



VLAD

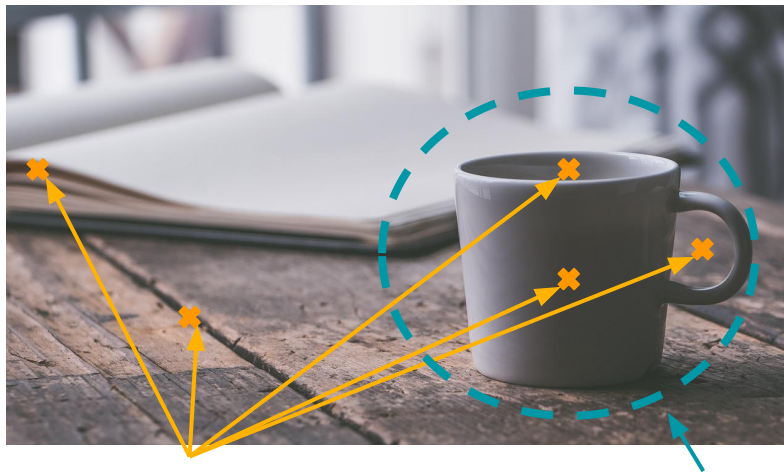
Récap

Description d'une image : exemple SIFT

SIFT(



) →





k_1 keypoints
taille D

keypoints de la tasse

SIFT repère des “points clés” dans l’image.
Chacun est décrit par un vecteur de taille d
Attention, nombre variable !

Description d'une image : exemple SIFT

SIFT() = tableau $[k_1, d]$

SIFT() = tableau $[k_2, d]$

Deux images décrites par SIFT, mais pas le même nombre de keypoints

Comment comparer les deux descripteurs??

La comparaison keypoint à keypoint est coûteuse en calcul...

Approche “Bag of Words” classique



L'approche BoW :
Un descripteur de taille constante,
qui décrit les éléments de l'image

- Liste de mots
- Pour chaque mot, on regarde s'il est présent dans l'image

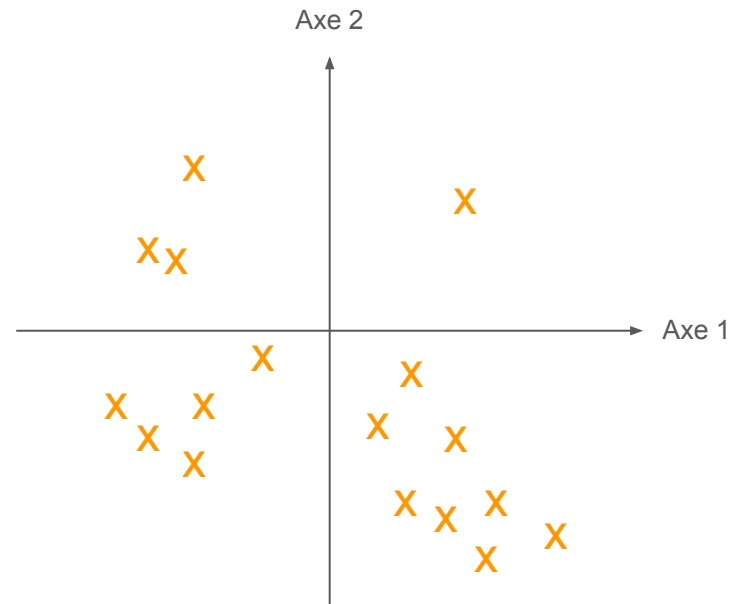
Descripteur BoW	
Objet ("word")	Nombre
Tasse	2
Chaise	3
Personne	0
Chien	0
Polytechnique Montréal le génie en première classe	0

Trouver les mots



On prend en vrac tous les keypoints
de tout un dataset

SIFT



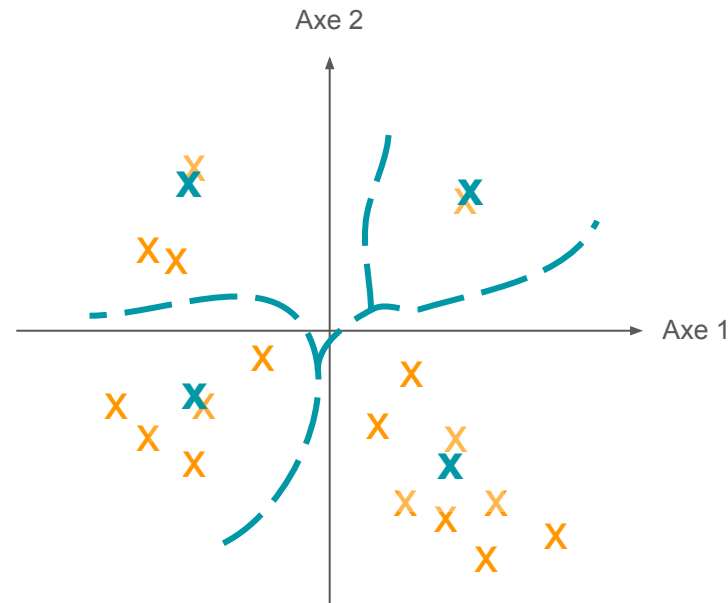
Représentation en 2D
(rappel : dimension d en pratique)

