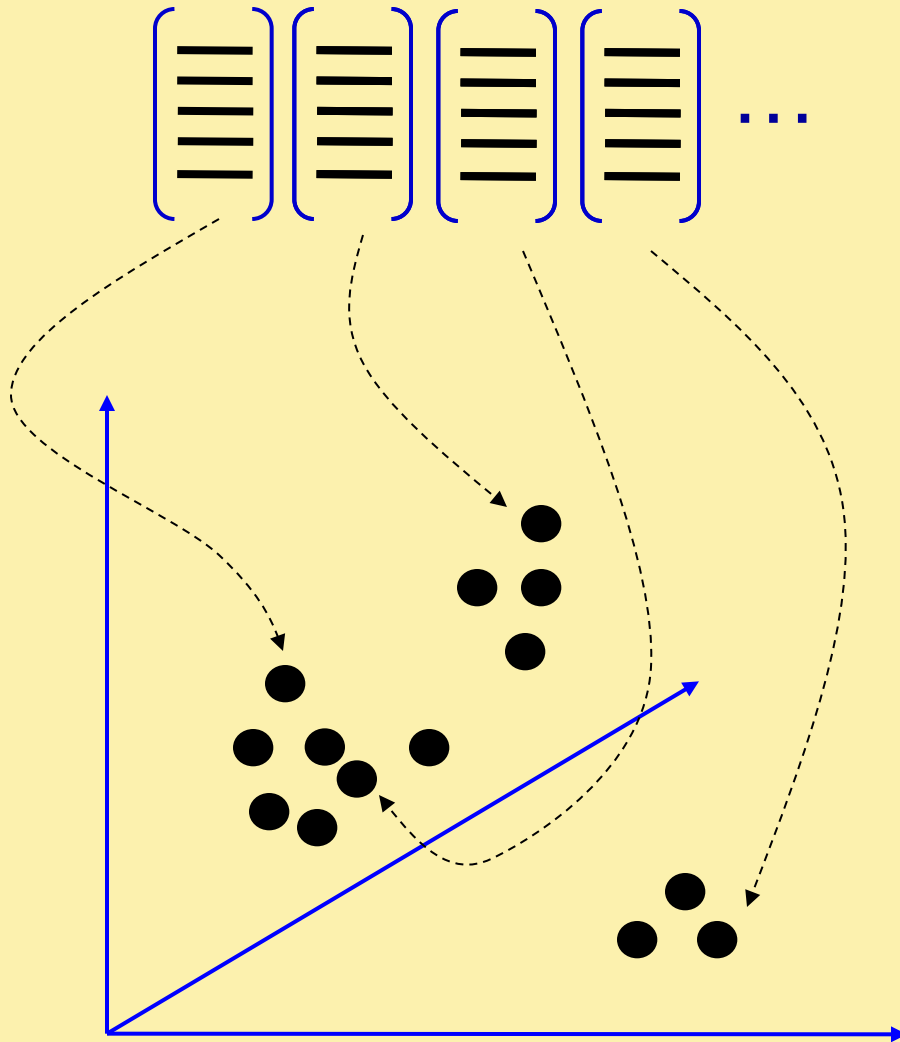
[Mata, Chum, Urban & Pajdla, '02]

[Sivic & Zisserman, '03]

