

6.S093 Visual Recognition through Machine Learning Competition

Aditya Khosla

Today's class

- Part 1: Competition details
- Part 2: Image representation lecture
 - Bag-of-words
 - Spatial pyramid
- Part 3: Feature extraction tutorial

Competition details: dataset



airplane



bicycle



car



person

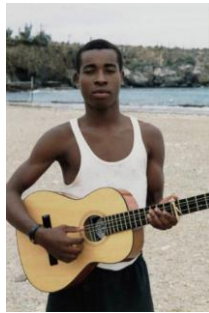
10 object categories



cup/mug



dog(s)



guitar



hamburger

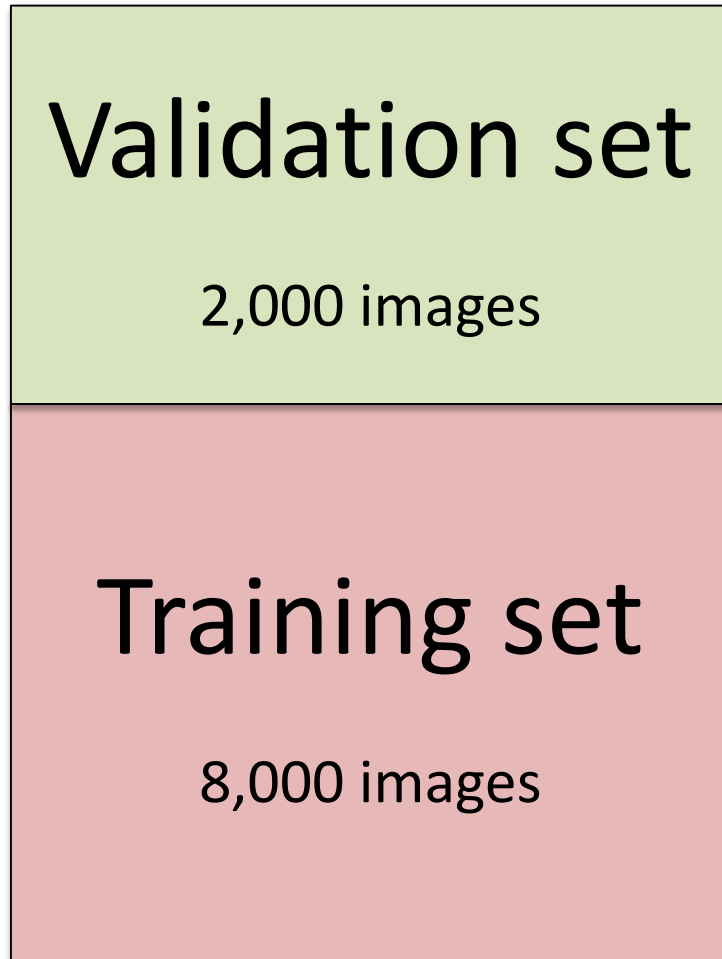


sofa

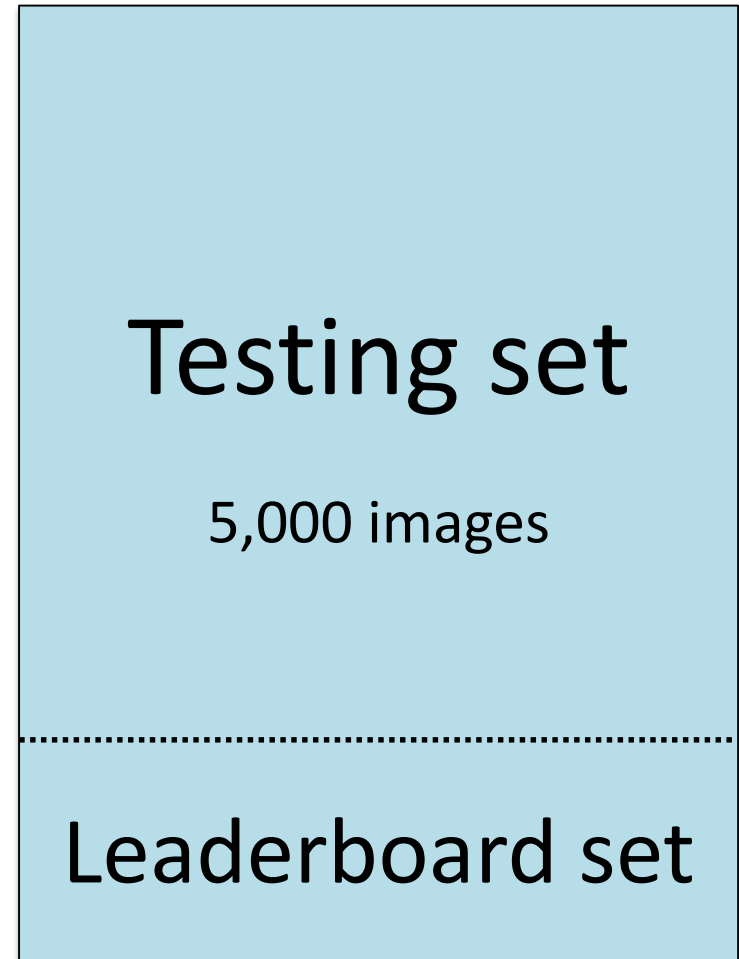


traffic
light

Competition details: dataset



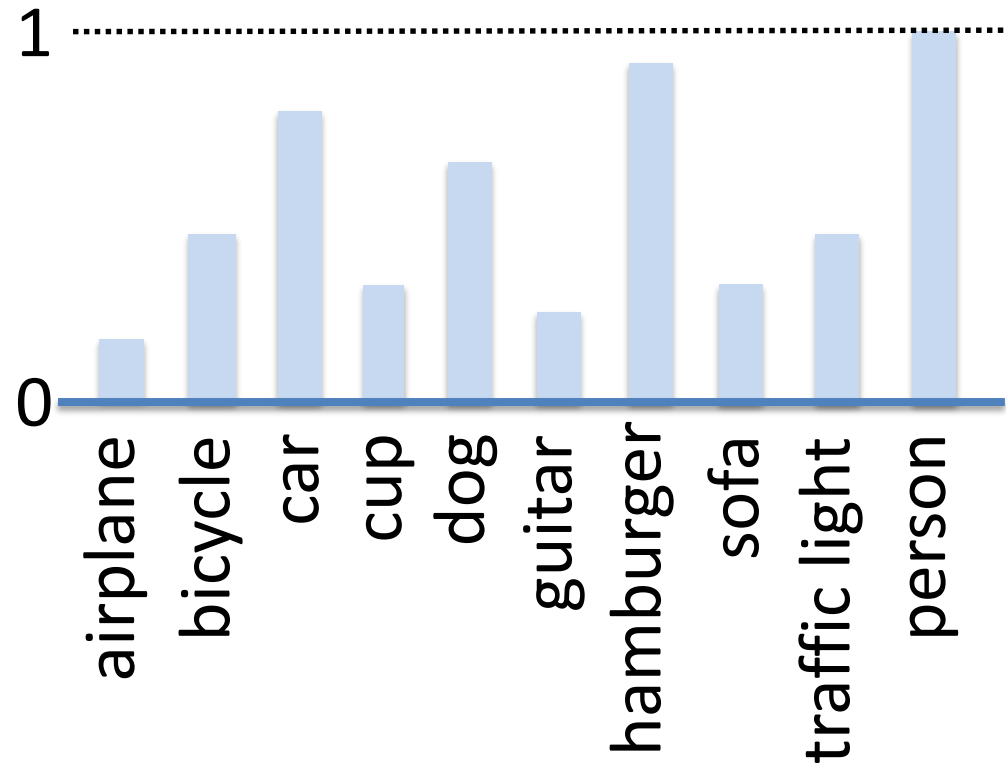
labels provided



NO labels provided

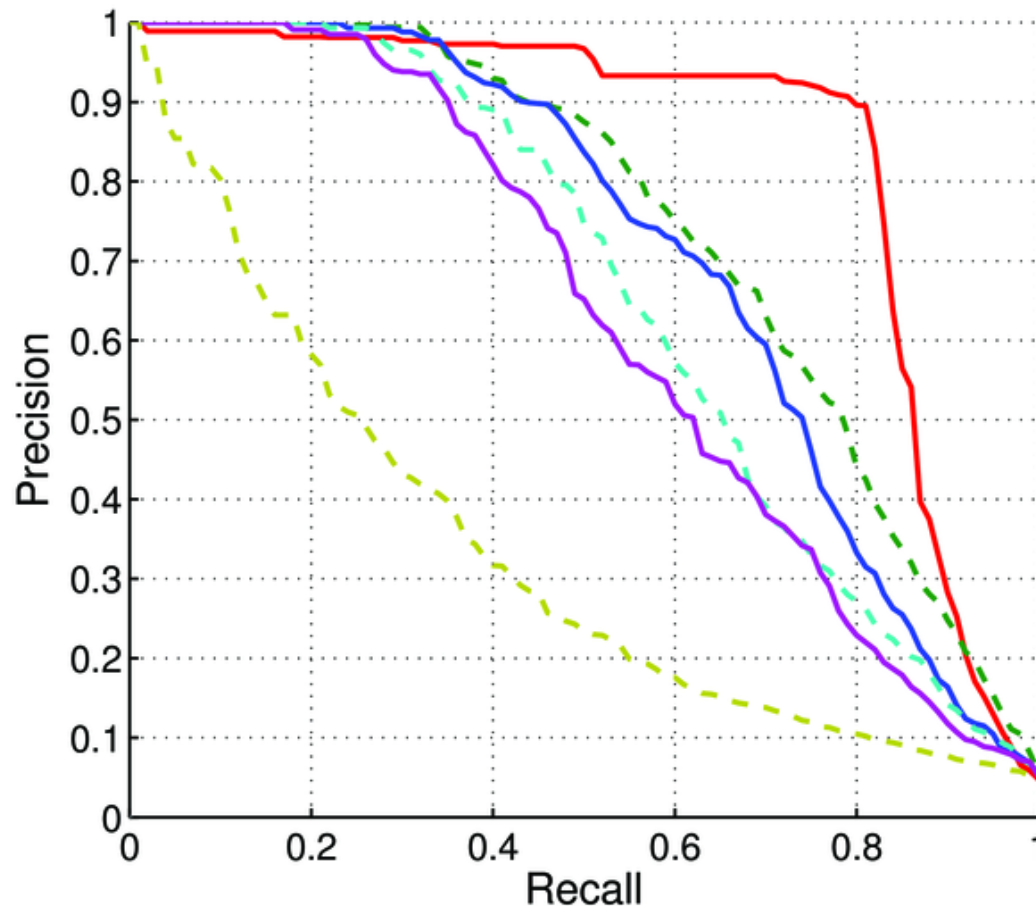
Competition details: submission

- For each image, you provide the probability of every class belonging in it (as returned by your algorithm)



Competition details: evaluation

- Average precision



Competition details: prizes

first



+
cash

second



+
cash

third

Cash

Competition details: thank you!



Microsoft

Google



Image representation: bag-of-words

Document representation: bag-of-words

- Order-less document representation: frequencies of words from a dictionary Salton & McGill (1983)

Document representation: bag-of-words

- Order-less document representation: frequencies of words from a dictionary Salton & McGill (1983)



Document representation: bag-of-words

- Order-less document representation: frequencies of words from a dictionary Salton & McGill (1983)

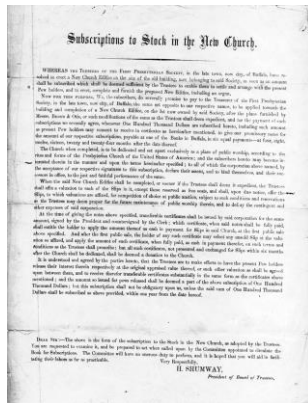


Document representation: bag-of-words

- Order-less document representation: frequencies of words from a dictionary Salton & McGill (1983)



Image representation: bag-of-words



document



2007-01-23: State of the Union Address

George W. Bush (2001-)

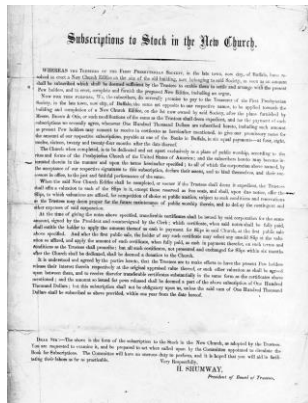
abandon accountable affordable afghanistan africa aided ally anbar armed army **baghdad** bless challenges chamber chaos
choices civilians coalition commanders **commitment** confident confront congressman constitution corps debates deduction
