

DIP: Final Project Report  
Image stabilization method using VMD & RE

鄭雅文                      梁健偉  
R06922097                  R06521521

### 1.Introduction

**Image stabilization** 是一個目前還滿常被討論的東西，其解決方法可以分為兩種，第一種是硬體型的，也就是從硬體著手在拍攝時同步把一些不必要的影像位置修正回來。另一種是軟體型的，也就是通過影像處理的方法，把影像進行處理從而達到修正的效果。如果我們可以通過利用影像處理的方法把一段影片處理後，從而達到一定預期的效果，那感覺會滿有趣的。

而常常我們在拍攝的時候，拍攝的人都會有一些不自然的移動和不自覺的抖動都會造成我們看到的影片影像的變得不清晰、觀感不佳。而硬體型的成本都較高，所以本次期末報告想要以一段影片使用影像處理進行修正從而穩定影像，降低影像的抖動程度。

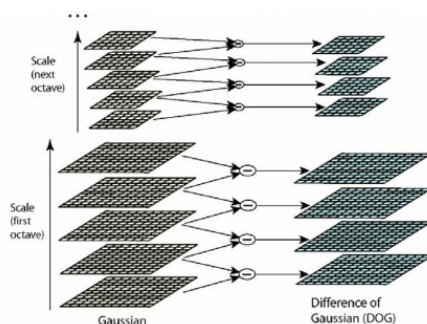
### 2.Feature Matching

論文中所使用的獲取特徵點的方法是 **SIFT + RANSAC**，而論文會用兩個結合的原因是因為在一些質量較差的照片、不容易分析的像素點時，會出現錯誤配對的情況，所以論文為了解決這個問題，在處理完 **SIFT** 得到特徵點後，再使用 **RANSAC** 把那些錯誤點去掉。

**SIFT** 是一個轉換的過程，它會把影像轉換成一些特徵值來描述。**SIFT** 分成以下步驟：

- 尺度空間極值的檢測 **Detection of scale-space extrema**

**SIFT** 裡使用 **DOG** 濾波來建立這個尺度空間。**DOG** 濾波就跟 **Gaussian** 濾波一樣，給定兩個不同尺度的 **Gaussian** 值，相減之後對原來的影像做濾波。



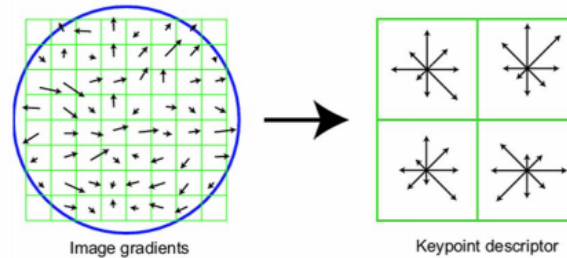
- 方向指定 **Orientation assignment**

這是為了接下來要拿來做配對用的。對於每一個特徵點我們要去計算它 梯度

的大小和方向。

- 對影像進行描述 Local image descriptor

已經找出特徵點的位置、尺度、方向，接下來第四步驟就是要把他進行描述。

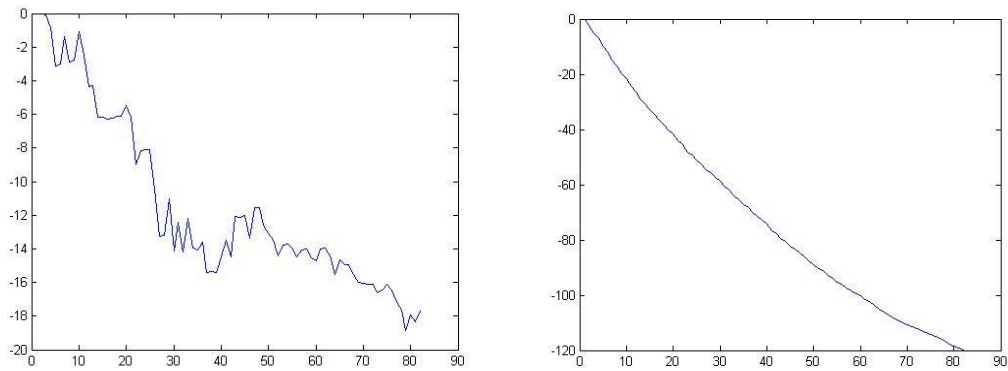


而我們在這次實作中只有取 SIFT 沒有使用 RANSAC，是因為我們在做完 SIFT 後發現，原本的 feature matching 的方式就已經沒有 mismatching 的問題，我們的影像特徵點只需要使用 SIFT 所得到得 matching point 即可，所以就沒有使用 RANSAC。