# 1. Feature detection and representation

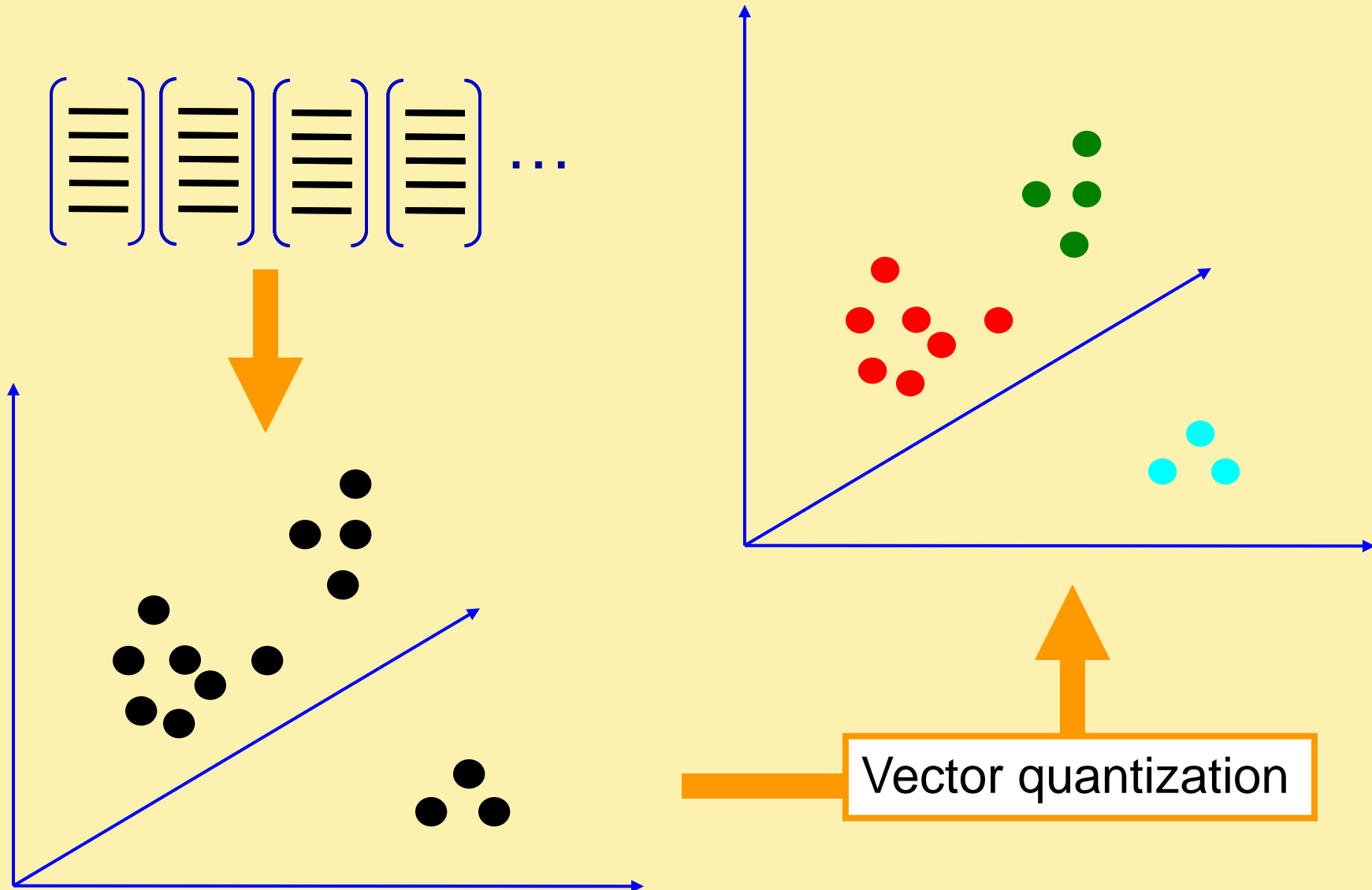


