VISUAL RECOGNITION - **WEEK 4**

Eric López, Gerard Martí, Sergio Sancho, Adriana Fernández

Master in Computer Vision, Barcelona
UAB, UOC, UPC, UPF

March 2015

Previous week follow-up

Classifier	Accuracy				
	No-SPM	SPM 2x2	SPM 3x1	SPM 2x2 + Int. Fusion	SPM 3x1 + Int. Fusion
LinearSVM	0.8289	0.8327	0.8413	0.8512	0.8587
HISVM	0.8550	0.8674	0.8580	0.8909	0.8599
SPMKernel SVM	-	0.8624	0.8537	0.8674	0.8636

As we mentioned in our previous slides, we did not have time to implement another type of spatial pyramid last week, so we did it during this week. Compared to our best results obtained in the previous week (with FAST+SIFT+ColorNaming, weights 0.35-0.65), we do not obtain a better performance than when working with our previous settings.

Contents

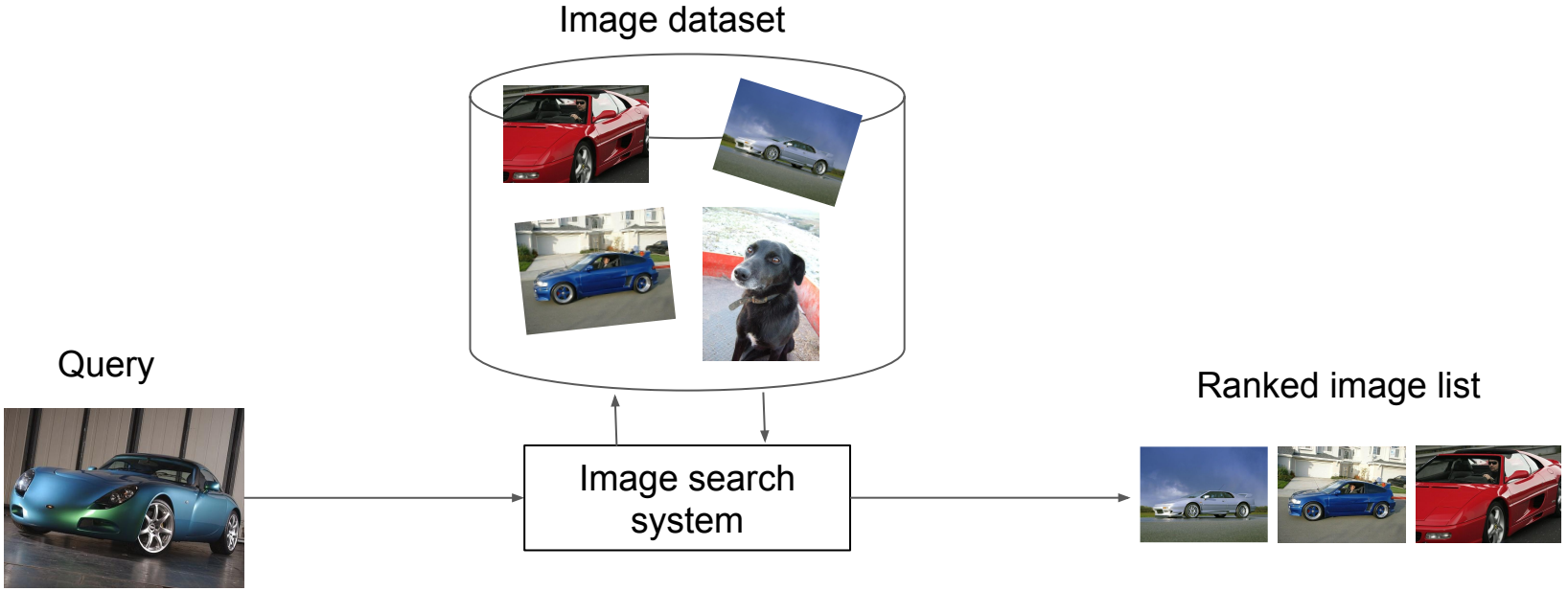
Introduction: Image retrieval

Query methods

Optional tasks

Visual results and Conclusions

Introduction: Image retrieval



Objective: find the most similar images to the query image.