deficit deliver democracy deploy dikembe diplomacy disruptions earmarks **economy** einstein elections eliminates
expand extremists falling faithful families freedom fuel funding god have haven ideology immigration impose

insurgents **iraq** islam julie lebanon love madam marine math medicine **moderation** neighborhoods nuclear offensive
palestinian payroll province pursuing **qaeda** radical regimes resolve retreat rieman sacrifices science sectarian senate

september **shia** stays strength students succeed sunni tax territories **terrorists** threats uphold victory
violence violent War washington weapons wesley

bag-of-words

Image representation: bag-of-words



document



2007-01-23: State of the Union Address

George W. Bush (2001-)

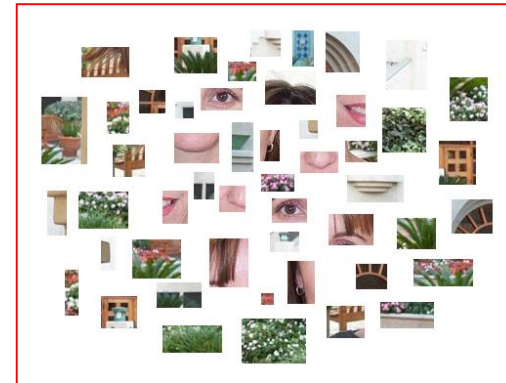
abandon accountable affordable afghanistan africa aided ally anbar armed army baghdad bless challenges chamber chaos
choices civilians coalition commanders commitment confident confront congressman constitution corps debates deduction
deficit deliver democratic deploy dikeembe diplomacy disruptions earmarks economy einstein elections eliminates
expand extremists falling faithful families freedom fuel funding god haven ideology immigration impose

insurgents iran **iraq** islam julie lebanon love madam marine math medicare moderation neighborhoods nuclear offensive
palestinian payroll province pursuing **qaeda** radical regimes resolve retreat rieman sacrifices science sectarian senate
september **shia** stays strength students succeed sunni tax territories **terrorists** threats uphold victory
violence violent war washington weapons wesley

bag-of-words



image



bag-of-visual words

Object



Bag of 'words'



Object



**Ugly bag of
'words'**



Object



**Stylish bag of
'words'**

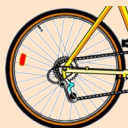
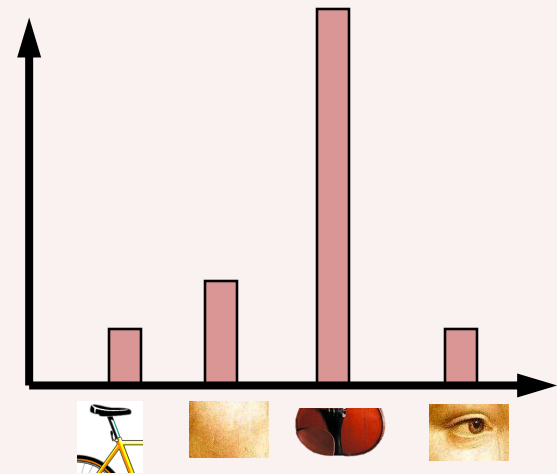
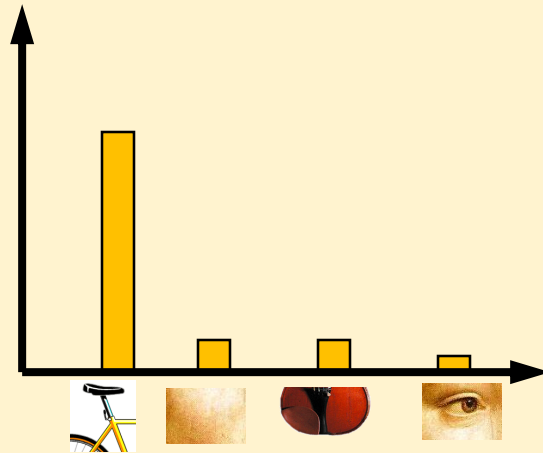
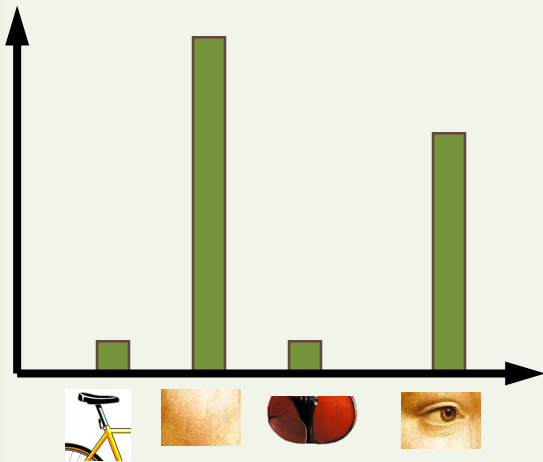


