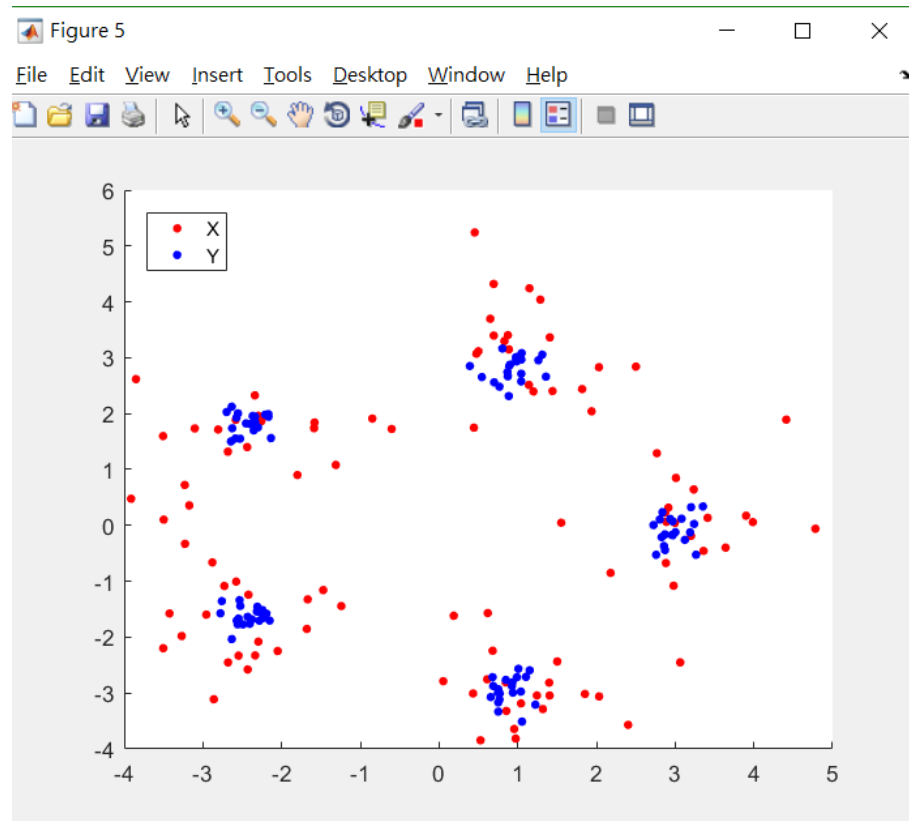
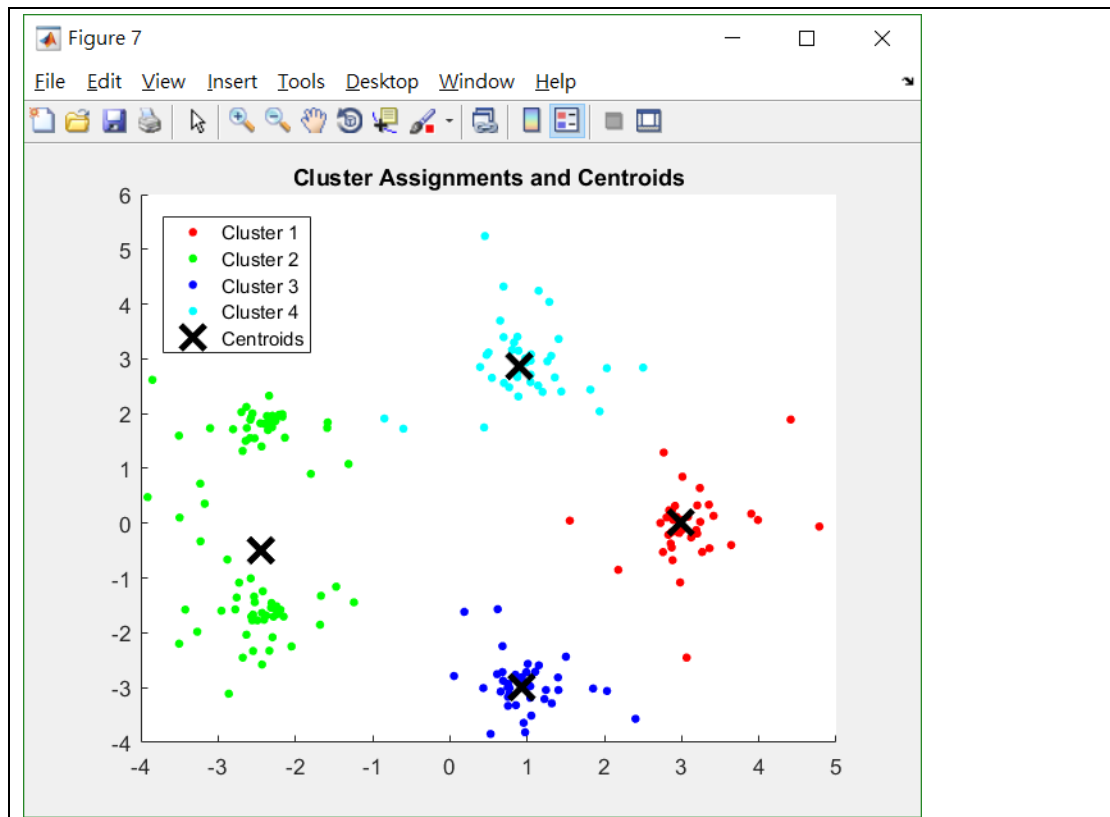


