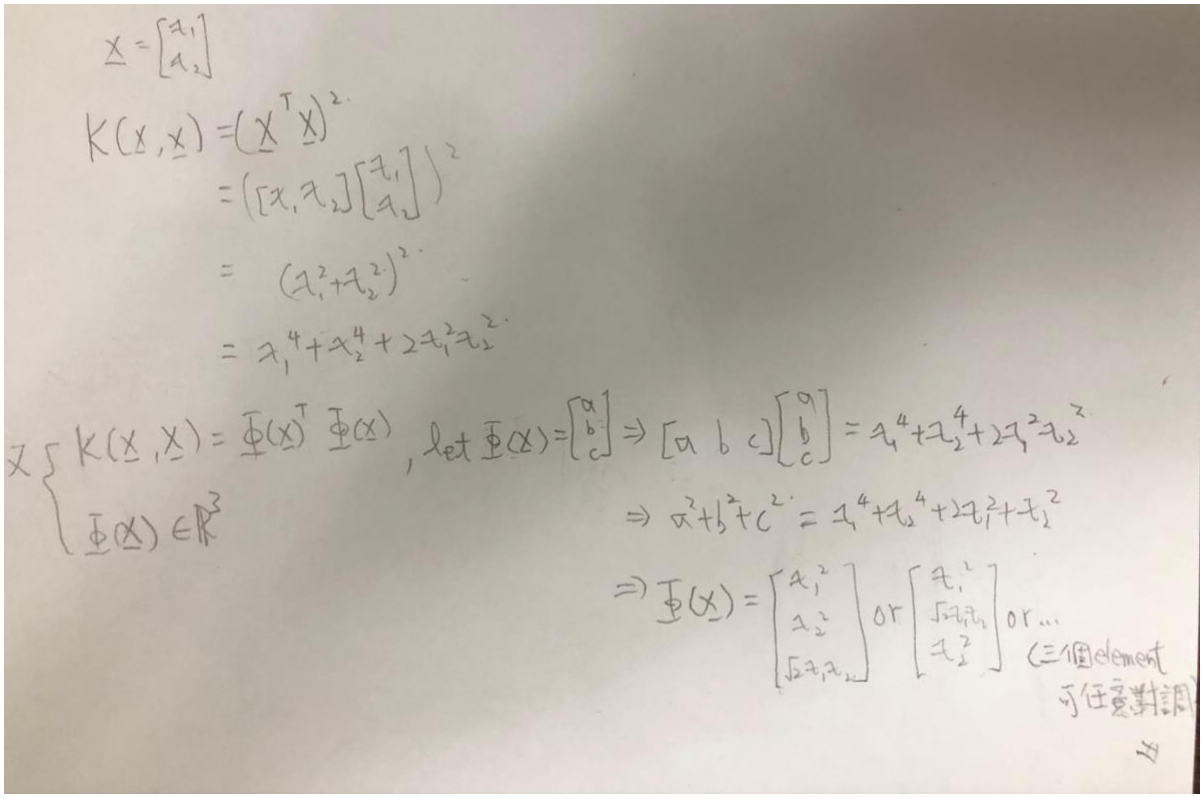
Deep Learning for Computer Vision
Homework 2

R06942065 電信一 黃宇平

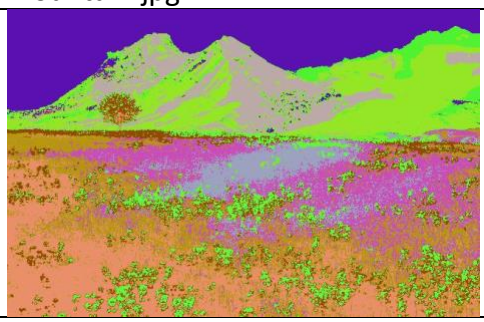
Problem 1: Kernel Trick



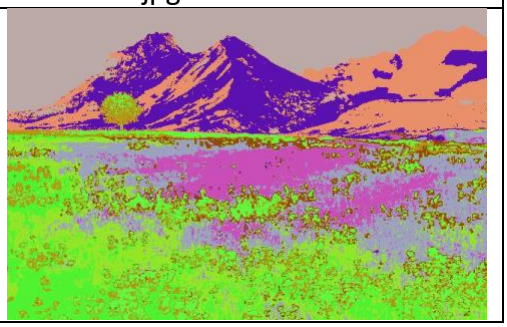
Problem 2: Color and Texture Segmentation