Introduction: Image retrieval

Query methods

Optional tasks

Visual results and Conclusions

Query methods

In the table we can see the results obtained for VLAD, Fisher and BoW using Fast NN or using product quantization. Always using FAST+SIFT and a vocabulary size of K = 1000 (except for fisher, which uses GMM and we use K = 64). We use this combination of detector+descriptor because is the combination that provided us with better results in the previous weeks. For product Quantization, we use adc as the distance function, as in our testing it has shown on par results with sdc distance but much better performance

Performance (Accuracy)	Fast NN			Product Q.		
	Vlad	Fisher	BoW	Vlad	Fisher	BoW
Top 1	0.69	0.78	0.67	0.66	0.51	0.52
Top 5	0.74	0.82	0.74	0.67	0.55	0.56
Top 10	0.72	0.80	0.72	0.66	0.55	0.57
Top 15	0.70	0.79	0.71	0.61	0.56	0.57
Top 20	0.69	0.78	0.70	0.60	0.56	0.56
Top 25	0.69	0.77	0.69	0.60	0.56	0.56

Introduction: Image retrieval

Query methods

Optional tasks

Visual results and Conclusions

Optional tasks - Using LSH

We tried to apply LSH to the product quantization method and to the different global descriptors and see if there is a difference in performance. Using the same parameters as before, we obtain the following results. In our experiments, LSH did not offer a gain in performance at all, it decreased our accuracy.

Performance (Accuracy)	Product Q.			Product Q. + LSH		
	Vlad	Fisher	BoW	Vlad	Fisher	BoW
Top 1	0.66	0.51	0.52	-	0.51	0.49
Top 5	0.67	0.55	0.56	-	0.53	0.52
Top 10	0.66	0.55	0.57	-	0.53	0.54
Top 15	0.61	0.56	0.57	-	0.54	0.55
Top 20	0.60	0.56	0.56	-	0.54	0.55
Top 25	0.60	0.56	0.56	-	0.54	0.54

Optional tasks - Using Inverted Files

We then tried to apply Inverted Files to Product quantization. In theory we expected to observe a major improvement in computation time, and that is what happens: we see that the performance per image increases a lot while doing inverted files, and the accuracy also seems to improve marginally.

Performance (Accuracy)	Product Q.			Product Q. + InvFile		
	Vlad	Fisher	BoW	Vlad	Fisher	BoW
Top 1	0.66	0.51	0.52	-	0.48	0.55
Top 5	0.67	0.55	0.56	-	0.56	0.59
Top 10	0.66	0.55	0.57	-	0.56	0.59
Top 15	0.61	0.56	0.57	-	0.56	0.59
Top 20	0.60	0.56	0.56	-	0.56	0.59
Top 25	0.60	0.56	0.56	-	0.55	0.58
Performance per image (s)	3.67 s	2.14s	3.37 s	-	1.62	1.42 s

Introduction: Image retrieval

Query methods

Optional tasks

Visual Results and Conclusions

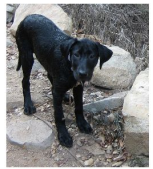
Visual results

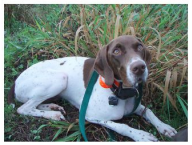
In the following slides we will show some of the queries obtained. These images were retrieved using Fisher + FastNN, our best configuration. As we can see we have chosen a Top-10 performance, to show our results.

QUERY










This query is really interesting: there are several dogs, and one of the dogs has really fluffy hair.

The system give us dogs which has a similar fluffy hair, but also dogs near patches of grass that have a similar texture.