Problem 1 (k-means)

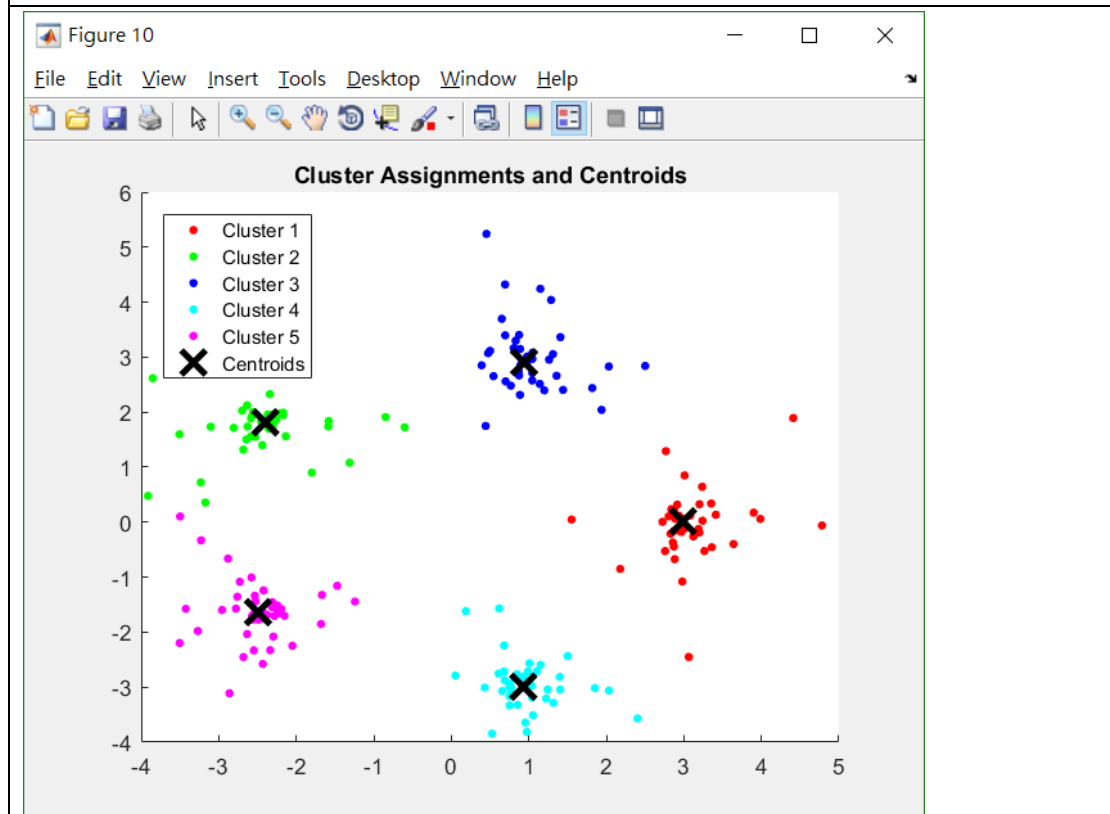
原始資料分布



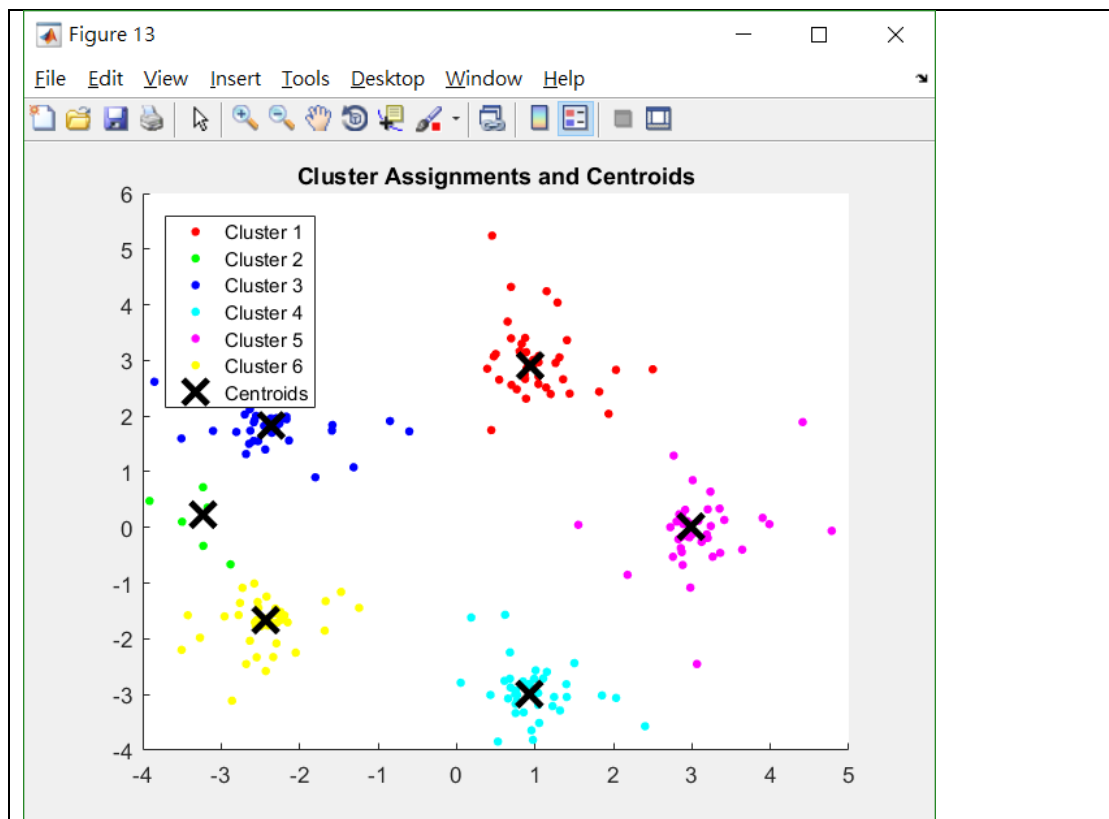