Object



Stylish bag of words





visual dictionary



Image representation: bag-of-words

1. Extract descriptors

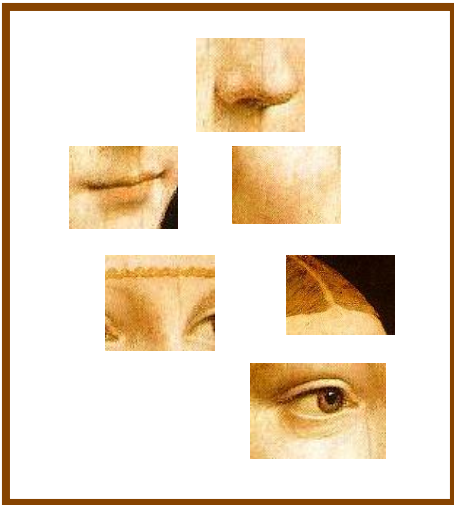


Image representation: bag-of-words

1. Extract descriptors
2. Learn “visual dictionary”

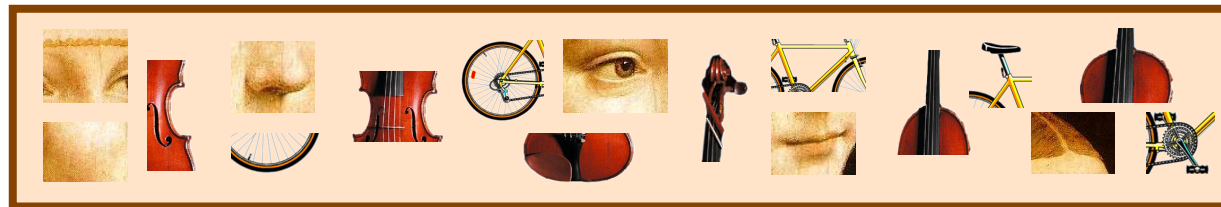


Image representation: bag-of-words

1. Extract descriptors
2. Learn “visual dictionary”
3. Quantize features using visual vocabulary

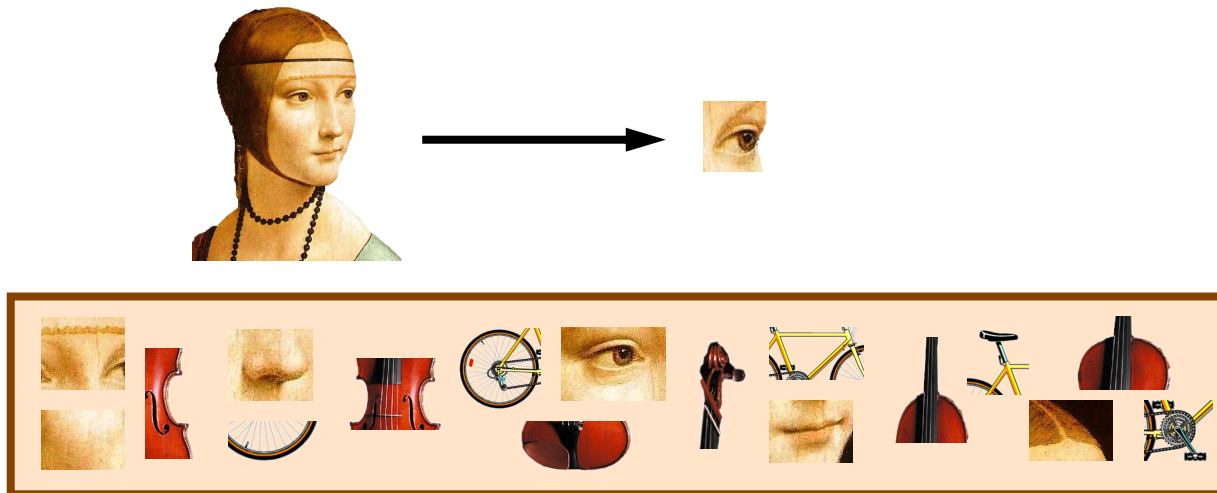


Image representation: bag-of-words

1. Extract descriptors
2. Learn “visual dictionary”
3. Quantize features using visual vocabulary

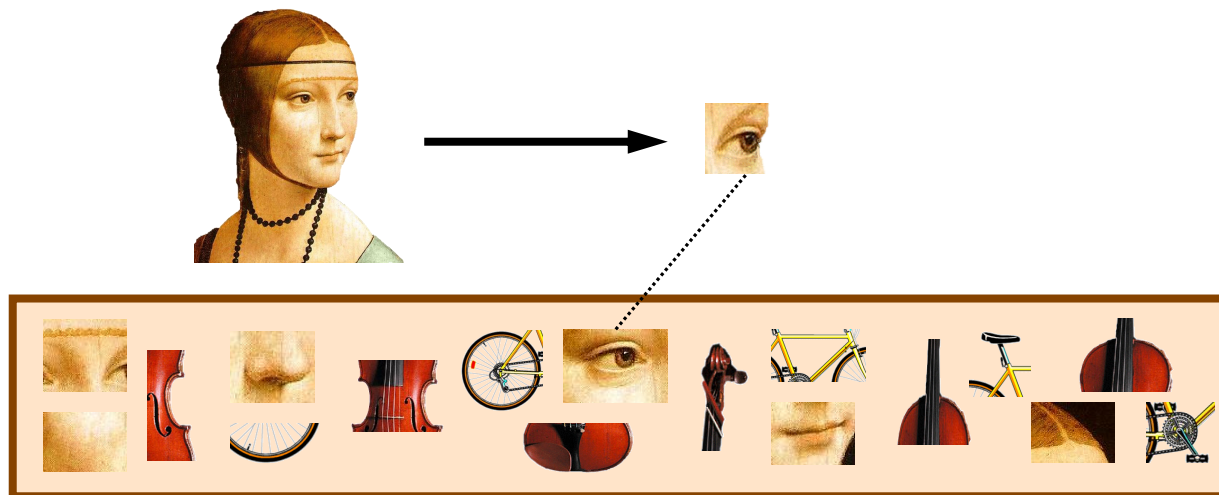
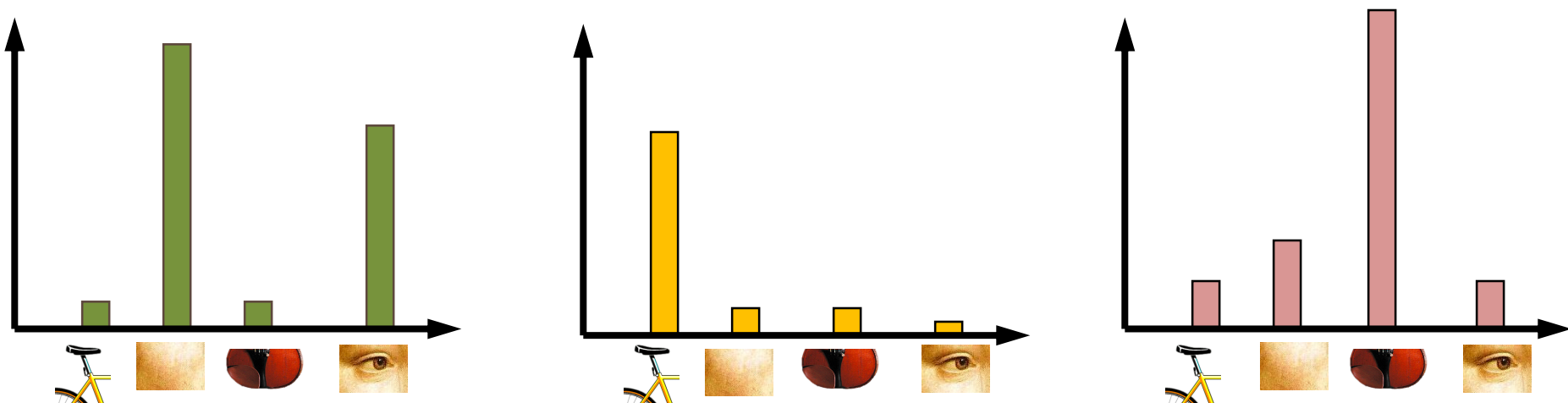