(a) Color Segmentation

(i) RGB color space

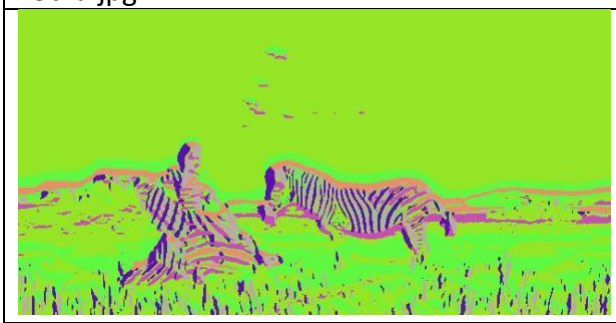
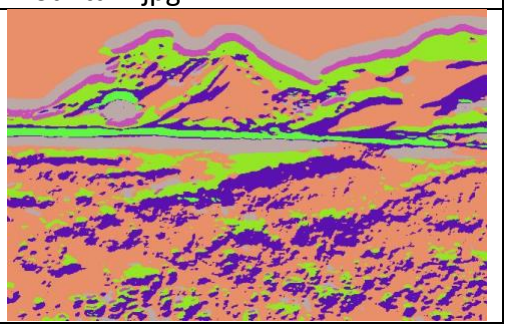
zebra.jpg	mountain.jpg
	

(ii) LAB color space

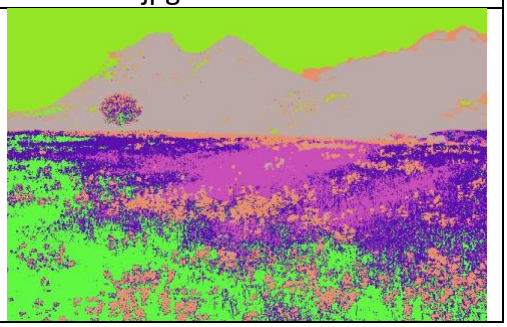
zebra.jpg	mountain.jpg
	

(b) Texture Segmentation

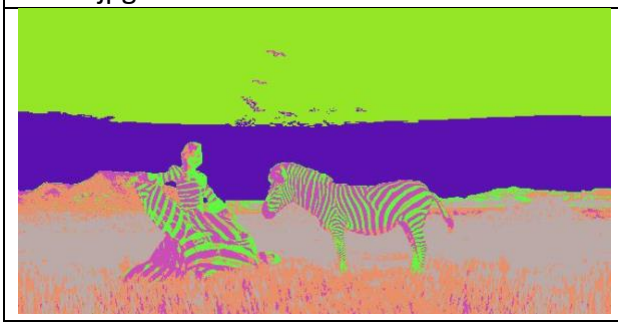
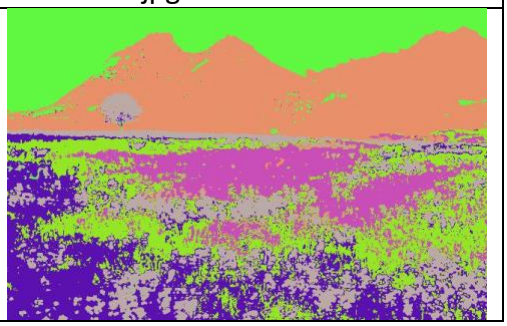
(i) Only texture features