$k = 4$



k = 5



k = 6



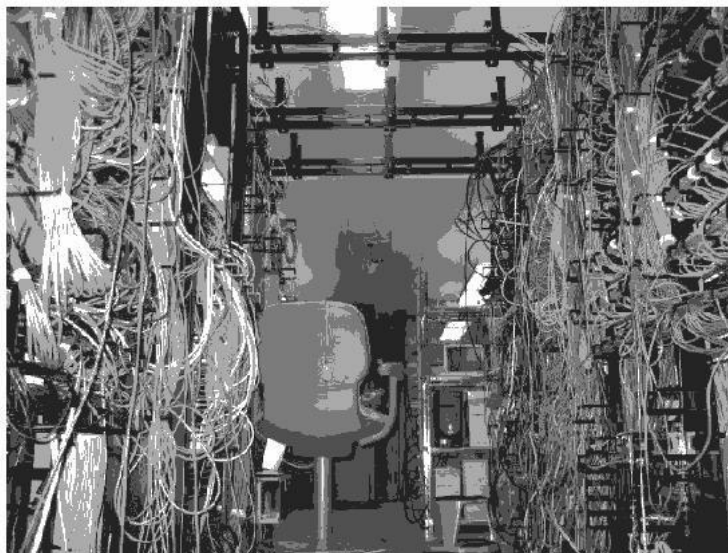
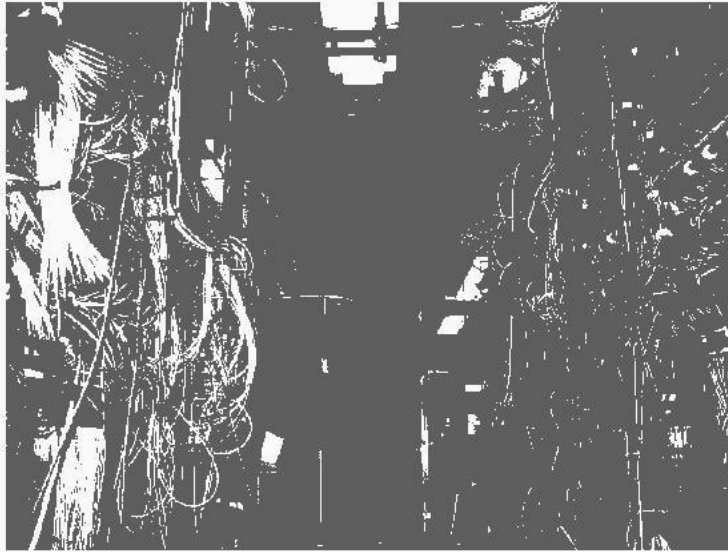
結果與討論:

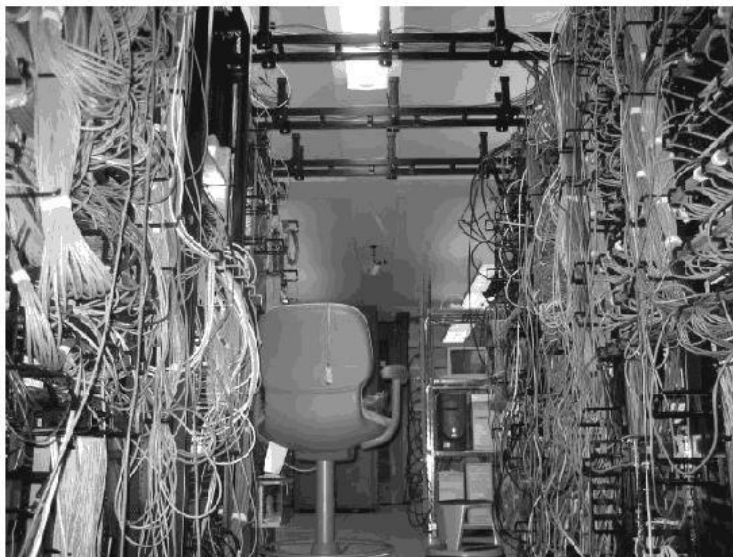
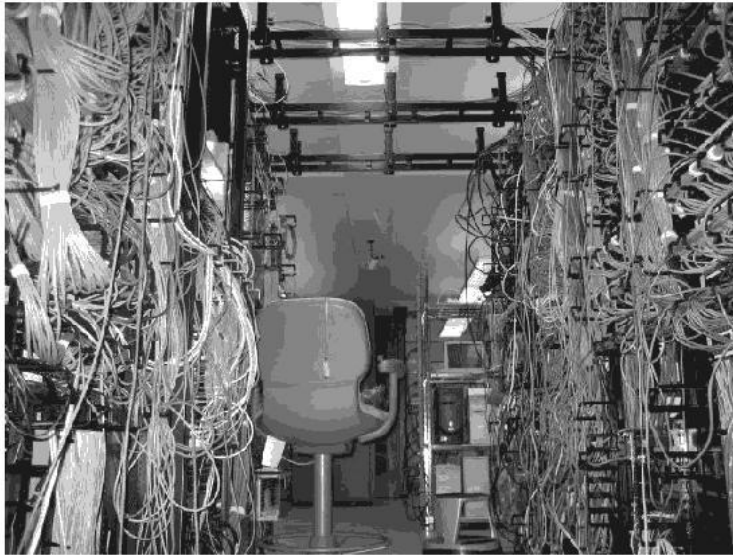
由上圖不同 K 值的 K-means 分布圖呈現，可以發現當 $k = 5$ 時，在視覺上的效果最好，也最合理。

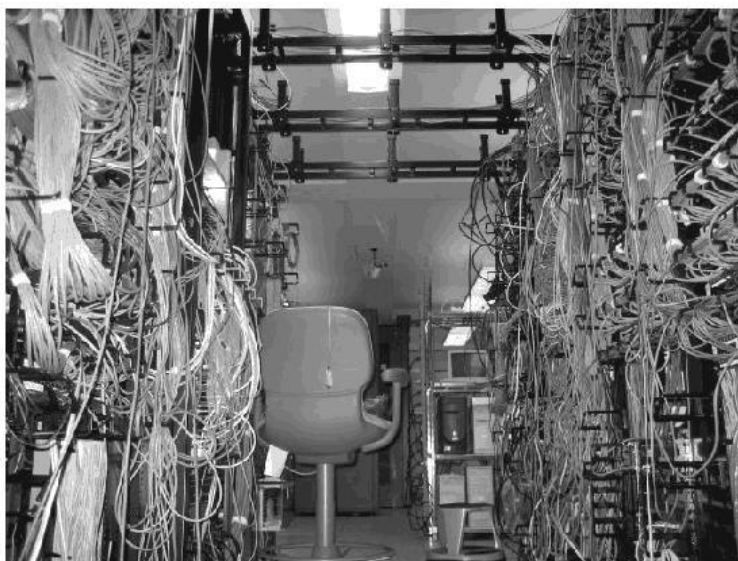
Problem 2 (k-means)

(a) Display the images after data compression using k-means clustering for different values of $k=2,5,10,15$, and 20.