Image representation: bag-of-words

1. Extract descriptors
2. Learn “visual dictionary”
3. Quantize features using visual vocabulary
4. Represent images by frequencies of “visual words”



1. Extracting descriptors

regular grid



interest points

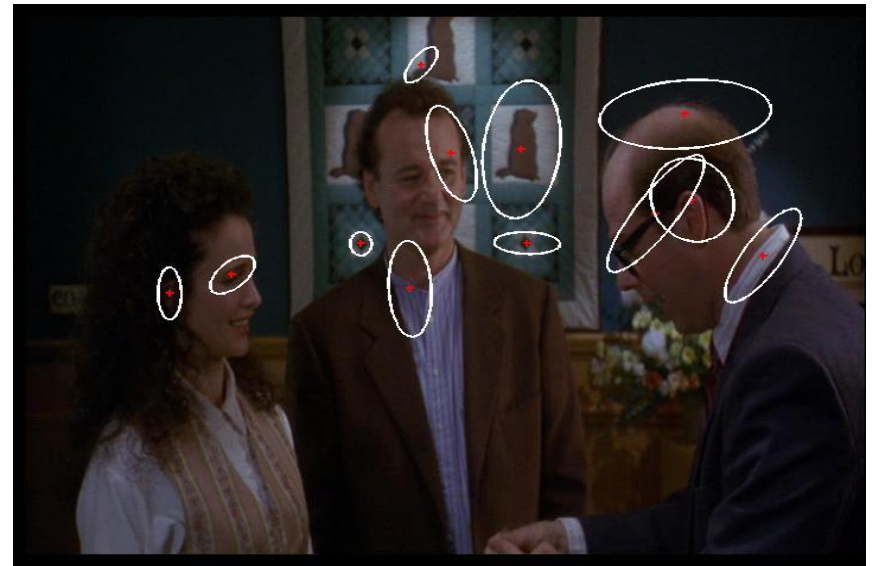


Image representation: yesterday

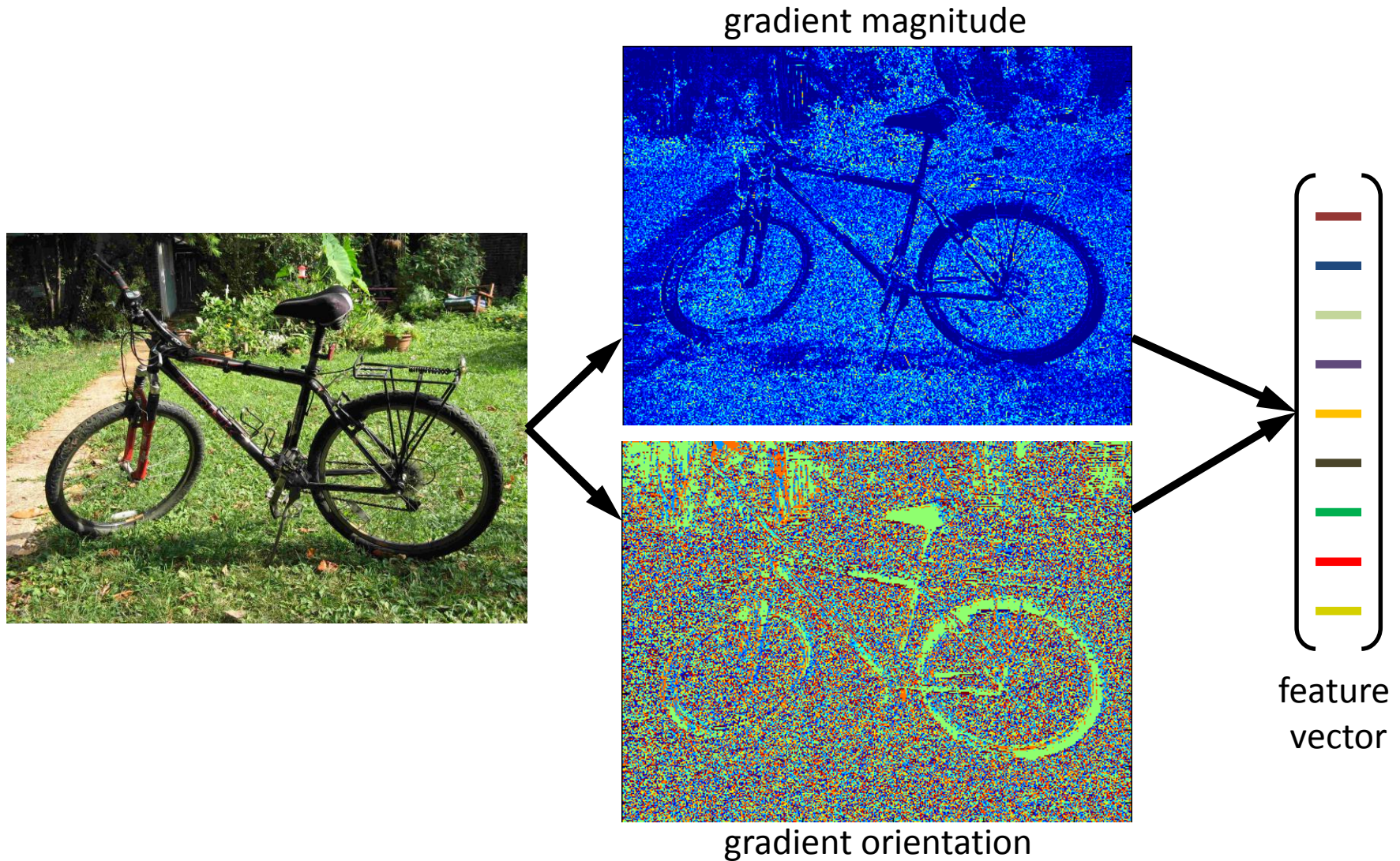
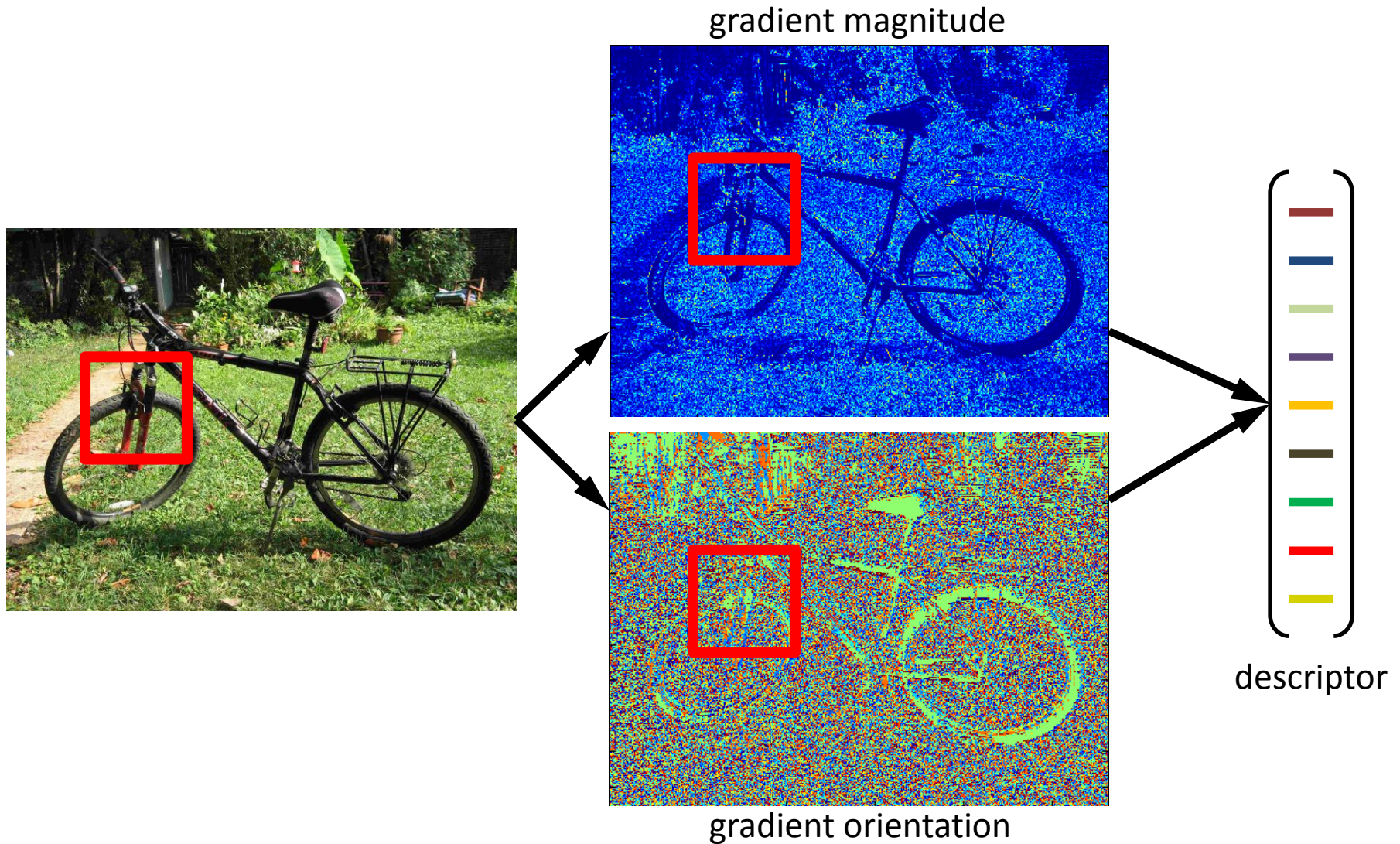
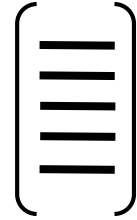