zebra.jpg	mountain.jpg
	

(ii - 01) With RGB color features

zebra.jpg	mountain.jpg
	

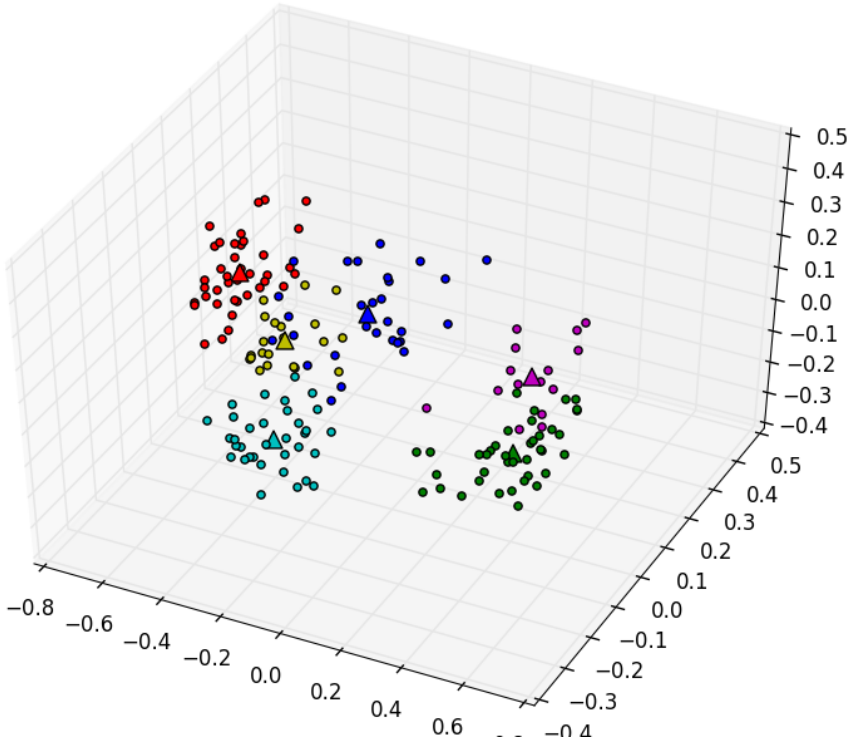
(ii - 02) With LAB color features

zebra.jpg	mountain.jpg
	

Problem 3: Recognition with Bag of Visual Words

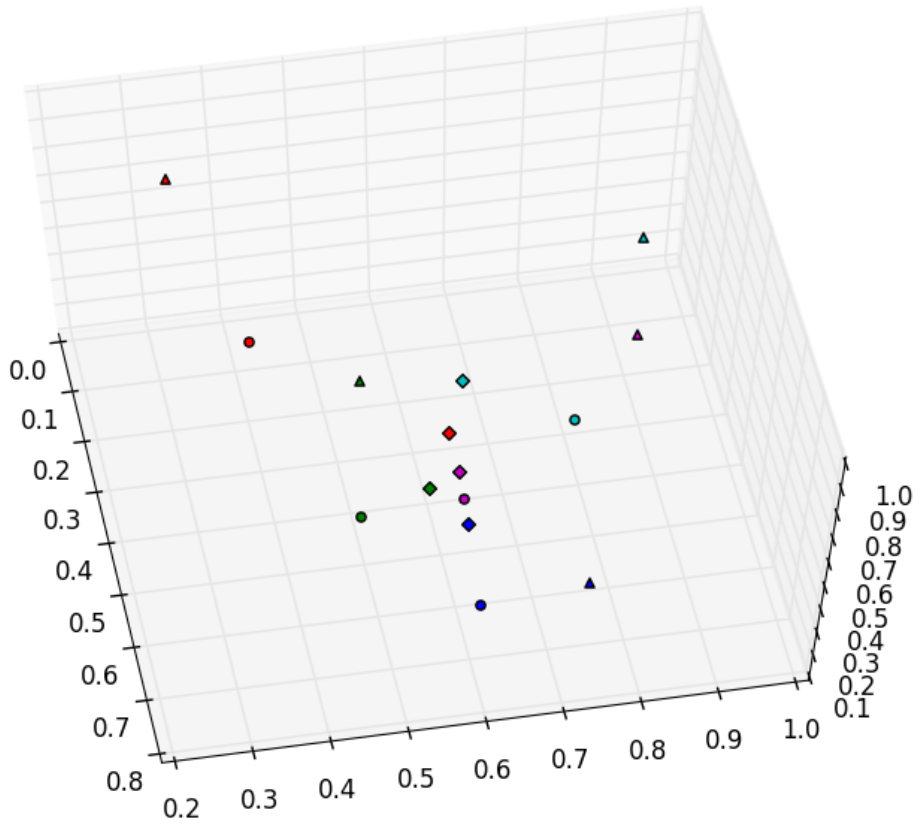
(a) Detect Interesting Points Using SURF

(b) Plot of Visual-Words in PCA Subspace



(其中較大的三角形 Cluster centroid(即 Visual-Word), 同色的小圓點為相同 Cluster 中的 interest points)

(c) BoW Plot



(其中相同顏色的 marker 表示同一張圖，三角形為 Hard-Sum BoW、圓形為 Soft-Sum Bow、菱形為 Soft-Max Bow)
由上圖，我發現不同張圖片的 Hard-Sum Bow 距離較遠，因此在這個 case 上，我猜測 Hard-Sum Bow 較有機會正確地區分不同的圖片。