灰階影像(分別為 $k=2,5,10,15$, and 20)

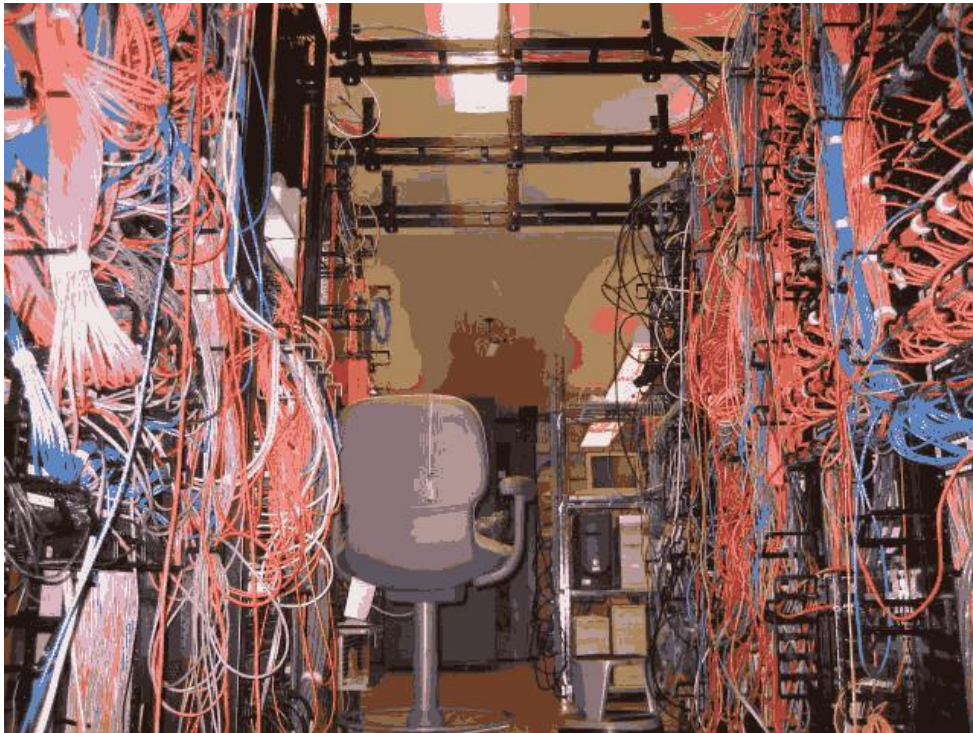


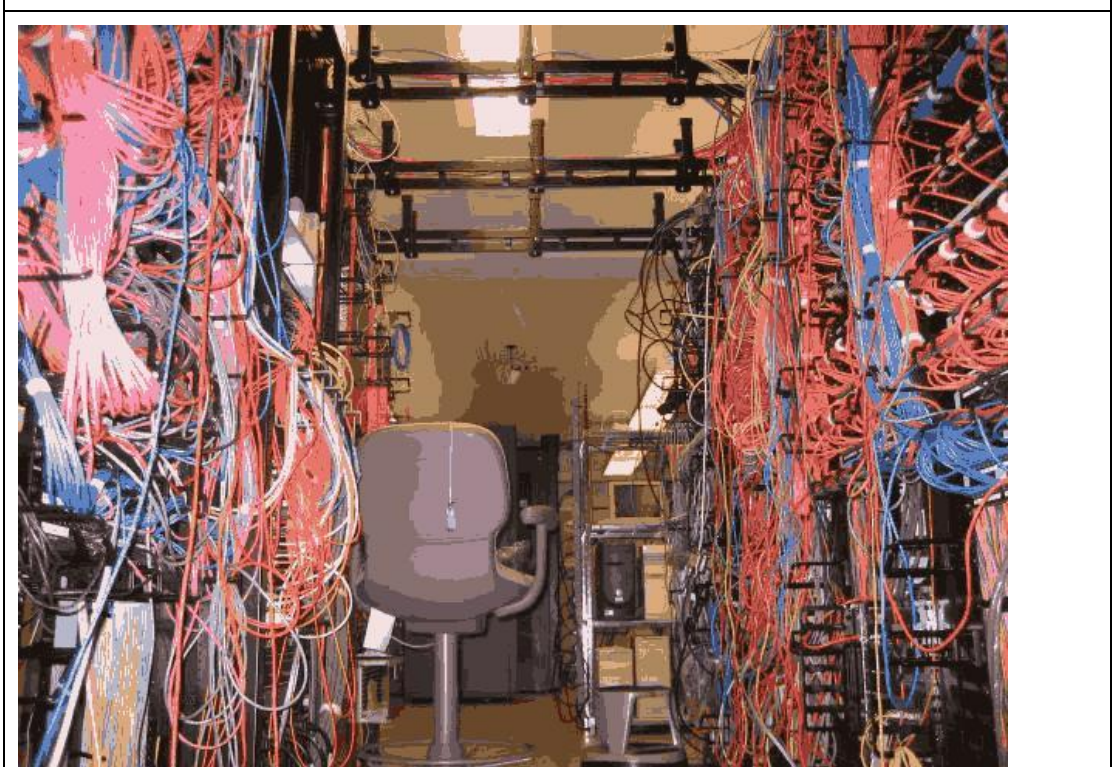