Visual results

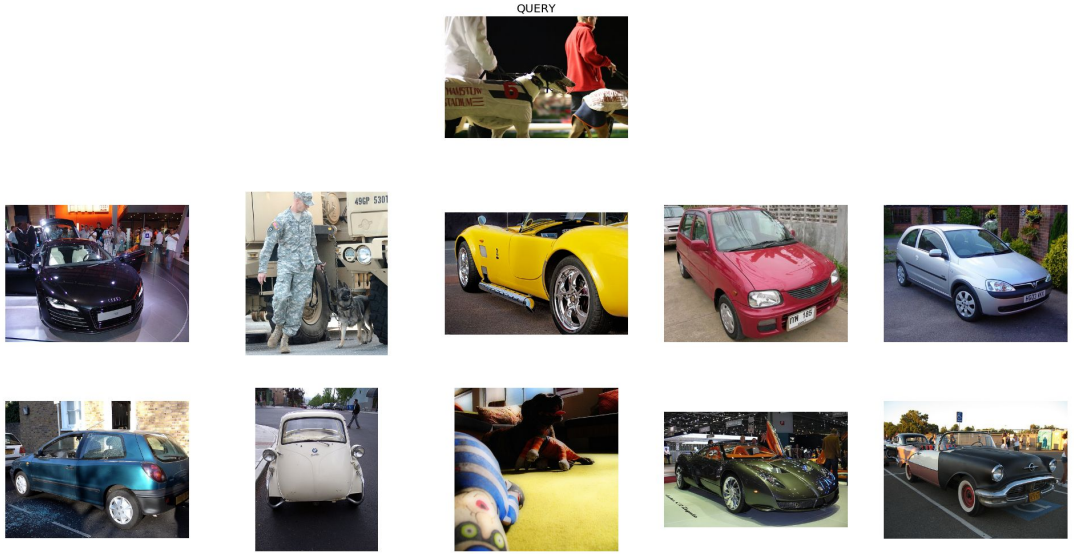
In the following slides we will show some of the queries obtained. These images were retrieved using Fisher + FastNN, our best configuration. As we can see we have chosen a Top-10 performance, to show our results.

This query is a closeup of a face of a dog.

Most of the images obtained are also of a closeup of a dog face, where the nose, eyes and ears are visible.

Visual results

In the following slides we will show some of the queries obtained. These images were retrieved using Fisher + FastNN, our best configuration. As we can see we have chosen a Top-10 performance, to show our results.



This query is an image of a dog dressed up, prepared for running in a competition.

The piece of cloth that the dog is wearing modifies the dog shape, and it seems that the classifier is confused by this and gives us cars in return. Also, the dog's perspective in that picture is not common.

Visual results

In the following slides we will show some of the queries obtained. These images were retrieved using Fisher + FastNN, our best configuration. As we can see we have chosen a Top-10 performance, to show our results.

QUERY










This query is an image of sports race car, which is being driven.

Almost all of the images returned are of sport cars (which have all the extra details and numbers in the side) and which are in motion. It seems that the system recognizes quite well not only the shape of the car, but also the details on the side.

Conclusions

- Due to the many errors in the code that was provided, we spend a great amount of time debugging it and we were not able to do as many experiments as we would have liked to do, in order to test a wider range of parameters and combinations. This may have hindered our results.
- Also, some of the experiments took a lot of time and memory (mainly Fisher), and the cluster most of the times did not work properly, with processes disappearing and programs not delivering any output at all. We have not been able to find the proper RAM memory to make it work.
- Our best combination is: **Fisher + FastNN, using FAST+SIFT and K = 64,** with accuracy: **Top5 = 0.82**, **Top10 = 0.80** and **Top25 = 0.77**.
- We expected an improvement applying product quantization, but as we observed, the results show that we are not able to achieve a better performance by using this method.
- Comparing product quantization, product quantization + LSH and product quantization + inverted files, the best results are obtained when inverted files are used, both accuracy and performance wise, but they are still very underwhelming. Fast NN seems to provide the best results.
- When we applied product quantization + LSH using VLAD we had huge number of execution errors and issues, and because of this we were not able to test with that specific configuration.

VISUAL RECOGNITION - **WEEK 4**

Eric López, Gerard Martí, Sergio Sancho, Adriana Fernández

Master in Computer Vision, Barcelona
UAB, UOC, UPC, UPF

March 2015