Image understanding

Assignment 2

Luigi Quarantiello

Code snippets

```python
#extract keypoints and descriptors for each image
def extract_sift_descriptor(set):
    kp = {}
    des = {}
    sift = cv2.xfeatures2d.SIFT_create()
    for key,img in set.items():
        kp[key], des[key] = sift.detectAndCompute(img, None)

    return kp, des
```

```python
#learn the 500-dimensional codebook using k-means clustering
def k_means(samples, k = 500, att = 20):
    kmeans = KMeans(k, verbose = 1, n_init = att)
    kmeans.fit(samples)

    return kmeans
```

```python
#build histogram using the codebook
def build_histogram(descriptors, kmeans):
    hist = np.zeros(shape=len(kmeans.cluster_centers_))
    predictions = kmeans.predict(descriptors)
    for i in predictions:
        hist[i] += 1

    return hist, predictions
```

Code snippets

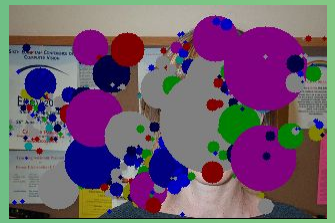
```
#Bag of Words representation
#list of lists - each list contains (word_id, word_freq) pairs for each visual word from the histograms
def format_corpus(set):
    total_bow = []
    for doc in set:
        single_doc_bow = []
        for id, freq in  enumerate(doc): single_doc_bow.append((id, freq))
        total_bow.append(single_doc_bow)

    return total_bow
```

```
#define the LDA model, using the number of different objects in the dataset as number of topics
lda = LdaModel(train_bow, num_topics = 9, alpha = 'auto', eta = 'auto', iterations = 1000, passes = 1000)
```

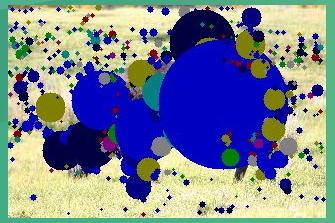
```
#get the most relevant topic for each word
def topic_per_word(set_clustered):
    doc_topics = {}
    for key, img in set_clustered.items():
        doc_topics[key] = []
        for id, word in enumerate(img):
            r = lda.get_term_topics(word, minimum_probability = 0)
            max_topic = max(r, key = lambda k: k[1])
            doc_topics[key].append(max_topic[0])

    return doc_topics
```

Results

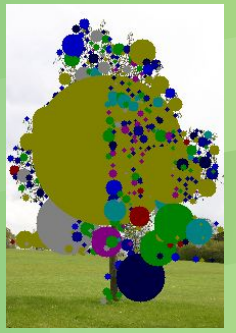


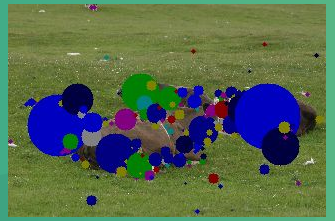





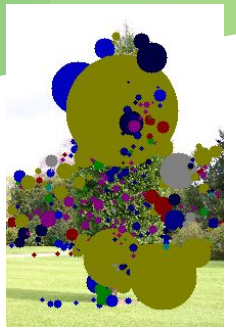












Final considerations

- Quite simple model to implement
- It requires few hyperparameters, in particular the number of topics and the α and β parameters

- It has good results on some of the subsets of images, while it seems not to understand the topic for other ones