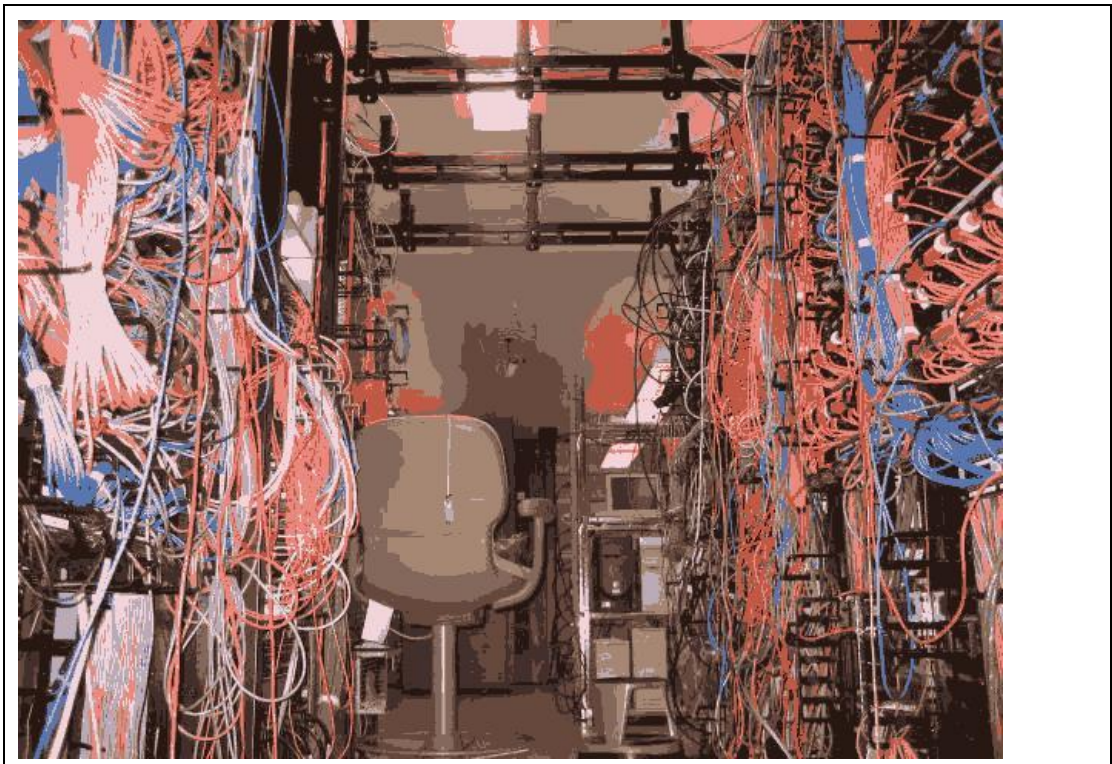




彩色影像 (分別為 $k=2,5,10,15$, and 20)







(b) What are the compression ratios for different values of k ?