Image representation: yesterday

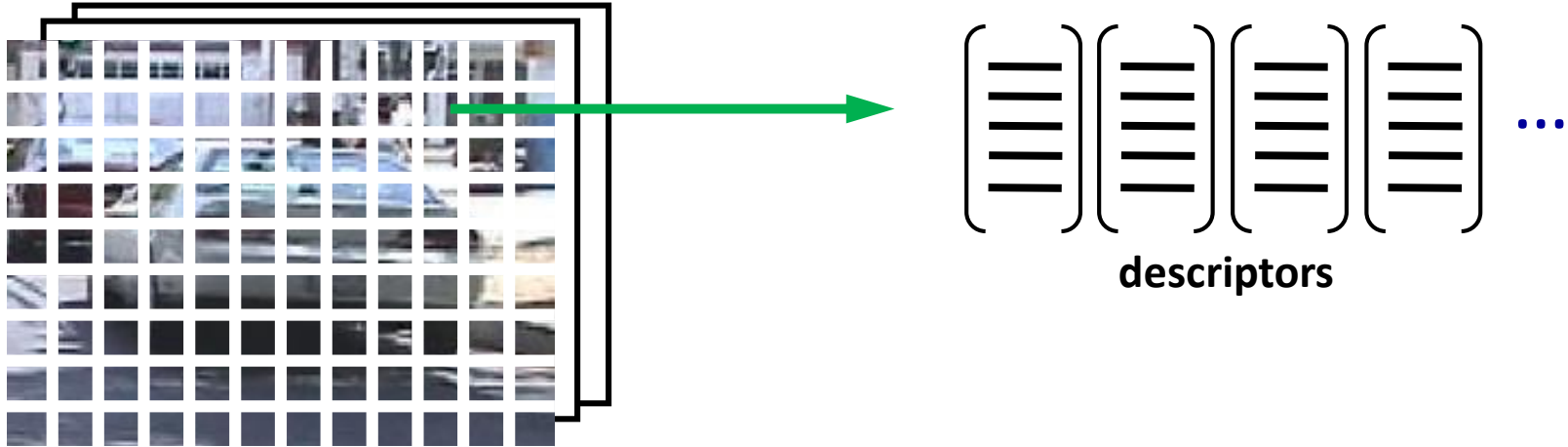


2. Learning “visual dictionary”

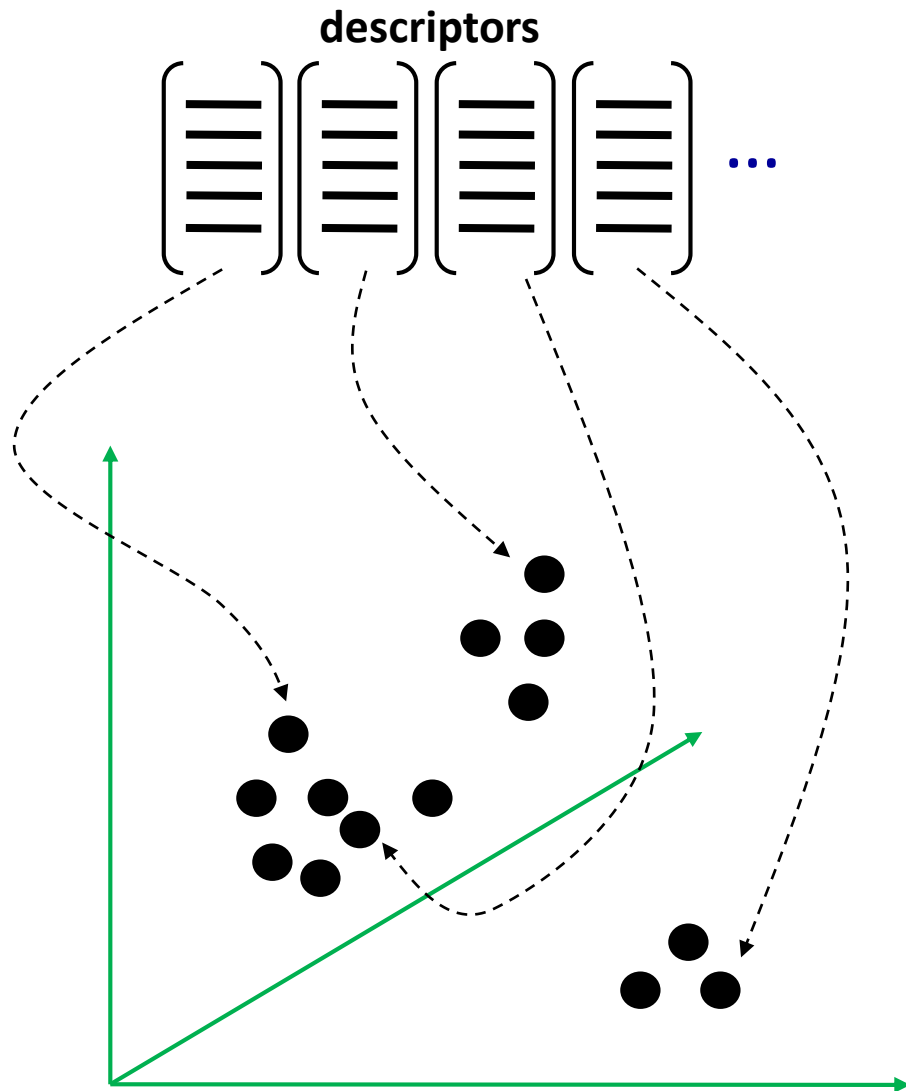


Compute
descriptor

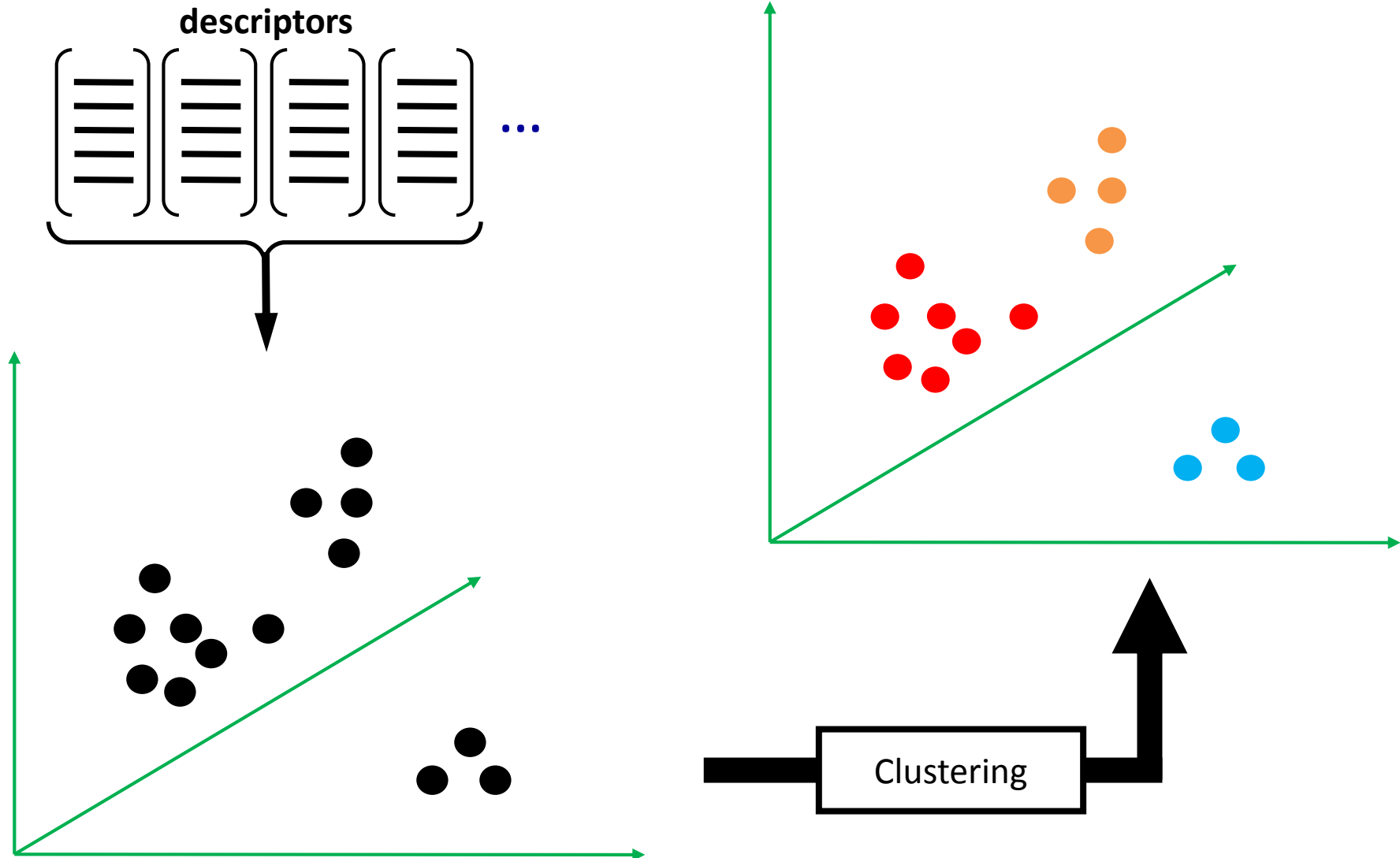
2. Learning “visual dictionary”