Trouver les mots



On prend en vrac tous les keypoints
de tout un dataset

SIFT



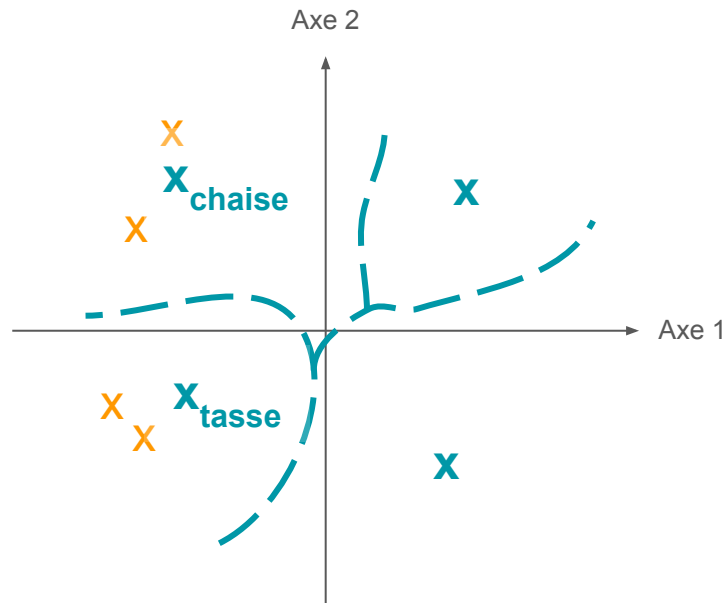
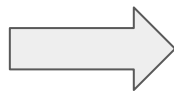
On détermine ensuite n clusters
Donc n centres de clusters = n mots
Avec k -means par exemple

Description d'une image

SIFT(



)



Pour décrire une image, on prend les keypoints via SIFT et on compare aux n centres / mots

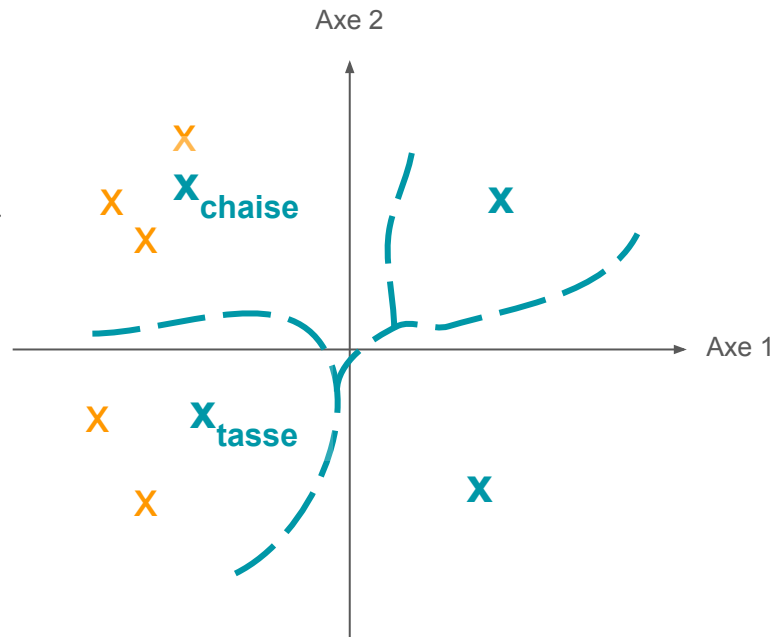
Remarque : en pratique pas de sens sémantique aussi fort que “tasse” ou “chaise”, c’est pour l’exemple !

Description d'une image : BoW naïf

SIFT(



)




Descripteur BoW Classique	
n Mots	Résultante
Tasse	2
Chaise	3
...	...

Avec le BoW classique, on compte juste le nombre de keypoints qui tombent dans tel ou tel cluster

Description d'une image : Méthode VLAD

Nuance avec VLAD : pour chaque cluster,

1. On regarde quels keypoints sont dedans
2. Pour chacun d'entre eux, on calcule la différence vectorielle
3. On somme ces différences

Descripteur BoW VLAD Taille $n * d$	
n Mots	Résultante
Tasse	
...	...

