

# Supplementary Material

BMVC 2015 – Experimental Evaluation of the Bag-of-Features Model for Unsupervised Learning of Images

## 1. Overall Test Setting

The test setting for this evaluation study followed a number of different steps. In all of these steps, algorithms or parameters can be varied. Table 1 presents all the algorithms tested and how to set these algorithms and parameters in the python script. The program<sup>1</sup> starts when the main.py script is called (in the bash command line). An example of a call for this script could be:

```
python main.py -p /Users/general/ImageDatabases/Coil-20-proc40 -k FAST -n
109 -s False -d FREAK -g SAMPLEP:0.05 -c KMEANS -m euclidean -t 500 -f NONE
-h NONE -a KMEANS -r 10
```

**Table 1 - Steps of the BoF model for image clustering including algorithms and steps.**

Stage							Small option	Long Option
<b>Dataset path</b>							-p	--path
<b>1.1) Keypoint Detector</b>	SIFT (DoG)	SURF (Harris)	STAR	FAST	ORB	RANDOM	-k	--keypnt
<b>1.2) Number of patches per image</b>	Dependant on the dataset						-n	--numpatch
<b>1.3) Same number of patches or different</b>	Same (True)	Different (False)					-s	--equalnum
<b>2) Descriptors</b>	SIFT	SURF	BRIEF	ORB	FREAK		-d	--imdes
<b>3) Image and Feature Sampling</b>	None	SAMPLEP	SAMPLEI				-g	--imsample
<b>4.1) Codebook construction algorithm</b>	K-Means	Mini Batch K-Means	Birch	RANDOMV	RANDOM		-c	--codebook
<b>4.2) Size of codebook</b>	Dependant on the dataset						-t	--size
<b>5) Feature selection</b>	Remove frequent features	Remove rare features	Remove frequent and rare features	None			-f	--fselec
<b>6) Histogram Normalization</b>	None	Tf-idf	Tf normalized	Td-idf variant	Tf-idf normalized	Binarization	-h	--histnorm
<b>7.1) Clustering algorithm</b>	DBSCAN	K-Means	Birch	HIERAR1 (Scipy)	HIERAR2 (Scikit-learn)		-a	--clust
<b>Repetitions</b>							-r	--rep

<sup>1</sup> <https://github.com/marianaAfonso/BOFClustering>

## 2. Steps of the Experimental Design

### 2.1 Detectors and Descriptors

For the detectors and the descriptors evaluation, fixed values for the codebook size and the proportion of keypoints used for codebook learning were chosen and can be found in Table 2.

**Table 2 - Feature selection results for Coil-20 Dataset**

Dataset	Codebook Size	Proportion of Keypoints used for Codebook learning
<b>Coil-20</b>	110	0.3
<b>Natural and Urban</b>	300	0.05
<b>Event</b>	500	0.05

#### 2.1.1 Detectors

##### 2.1.1.1 SIFT

The SIFT descriptor is used with different parameters for the three datasets. Table 3 shows the different parameter settings:

**Table 3 - Feature selection results for Coil-20 Dataset**

Dataset	Contrast Threshold	Edge Threshold
<b>Coil-20</b>	0.01	4.5
<b>Natural and Urban</b>	0.01	10
<b>Event</b>	0.01	5

##### 2.1.1.2 SURF

The SURF descriptor is also used with different parameters for the three datasets. Table 4 shows the different parameter settings:

**Table 4 - Feature selection results for Coil-20 Dataset**

Dataset	Hessian Threshold
<b>Coil-20</b>	450
<b>Natural and Urban</b>	50
<b>Event</b>	80

##### 2.1.1.3 STAR

The STAR descriptor is used with the Threshold = 0 for the three datasets.

#### 2.1.4 FAST

The FAST descriptor is also used with different parameters for the three datasets. Table 5 shows the different parameter settings:

**Table 5 - Feature selection results for Coil-20 Dataset**

Datasets	Threshold
<b>Coil-20</b>	21
<b>Natural and Urban</b>	17
<b>Event</b>	35

#### 2.1.5 ORB

The ORB descriptor is also used with different parameters for the three datasets. Table 6 shows the different parameter settings:

**Table 6 - Feature selection results for Coil-20 Dataset**

Datasets	Max Features	Patch size
<b>Coil-20</b>	50	15
<b>Natural and Urban</b>	500	15
<b>Event</b>	1000	11

#### 2.1.6 RANDOM

The RANDOM detector uses the DENSE detector from the OpenCV library with  $\text{initXyStep} = 5$  pixels.

### 2.1.2 Descriptors

Apart from the BRIEF descriptor that was used with the number of bytes =32, all the other descriptors were used with their Default parameters.

## 2.2 Image and Feature sampling

The method SAMPLEP is a general random sampling of all the features obtained from the images. The parameter used for this sampling method is the percentage of features that will be used to create the codebook of the BoF model.

Alternatively, the SAMPLEI method, which was proposed in this work, intends to minimize the influence of the different number of keypoints per image for the construction of the codebook. The way it does this task is by selecting fewer features from the images that have more than the average number of features. The percentage of features that is selected follows the curve presented in Figure 1.