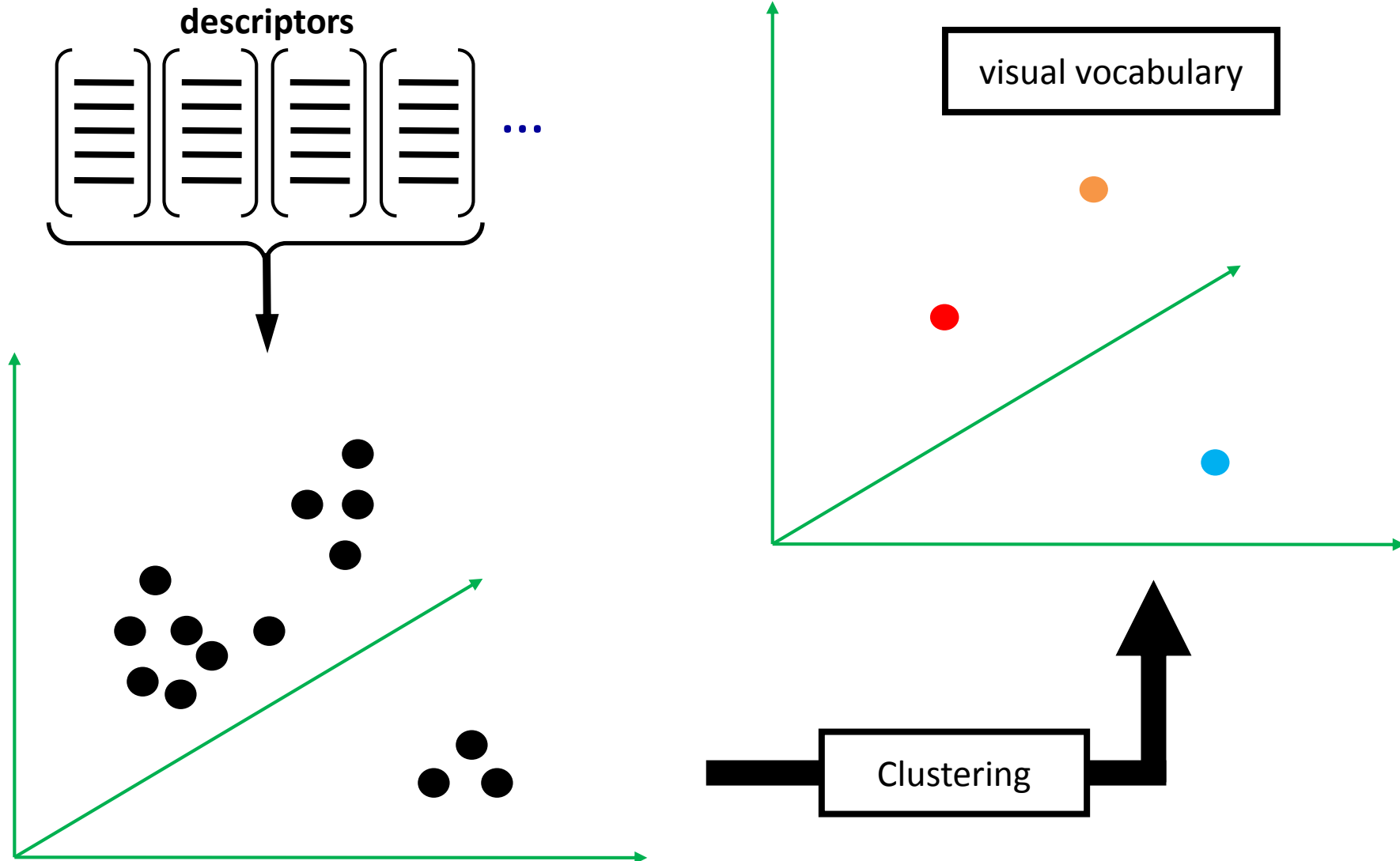
2. Learning visual dictionary



2. Learning visual dictionary



2. Learning visual dictionary



Example visual vocabulary

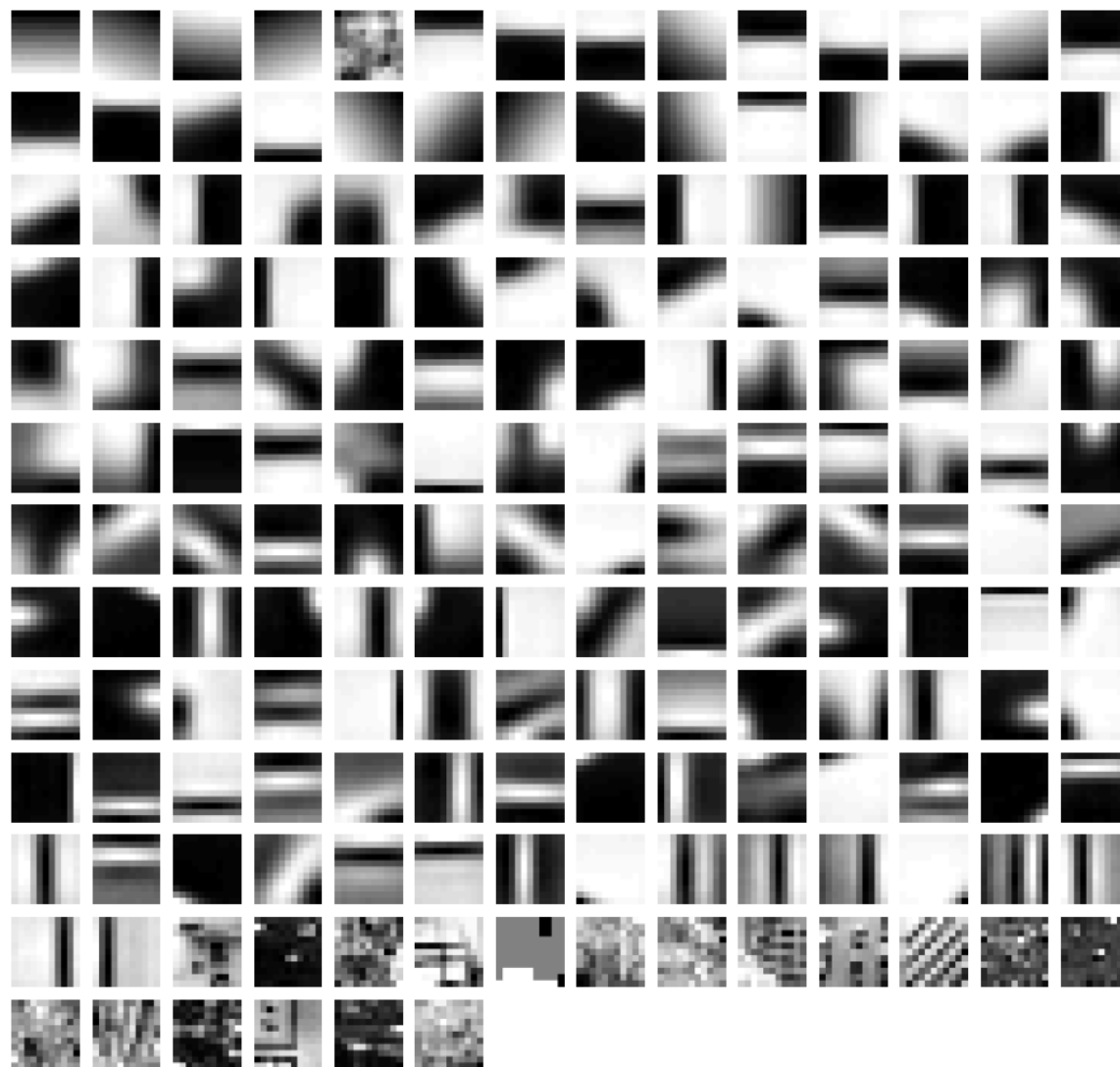


Image patch examples

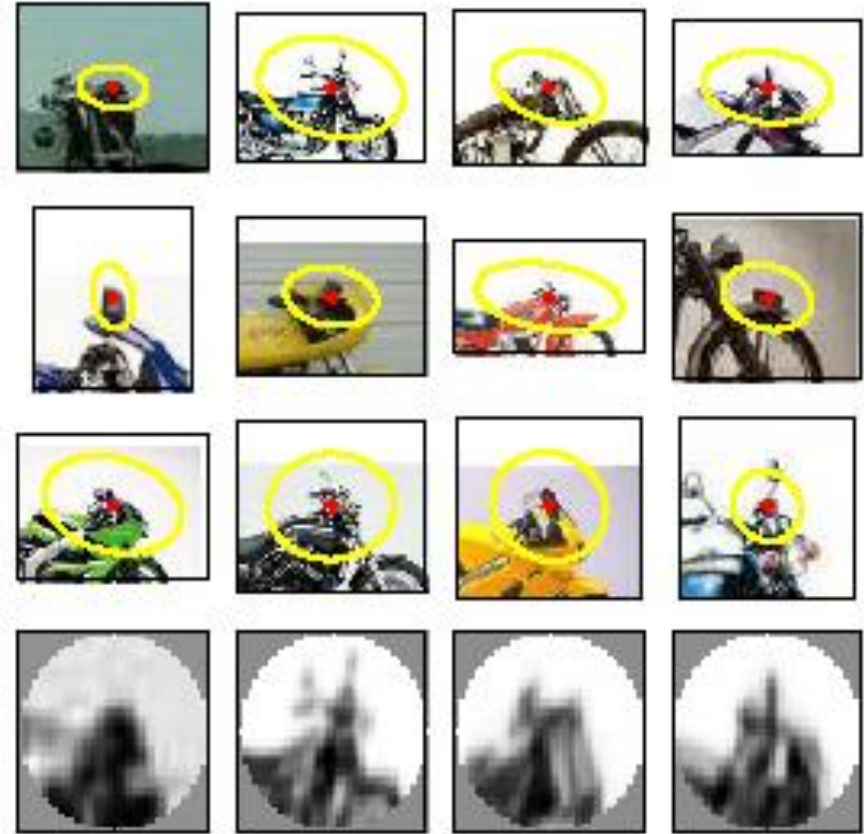
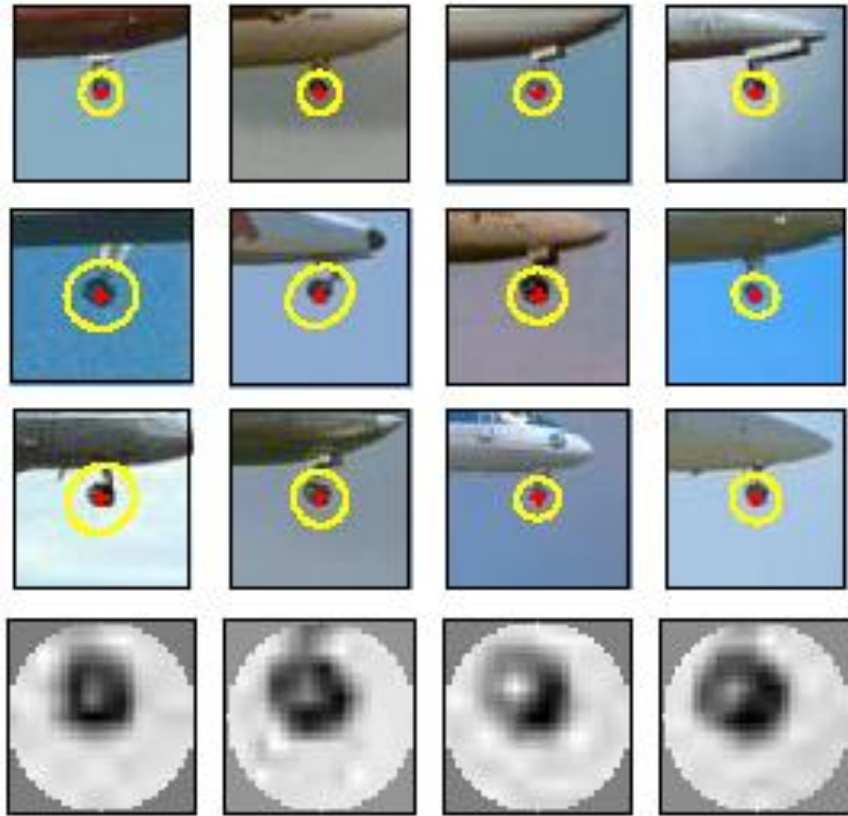