(d) k-NN on BoW Features
(i) 分類結果為：

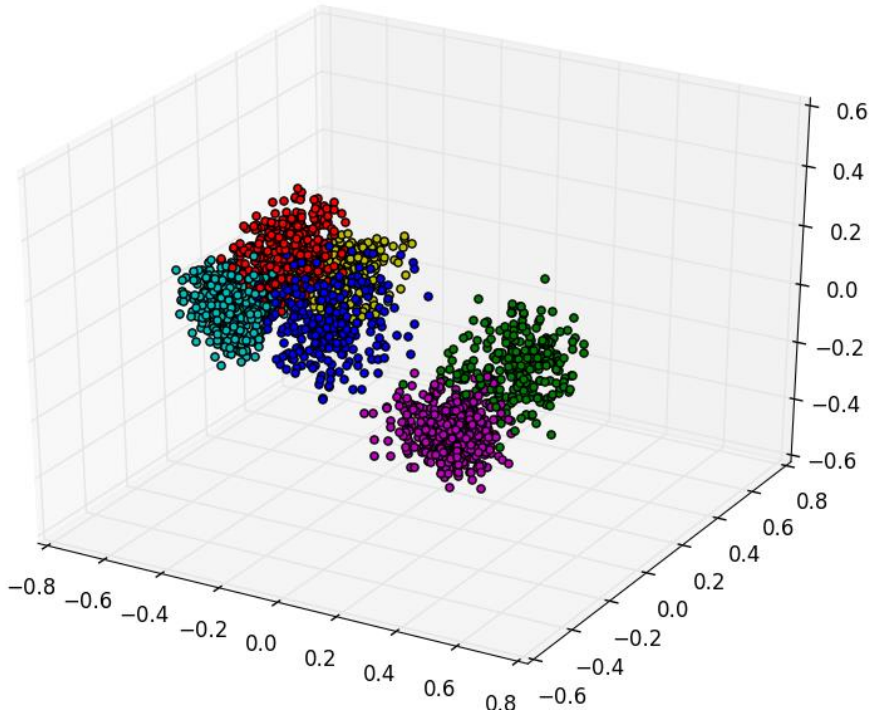
	Hard-Sum Bow	Soft-Sum Bow	Soft-Max Bow
Accuracy	0.248	0.244	0.234

此分類結果雖然和我預期的相同，但相距並不明顯，且三者準確度都不高，這可能是因為選用的 Visual words 太少(這是為了要能夠在 3 維空間中畫出結果)

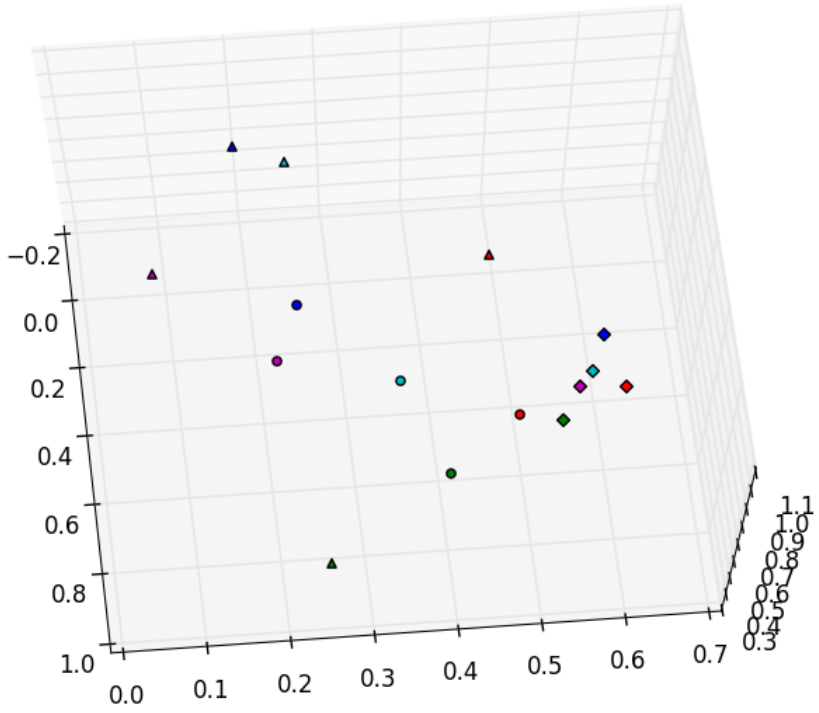
(ii)

(a) 的結果不會改變(因為沒有用到 training data)

(b)



(c)



(d) 我發現使用 train-100 的結果明顯較 train-10 好。除此之外，若我進一步調整選用的 Visual-Word 數量，Accuracy 還可再更高：

	Hard-Sum Bow	Soft-Sum Bow	Soft-Max Bow
Accuracy(c=3)	0.33	0.376	0.308
Accuracy(c=10)	0.456	0.452	0.424
Accuracy(c=25)	0.482	0.48	0.472
Accuracy(c=50)	0.526	0.496	0.49

由上表我還發現 Performance 大致上 Hard-Sum BoW > Soft-Sum Bow >Soft-Max BoW，不過此差距會隨選用的 Visual-Word 數量增加而越來越小