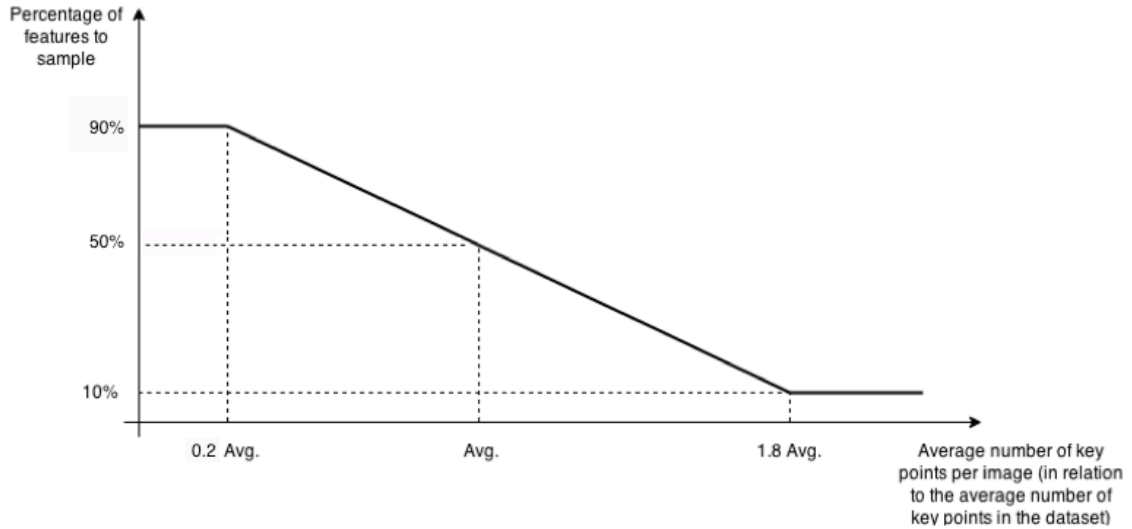


Figure 1 - SAMPLEI sampling curve for the features of the images

## 2.3 Feature Selection

Tables 7 to 9 present the results for the feature selection applied to the visual words of the codebook for each dataset.

The tables show the minimum frequency that a visual word appears, the maximum frequency that a visual word appears, and the average NMI score for that configuration.

Table 7 - Feature selection results for Coil-20 Dataset

Fmin	Fmax	Avg. NMI
0	0,65	<b>0,80</b>
0	0,7	<b>0,82</b>
0	0,75	<b>0,82</b>
0	1	<b>0,80</b>
0,25	0,65	<b>0,80</b>
0,25	0,7	<b>0,81</b>
0,25	0,75	<b>0,82</b>
0,25	1	<b>0,81</b>
0,3	0,65	<b>0,80</b>
0,3	0,7	<b>0,80</b>
0,3	0,75	<b>0,81</b>
0,3	1	<b>0,81</b>
0,35	0,65	<b>0,79</b>
0,35	0,7	<b>0,79</b>
0,35	0,75	<b>0,80</b>
0,35	1	<b>0,78</b>
0,4	0,65	<b>0,77</b>
0,4	0,7	<b>0,77</b>
0,4	0,75	<b>0,79</b>
0,4	1	<b>0,78</b>

**Table 8 - Feature selection results for Natural and Urban dataset**

Fmin	Fmax	Avg NMI
0	0,8	<b>0,24</b>
0	0,85	<b>0,25</b>
0	0,9	<b>0,26</b>
0	0,95	<b>0,27</b>
0	0,995	<b>0,27</b>
0,01	0,8	<b>0,23</b>
0,01	0,9	<b>0,27</b>
0,01	0,995	<b>0,27</b>
0,01	1	<b>0,27</b>
0,03	0,8	<b>0,23</b>
0,03	1	<b>0,28</b>
0,05	0,8	<b>0,23</b>
0,05	1	<b>0,28</b>
0,08	1	<b>0,28</b>
0,1	0,9	<b>0,26</b>
0,1	0,995	<b>0,27</b>
0,1	1	<b>0,28</b>
0,15	1	<b>0,27</b>
0,2	0,9	<b>0,26</b>
0,2	0,995	<b>0,28</b>
0,2	1	<b>0,27</b>

**Table 9 - Feature selection results for the Event dataset**

Fmin	Fmax	Avg. NMI
0	0.8	<b>0,42</b>
0	0.9	<b>0,42</b>
0	0.95	<b>0,41</b>
0	0.995	<b>0,42</b>
0.05	1	<b>0,42</b>
0.05	0.8	<b>0,42</b>
0.05	0.9	<b>0,42</b>
0.05	0.95	<b>0,43</b>
0.05	0.99	<b>0,42</b>
0.05	0.995	<b>0,42</b>
0.1	1	<b>0,42</b>
0.1	0.8	<b>0,41</b>
0.1	0.9	<b>0,42</b>
0.1	0.95	<b>0,42</b>
0.1	0.995	<b>0,42</b>
0.15	1	<b>0,42</b>
0.15	0.8	<b>0,42</b>
0.15	0.9	<b>0,42</b>
0.15	0.95	<b>0,42</b>
0.15	0.995	<b>0,42</b>
0.2	1	<b>0,42</b>
0.2	0.8	<b>0,41</b>

Mariana Afonso, [marianafza@fe.up.pt](mailto:marianafza@fe.up.pt), Luis Teixeira, [luisft@fe.up.pt](mailto:luisft@fe.up.pt)

0.2	0.9	<b>0,42</b>
0.2	0.95	<b>0,42</b>
0.2	0.995	<b>0,42</b>