

Shallow representation - From Shallow to Deep representation for multimedia data - Lecture 4 : From BOVW to VLAD and Fisher Vector

Johnny Nguyen¹

Abstract—This electronic document permits me to synthesis the fourth course of analysis and indexation.

I. INTRODUCTION

Dans cette partie, nous parlerons du Bag of Visual Word, du VLAD et du Fisher Vector.

II. BAG OF WORDS

Il s'effectue en deux tapes :

- Projecteurs des descripteurs locaux : vocabulaire,
- Statistique sur les descripteurs : mot visuel.

Les dsavantages sont :

- Impossibilit d'avoir des donnes d'entrainement,
- Beaucoup de perte d'informations.

L'utilisation du K-means est fortement conseill. Il consiste compter le nombre de descripteur par rgion de voronoi.

III. LE VLAD

Le but est d'obtenir la direction de chaque points d'un cluster et de les ajouter entre elles.

IV. FISHER VECTOR

Rcupre le plus d'informations avec les statistiques en utilisant le gradient.

V. CONCLUSION

Le BOW est infrieur au VLAD qui est infrieur au Fisher Vector. La moyenne correspond au VLAD et la variance au Fisher Vector. C'est une mthode d'aggrgation.

ACKNOWLEDGMENT

Thanks to Frederic Precioso for his work.

REFERENCES

- [1] <https://moodle.polytech.unice.fr/course/view.php?id=31>