將 jpg 檔經過 k-means 分類後，再將處理過後的影像儲存成 jpg 檔，由檔案大小計算其壓縮比(如下表)，可以發現彩色影像的壓縮效果並不明顯，而灰階影像的壓縮效果較好。

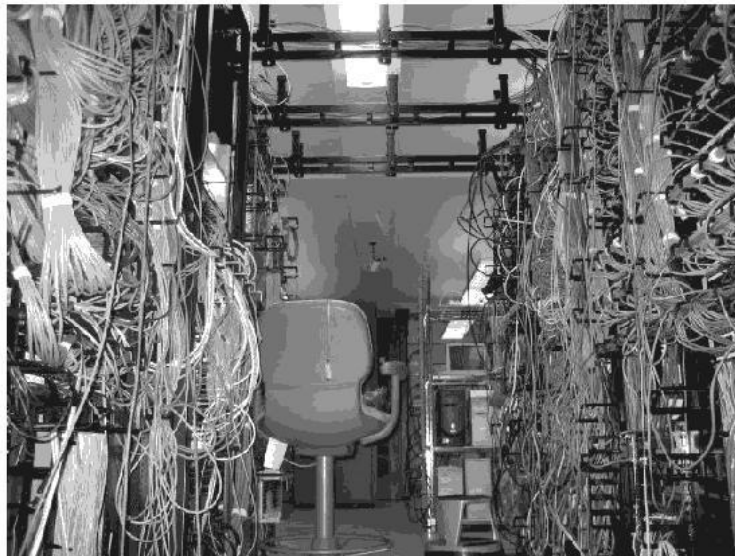
但是從色彩數量來說， k 越小的話照道理壓縮比要越大，但是由於儲存的影

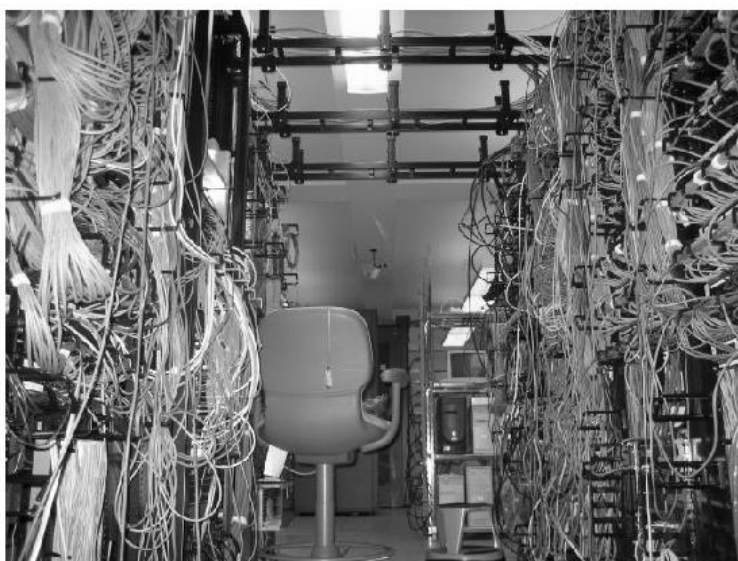
像是由 matlab 直接存成 jpg 檔案，沒有經過演算法壓縮相同顏色的區塊，因此才會使得影像儲存的大小和原始差不多。

壓縮比	原始/處理(彩色影像)	原始/處理(灰階影像)
K = 2	92.4/82.6(KB) = 1.1186	213/54.3(KB) = 3.9227
K = 5	92.4/92.4(KB) = 1	213/75.1(KB) = 2.8362
K = 10	92.4/92.8(KB) = 0.9957	213/71.0(KB) = 3
K = 15	92.4/92.3(KB) = 1.0011	213/69.9(KB) = 3.0472
K = 20	92.4/92.8 (KB)= 0.9957	213/69.4(KB) = 3.0692

(c) You will see that there is a trade-off between degree of compression and image quality. What would be a good value of k for each of the two images?