## 2. Codewords dictionary formation

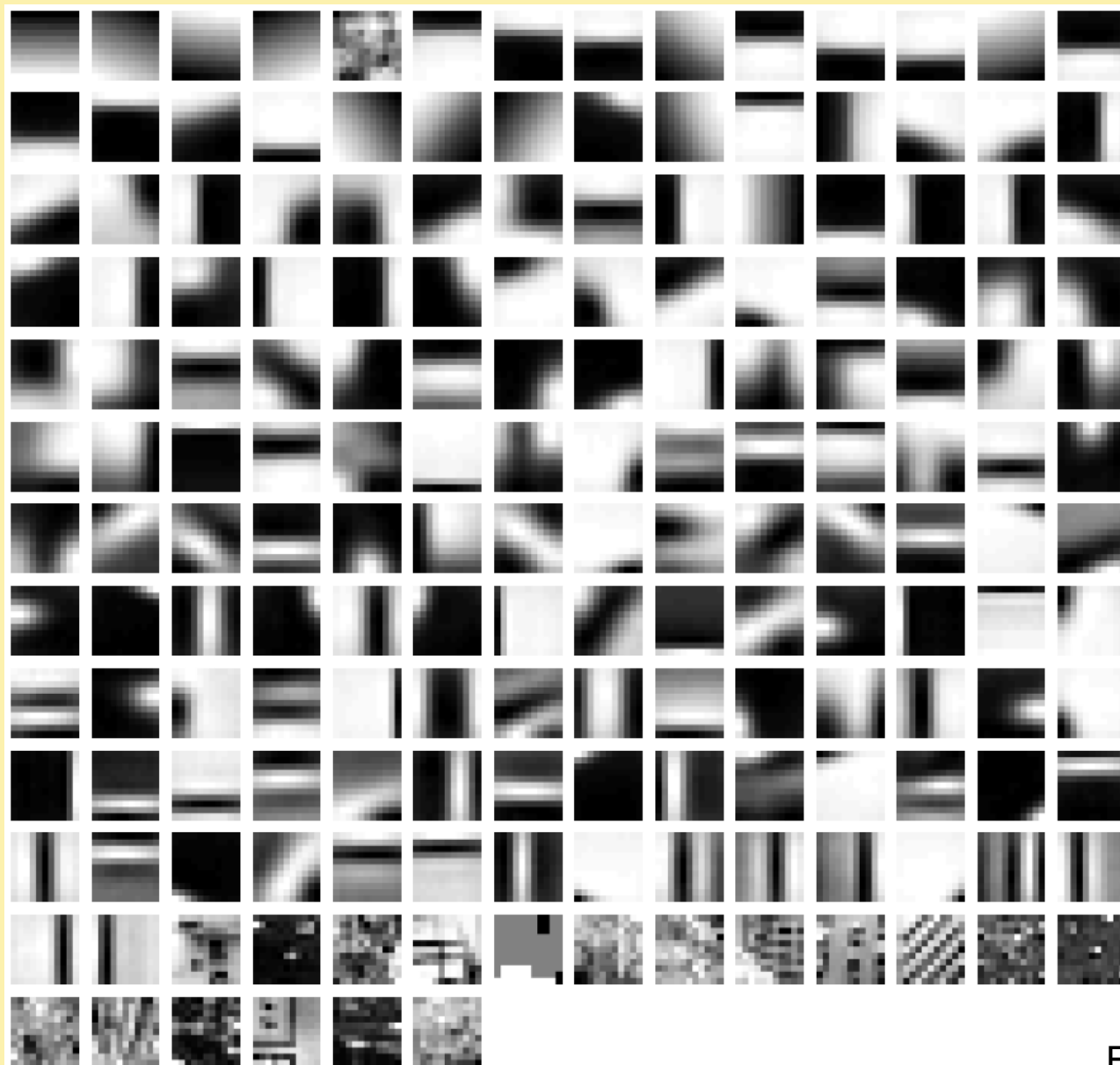




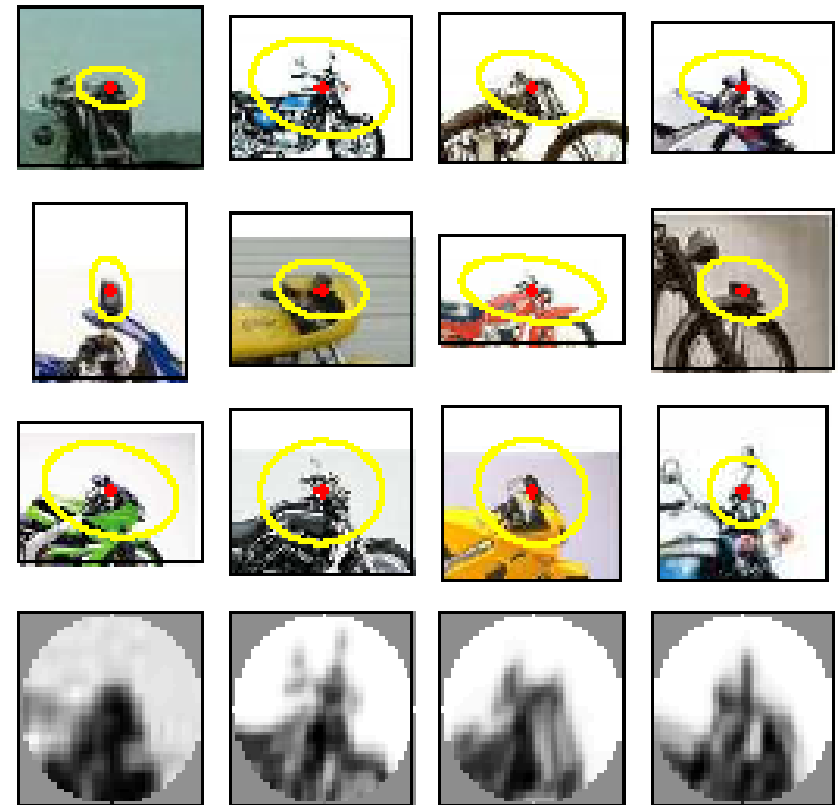
