



# Vision and Cognitive Services

SCP9087563 - LM CS,DS,CYB,PD,CE

Bag-of-visual-words for image classification



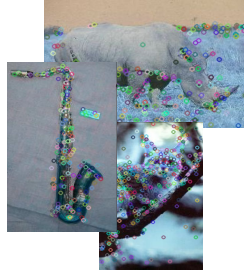
# Bag-of-visual-words for image classification

## Training



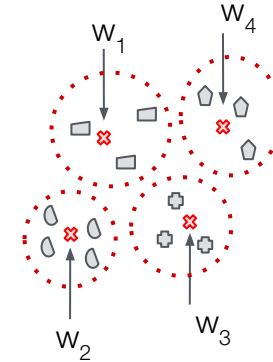
Training  
images

Step 1



Features extraction (e.g., SIFT)

Step 2



Vocabulary generation (K-Means)

Step 3



Image 1

Image 2

...

Image N

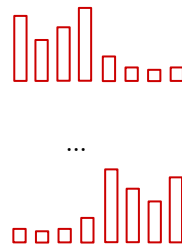
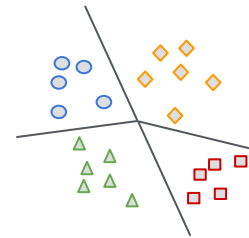


Image representation (frequency of words)

Step 4



Classifier training (SVM)

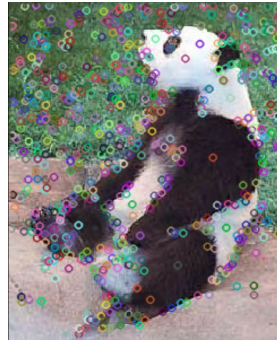
# Bag-of-visual-words for image classification

Testing



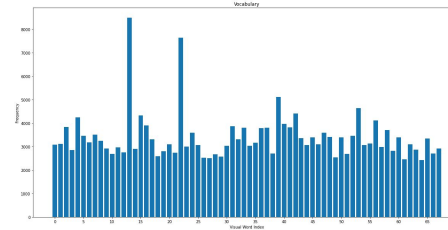
Test image

Step 1



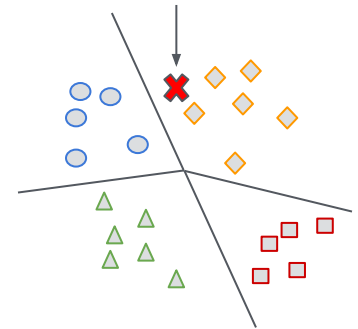
Features extraction (SIFT)

Step 2



Bag-of-visual-words

Step 3



Prediction

# Bag-of-visual-words for image classification

**Dataset:** *Caltech101* (trimmed version)

**Ver. 1** (*4\_ObjectCategories*): *airplanes, cars, faces, motorbikes*;

**Ver. 2** (*15\_ObjectCategories*): *bonsai, butterfly, crab, elephant, euphonium, faces, grand\_piano, joshua\_tree, leopards, lotus, motorbikes, schooner, stop\_sign, sunflower, watch.*

**Training set:** 30 images per category.

**Testing set:** 50 (or <50) images per category.

**Performance evaluation metrics:** accuracy, precision and recall.