### 3.Global Motion Vector

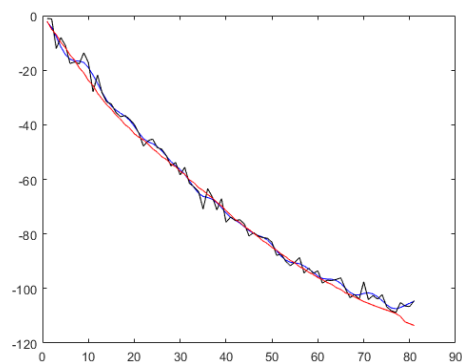
在影像處理中，全域移動向量(Global Motion Vector)是常被用來輔助處理的資訊，代表影像在連續的影像中每一個 frame 與 frame 之間的位移。而在計算 GMV 的過程中，因為並不是每個 frame 之間的移動都相等，可能很小、很大，甚至有些特徵點沒有位移，但是因為 GMV 是所有特徵點的平均位移，如果把沒有位移的特徵點也除掉，會造成 GMV 有額外的不穩定情形，經過實驗之後，決定只把有位移的特徵點納入 GMV 計算，就會比較穩定。



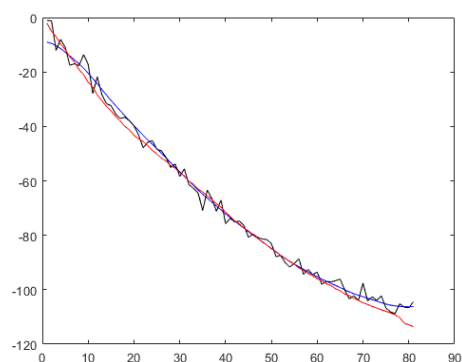
#### 4.VMD

而 VMD 主要有四個參數,分別是  $\alpha$  - 代表數據保真度約束的平衡參數,  $\tau$  - 雙重上升的時間步長 (選擇 0 表示噪音鬆弛),  $K$  - 要恢復的模式數量,  $\text{tol}$  - 容忍度; 通常在  $1e-6$  左右。

而我們選用的參數與論文的參數一樣,其  $\alpha = 100$ 、 $\tau = 0$ 、 $k = 5$ 、 $\text{tol} = 1e-6$ 。而我們通過數次試驗,發現  $\alpha$  跟論文一樣設為 100 時,效果較好。其輸出如下圖,雖然曲線看起來較不平滑。



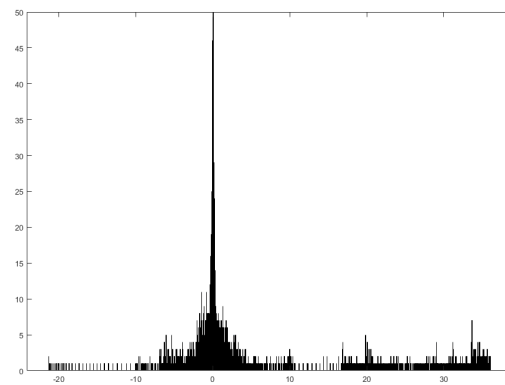
而下圖為  $\alpha$  等於 2000 時之輸出,其曲線會看起來比較平滑,但分解出來的效果並不如上述  $\alpha = 100$  時好。



並且我們也有對於”IMG\_3715”影片做 VMD，而這種 intentional motion 為水平移動，jitter motion 為垂直移動的影片，偵測出 GMV 之後直接把 jitter motion 全部減掉即是理論上的穩定影片，所以不適用 VMD 分解，不太算是本篇論文適用的方法。

