Shallow representation - From Shallow to Deep representation for multimedia data - Lecture 3 : BoW : coding pooling and spatial pyramid

Johnny Nguyen[1]

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

I. Introduction

The goal of this course is to solve the problem with local information. We will retrieve the same object or category. Amazon Go Seattle search with automatic inventory. We have reference image and we want to retrieve the same category. We will focus to be light on the memory.

II. Bag of Words

We can play with the semantic matching. We can replace all category of for example *bank* or *trader* in a text by the word *economic*. We are going to do the same process for images to understand which category the image is.

- *Stemming* : variation of the same word,
- *Lemmatization* : word with the same family,
- *Vector space model* : represent words in a space,
- *Distance of kh2* : normalize by the number of words on document.

In the case of an image descriptor point it is the same as a word.

III. Spatial pyramid

We want to extract the best point on a descriptor. This method doesn't use every point on descriptor but only the best, it corrects the error on an image.

ACKNOWLEDGMENT

Thanks to Frederic Precioso for his work.

References

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

*This work was not supported by any organization