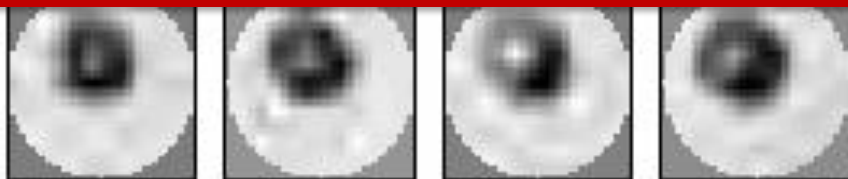


Image patch examples



How to choose the vocabulary size?



Bag-of-words: limitations

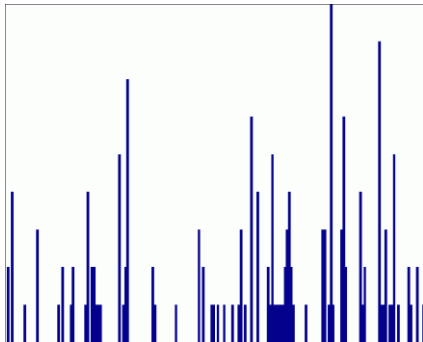
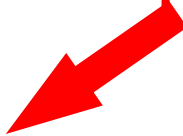
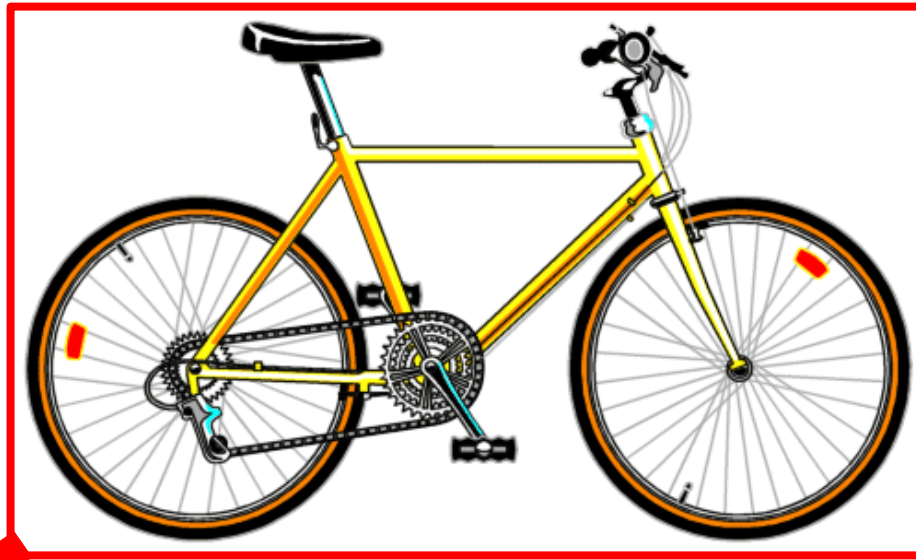
- What about the structure of the image?