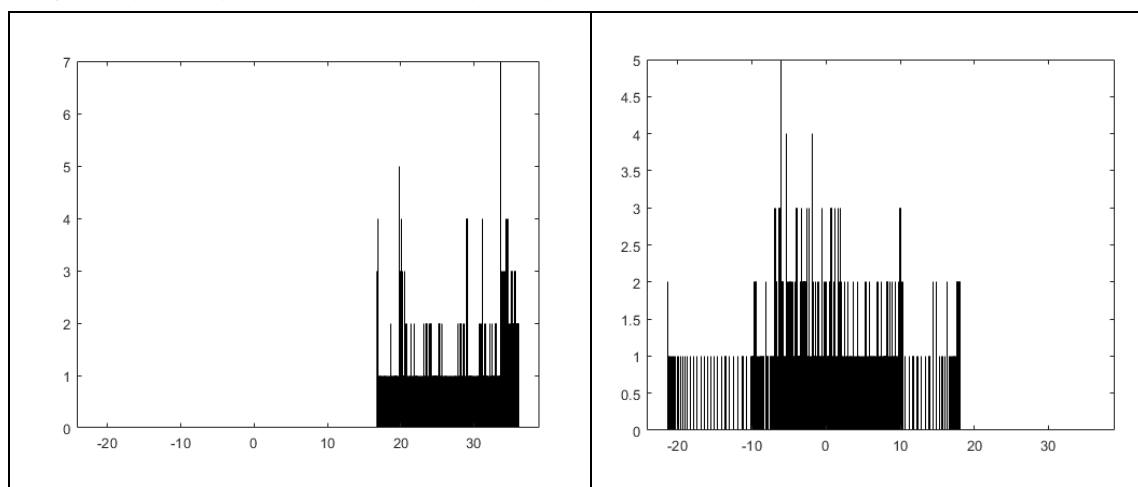
### 5.Relative entropy

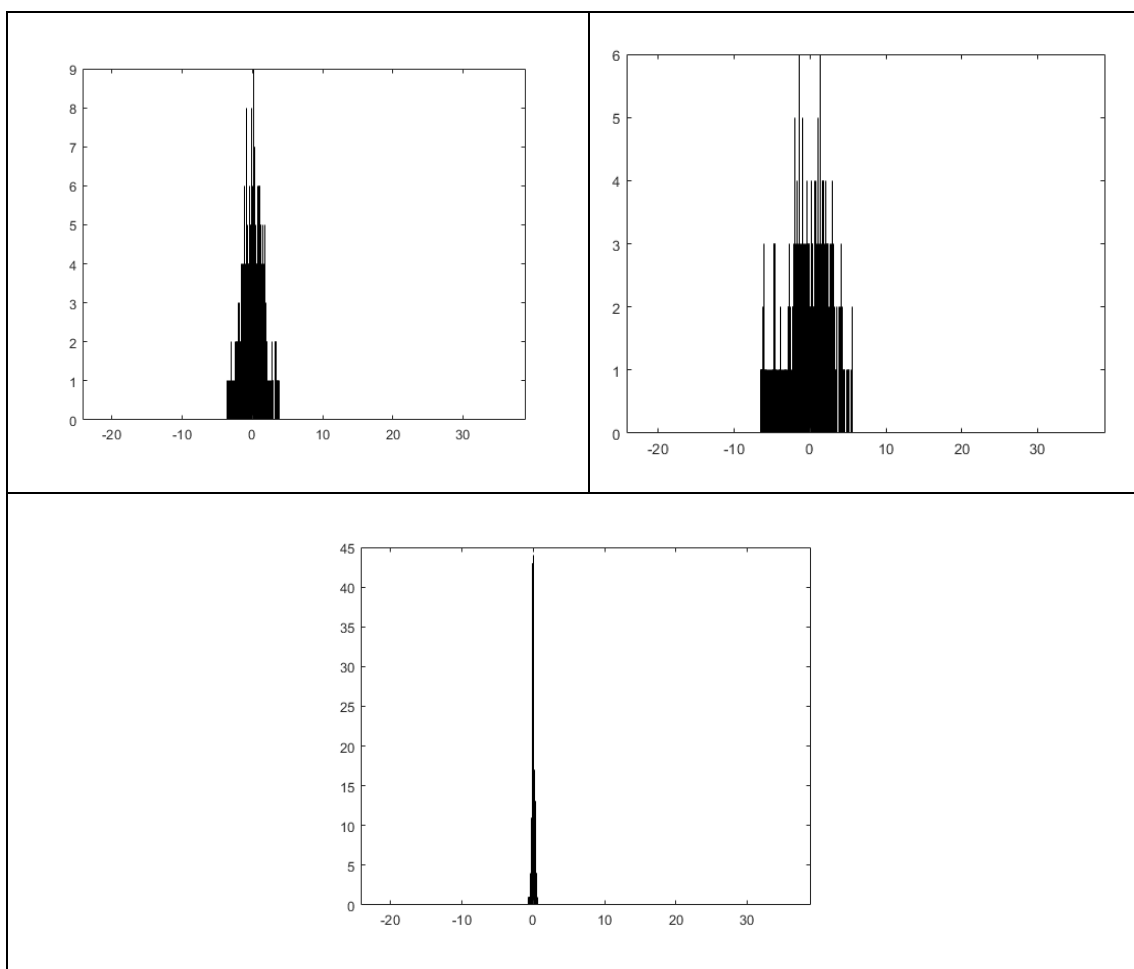
在計算 Relative entropy 時，因為要處理的資料是一個一維陣列，但是 Relative entropy 是取機率分佈為 input，例如現在第一個 frame 到第二個 frame 的 GMV 是-1.4173，第二個 frame 到第三個 frame 的 GMV 是-1.4714，那-1.4173 的機率分佈就是所有的 frame 分之一，-1.4714 的機率分佈也是所有的 frame 分之一。因為我們的 data 是浮點數，所以更難有相同的 y 出現。所以這種情況就要用 histogram 的方式處理機率分佈問題：