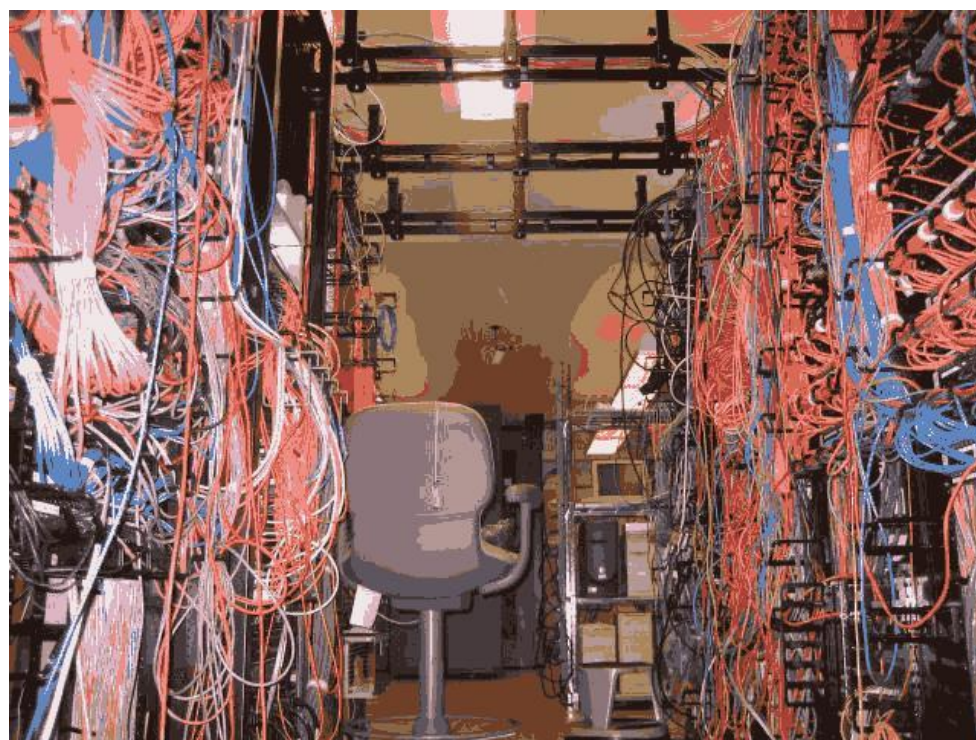
灰階:

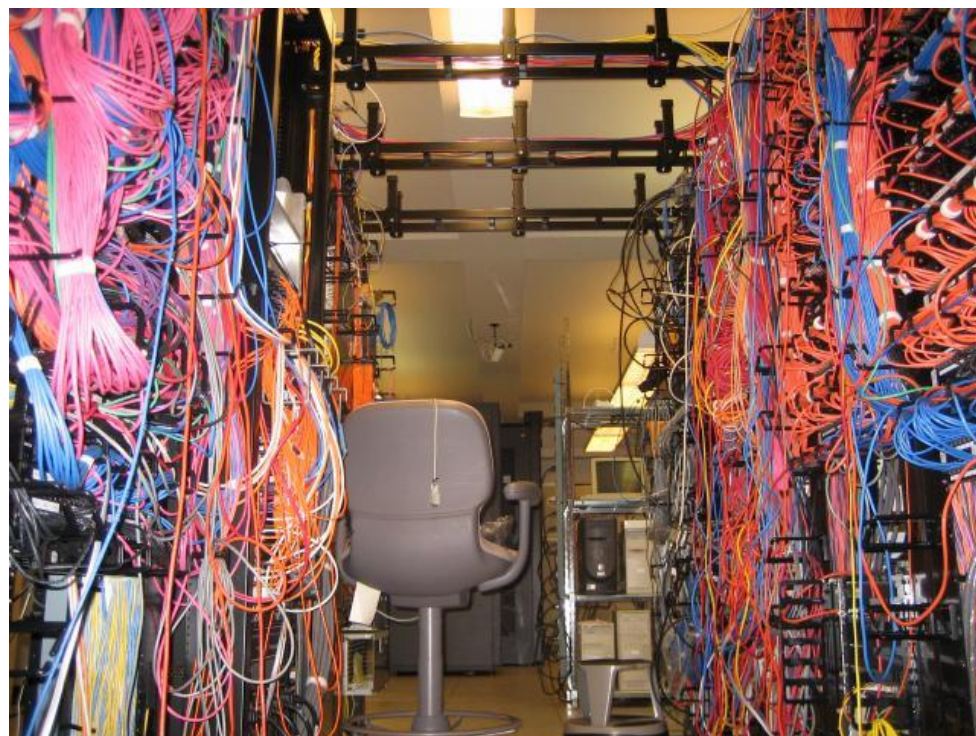




上上圖為 $k=10$ 的時候，上圖為原圖。可以看到雖然細部的地方不太清楚，但是整體還是可以看出是什麼東西，例如：椅子、電線。

彩色：





上上圖為 $k=10$ 的時候，上圖為原圖。已這張圖來看，由 K-means 壓縮後，可以看出藍色和紅色電線的差別，因此我認為 $k=10$ 有達到同時保留資訊和壓縮效果。