

HW3 Q3、Q4解說

2016年12月20日 下午 03:27

Q3. SIFT Visual Words

要完成這項作業大概分成幾個步驟：









步驟一、將每張圖像的**SIFT FEATURE**抓取出來

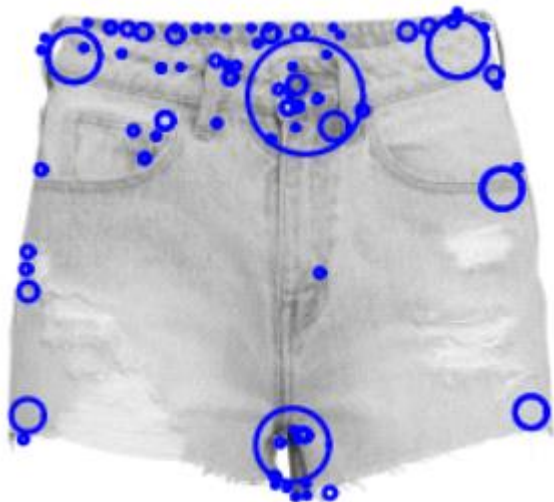
作法：

先到 <http://www.vlfeat.org/> 安裝好VLfeat並設定好環境變數

利用給同學的sift.py及SIFT_Example.py將每張圖片的Sift找出來

會產生下圖所示檔案

 ukbench00003.sift	2016/12/19 下午 ...	SIFT 檔案	14 KB
 ukbench00004.sift	2016/12/19 下午 ...	SIFT 檔案	26 KB
 ukbench00005.sift	2016/12/19 下午 ...	SIFT 檔案	14 KB
 ukbench00006.sift	2016/12/19 下午 ...	SIFT 檔案	10 KB
 ukbench00007.sift	2016/12/19 下午 ...	SIFT 檔案	12 KB
 ukbench00008.sift	2016/12/19 下午 ...	SIFT 檔案	7 KB
 ukbench00009.sift	2016/12/19 下午 ...	SIFT 檔案	18 KB
 ukbench00010.sift	2016/12/19 下午 ...	SIFT 檔案	102 KB
 ukbench00011.sift	2016/12/19 下午 ...	SIFT 檔案	132 KB
 ukbench00012.sift	2016/12/19 下午 ...	SIFT 檔案	5 KB



1	55.2689	17.2882	1.15353	4.22277	0	1	28	45	7	1	4	1	85	8	8	24	7	0	0	8	202
	6	0	0	0	0	0	40	49	5	0	0	0	0	4	5	0	1	6	12	12	21
	202	24	0	0	0	0	6	60	9	0	0	0	0	0	4	1	1	0	2	8	89
	32	202	12	0	0	0	1	27	52	1	0	0	0	0	1	7	0	24	64	2	0
	0	16	21	202	95	12	0	0	0	0	8	33	6	0	0	0	0	0	0	3	
2	141.512	31.366	0.983269	3.49967	1	0	0	15	140	5	0	0	140	18	0	0	81	6	0	1	
	116	5	0	0	4	7	9	0	0	0	0	5	30	8	0	3	1	0	9	140	8
	90	14	0	0	3	22	13	7	0	1	0	0	3	45	13	0	4	0	0	5	140
	4	61	56	2	0	1	2	29	26	24	8	50	1	0	0	23	13	0	3	0	0
	95	9	0	30	68	24	1	1	1	0	1	16	34	140	2	1	0	0	0	0	
3	49.5254	38.6691	0.982662	3.26293	0	0	4	39	90	0	0	0	51	2	0	35	160	1	0	1	
	160	11	0	2	14	0	0	10	51	6	1	0	0	0	1	12	0	0	0	22	92
	2	0	0	160	9	1	8	20	0	2	25	36	0	0	0	1	20	50	0	0	7
	160	2	0	0	160	16	1	11	27	0	1	16	31	2	2	0	0	0	12	23	1
	1	27	160	1	0	0	160	10	1	6	24	0	0	8	25	2	0	0	1	7	8
4	49.5254	38.6691	0.982662	6.23662	0	3	8	11	13	3	1	0	6	2	2	1	140	37	0	0	
	155	20	0	0	37	15	1	6	87	18	0	0	1	0	0	1	0	0	9	39	24
	2	3	155	25	0	0	30	18	2	12	86	19	2	0	0	0	3	9	0	0	6
	155	74	1	4	155	26	0	0	24	11	1	9	69	19	13	2	0	0	1	5	0
	7	0	0	155	70	0	1	155	29	0	0	11	3	0	8	60	2	1	1	1	0
5	169.046	55.9391	0.996793	2.99441	9	1	0	21	164	0	0	0	164	9	0	4	34	0	0	0	
	95	4	0	1	3	2	1	2	14	7	0	1	5	0	0	1	12	0	0	5	164
	9	101	15	0	4	9	8	2	6	21	3	0	0	7	4	3	9	16	0	0	13
	0	17	89	6	1	4	9	16	3	15	10	7	1	1	5	5	5	16	12	0	0
	38	1	0	14	72	12	5	10	2	0	0	4	9	13	2	1	6	0	0	0	
6	16.9892	75.5521	0.942717	3.56352	5	1	0	0	0	5	23	10	138	12	0	0	0	2	4	10	
	206	25	0	0	0	0	9	51	5	2	0	0	0	1	7	23	11	2	0	0	3
	40	206	6	0	0	0	0	45	76	1	0	0	0	0	0	2	5	28	11	6	1

而每一個sift檔都會標示它的visual word，一行就代表一個圈圈。

步驟二、將所有的圈圈（Word）一起下去分群（使用K-means）

這裡要注意的，是全部1006張的圖片的圈圈（blob）一起做分群，不是個別圖片去做分群。而分成幾群大家可以自己實驗看看怎樣的效果比較好。這樣每一個Word都可以被分到其中一個群。

步驟三、回到每一張圖，替virtual word標記對應的編號群Label

意思就是：某圖有Virtual Word：a0 b0 c0 a1 a2 b2 b3

而分群將a0 a1 a2分到了A群、將b0 b2 b3分到了B群、將c0分到了C群

所以這張圖片就會變成 A B C A A B B

步驟四、建立histogram、計算相似度

上述僅提到一張圖，若N張圖片共有A~G群

就可以將某圖表示成：

A B C D E F G

3 3 1 0 0 0 0

這就是這張圖的feature了，即可拿來與其他圖片進行相似度計算。

Q4. SIFT Visual Words using stop words

stop words意思就是每張圖片都有同樣的特徵。用文本來解釋，就是不具代表性的詞彙，比如說每個文章都會有「你、我、他、的、如果、或許、那個……」，在分析的時候把這些詞彙拿掉，文本留下的特徵就會比較有代表性。

因此在SIFT Visual Words也是一樣，如果全部的圖片中，每張圖片都有類似

或相同的blob，那麼就要把它忽略，這樣留下來的Virtual Word才會有代表性。

具體來說，上述提到已建立好的histogram，如果發現幾乎每一張圖都有A群，那就可以合理推斷A是Stop Word。本作業希望同學把前10%的the most common [top 10%] visual words就是同樣的意思。