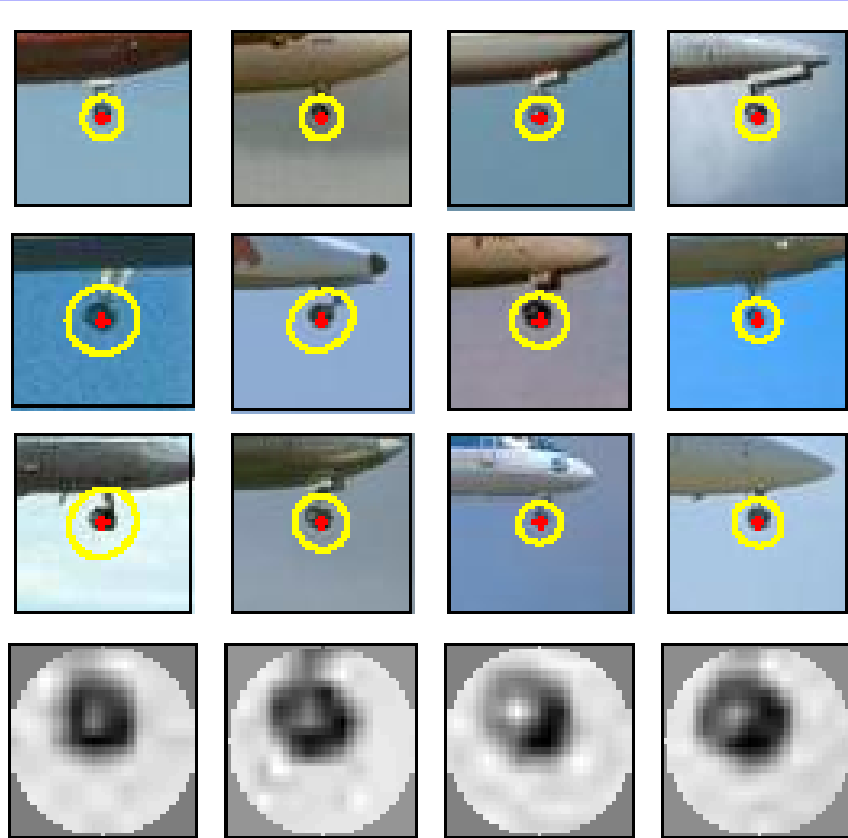
## 2. Codewords dictionary formation