=?



Image representation: spatial pyramids



level 0

Image representation: spatial pyramids

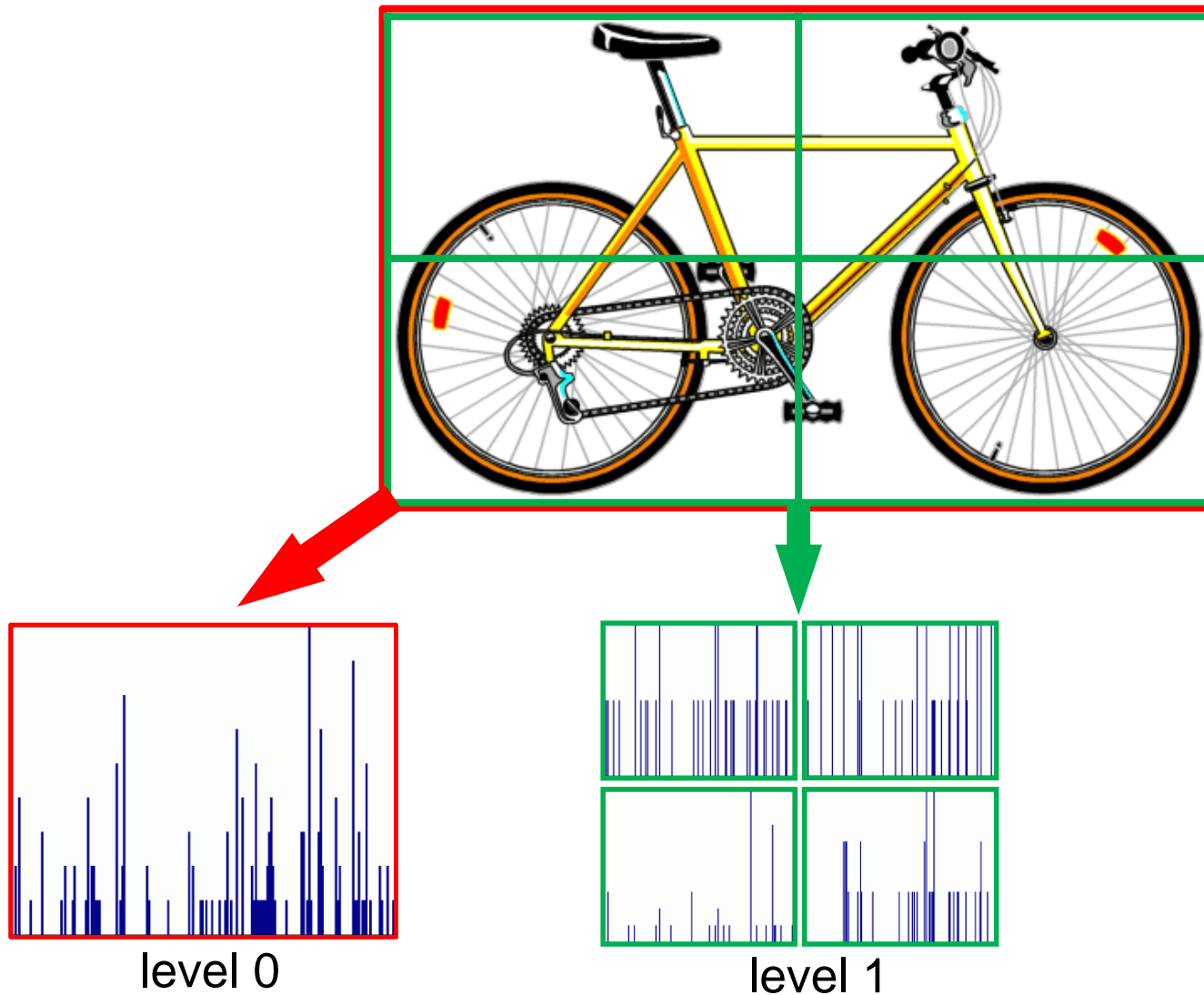
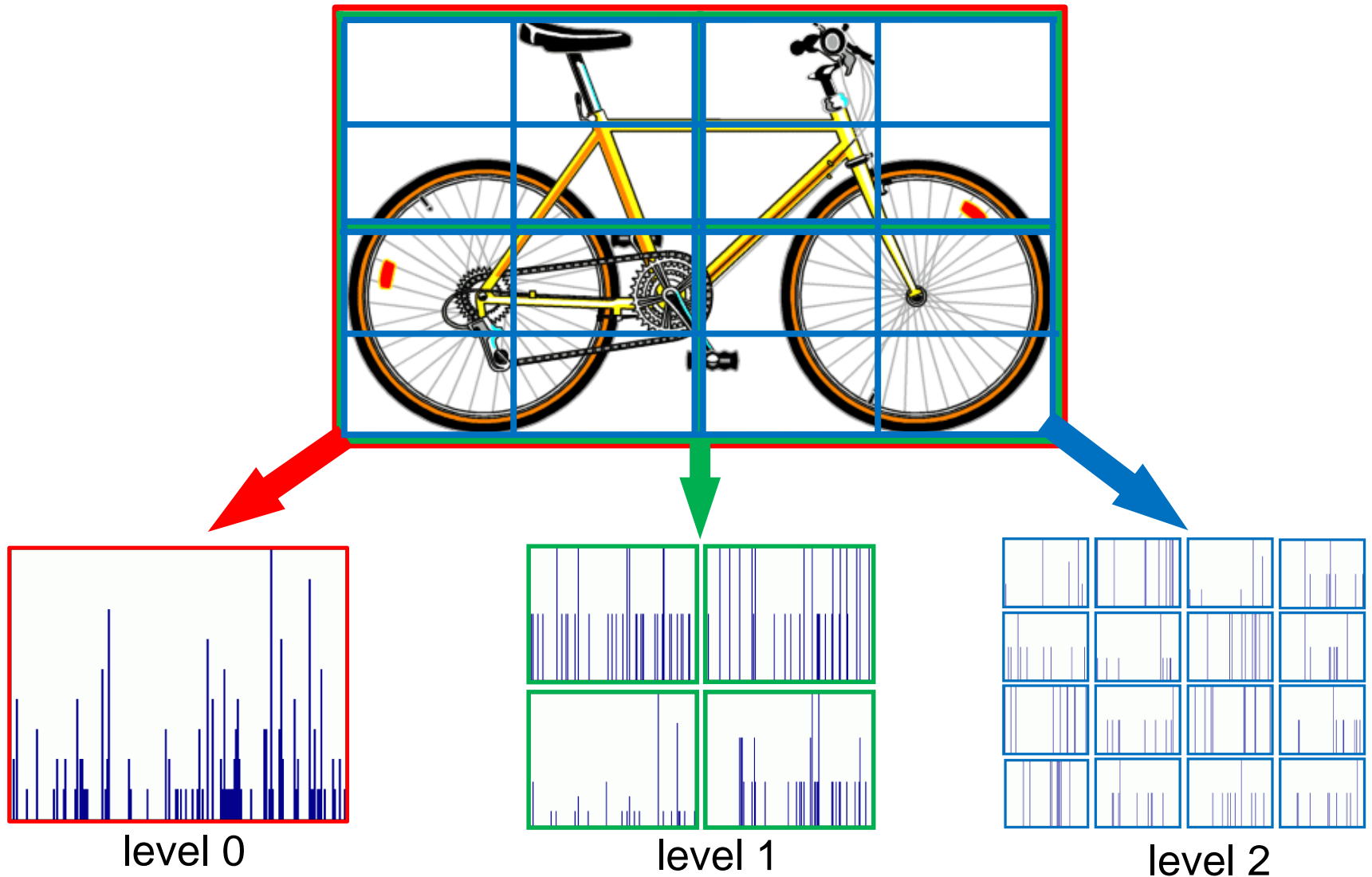


Image representation: spatial pyramids



Tutorial