如圖：先將所有的訊號做成 histogram，間距自訂，每個間距出現的次數除以有幾個 frame 的間距，即是他出現的機率。

但是如果將 5 種 mode 分開，會得到下圖：





因為有些 mode 的分佈範圍沒有 match 在一起，所以 relative entropy 仍然為 0。所以必須再將這些分佈 normalize 到 0 到 1 之間，自行定義中間該有幾個 bins，再對這些分佈取 relative entropy，才能得到較合理的數據。

在經過上述的 RE 的算法我們所得出的矩陣如下：

Q\P	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
Mode 1	0	0.7953	0.7036	0.5838	0.4187
Mode 2	0.2979	0	0.2961	-0.0030	0.0197
Mode 3	0.7104	0.8718	0	0.5032	0.3437
Mode 4	0.5696	0.5444	0.2690	0	0.3295
Mode 5	0.4784	0.8386	0.3979	0.5401	0

先前我們是水平看，發現 mode 1 和 mode 2 不太相似，但改為垂直看之後就會發現 mode 1 和 mode 2 相近。對照回原來的公式：

$$D_{\text{KL}}(P\|Q) = - \sum_i P(i) \ln \frac{Q(i)}{P(i)}.$$

等價於

$$D_{\text{KL}}(P\|Q) = \sum_i P(i) \ln \frac{P(i)}{Q(i)}.$$

這裡的算法應該是取 mode 1 為 P，mode 2~5 為 Q，才能得到如 paper 上 mode 1 和 mode 2 較相近，和其他 mode 距離較大的結果。

## 6.Data select

在選擇怎樣的影片資料時，就遇到了一些問題。因為原本想用車載裝置錄影片，結果發現像是行車記錄器那種只拍攝前景的影片的位移不是這篇 paper 想要解決的問題，而且有其他更容易的方式可以解決：

[https://youtu.be/15v\\_9VJa020?t=3m49s](https://youtu.be/15v_9VJa020?t=3m49s)

所以我們討論出這篇 paper 適用的影片內容應該是相機要用 pan 或是 tilt 的方式做上下/左右移動，且如果只做一維的 Global Motion Vector 的話，noise 應該也要在同個方向上，才能分類出 intentional motion 跟 jitter motion。

後來發現想要找 ground truth 的方法很難製造震動，有些製造震動的方法無法取得 ground truth，所以決定用一段平穩的影片製造震動的 noise，再做還原，跟原來的影片做比較。這樣不但可以跟最佳的結果做比較，也可以得到 noise 的 ground truth(如果不能得到 intentional motion 的 ground truth，就取 jitter motion 的 ground truth)。

## 7.result

最後，本次報告所做實驗結果都放在 output 文件中，裡面有使用到加入垂直 noise、水平 noise 的影片以及修正後的影片。

可以看到影片在經過處理後，大部份的抖動都被修正回來了，這表示我們的方法是有被成功的應用在這些影片上，也能夠有效的解決這個問題。

## Reference

1. Hao, D., Li, Q., & Li, C. (2017). Digital Image Stabilization Method Based on Variational Mode Decomposition and Relative Entropy. *Entropy*, 19(11), 623.  
<http://www.mdpi.com/1099-4300/19/11/623/pdf>
2. Dragomiretskiy, K., & Zosso, D. (2014). Variational mode decomposition. *IEEE transactions on signal processing*, 62(3), 531-544.  
<https://ieeexplore.ieee.org/document/6655981/>
3. Mathematical statistics, the Kullback–Leibler divergence (also called **relative entropy**)  
[https://en.wikipedia.org/wiki/Kullback%E2%80%93Leibler\\_divergence](https://en.wikipedia.org/wiki/Kullback%E2%80%93Leibler_divergence)
4. Scale-invariant feature transform  
[https://en.wikipedia.org/wiki/Scale-invariant\\_feature\\_transform](https://en.wikipedia.org/wiki/Scale-invariant_feature_transform)