## 2. Codewords dictionary formation



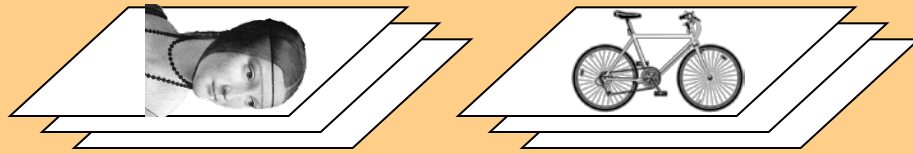
# Image patch examples of codewords



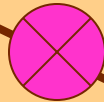
### 3. Image representation



# Representation



1. feature detection  
& representation



2. **codewords dictionary**

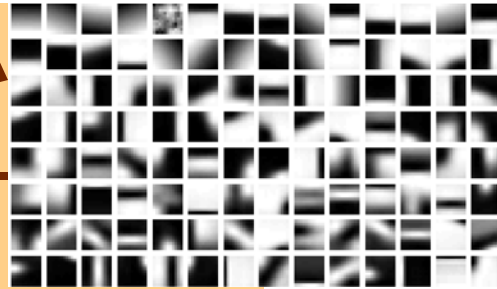
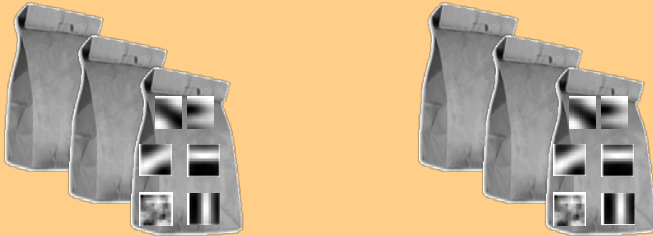
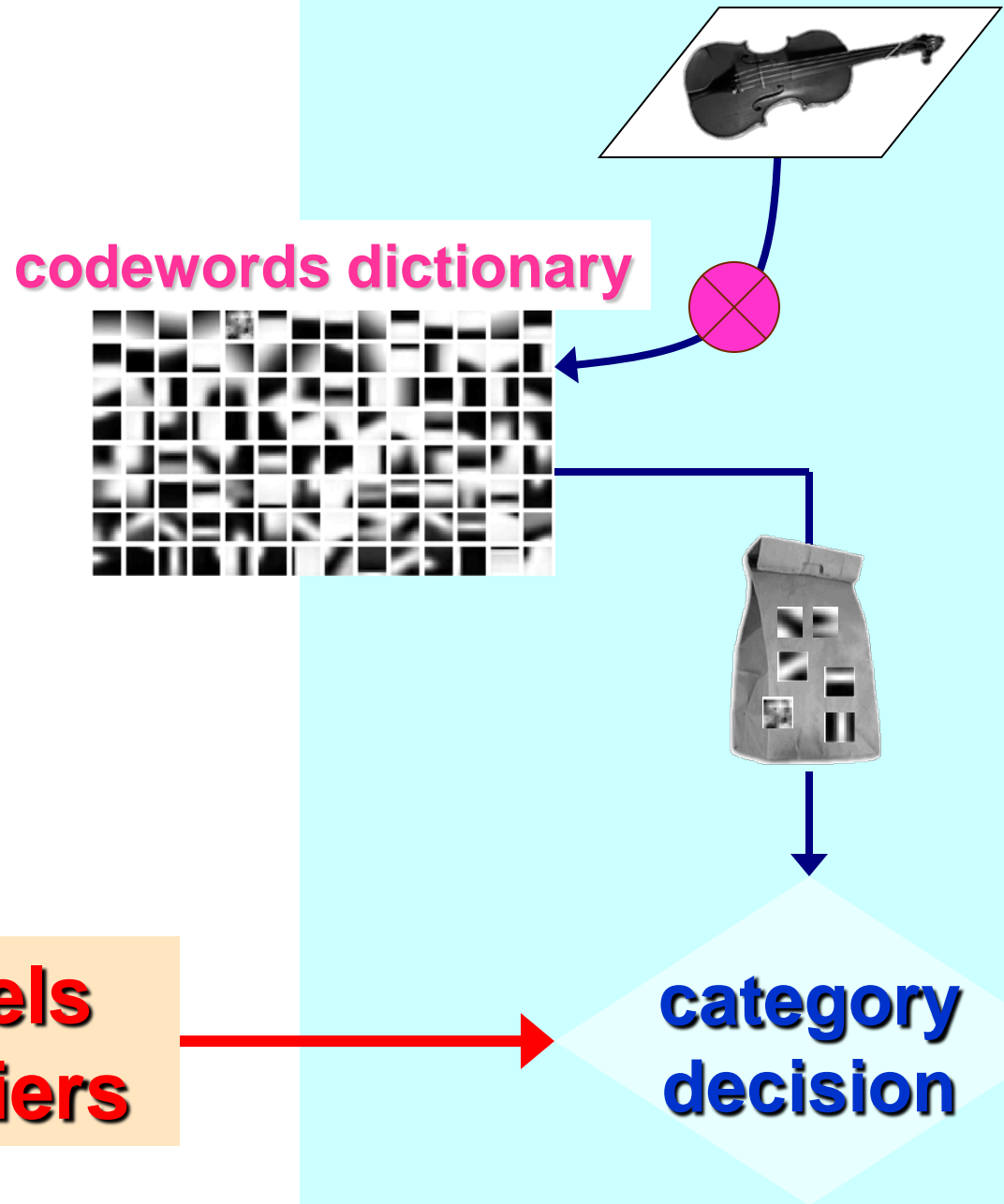


image representation

3.



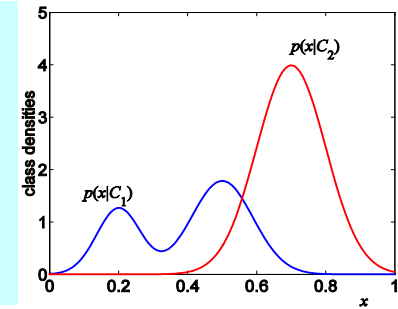
# Learning and Recognition



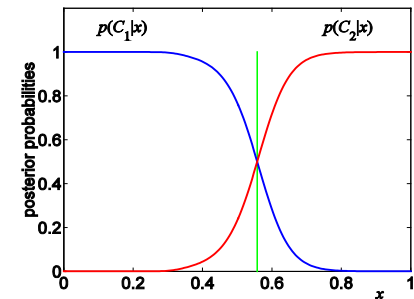
**category models  
(and/or) classifiers**

# Learning and Recognition

1. Generative method:  
- graphical models



2. Discriminative method:  
- SVM



**category models  
(and/or) classifiers**