

以區域基礎的光學流動倒投影技術求取真實的影像移動向量

詹寶珠^{*1}、黃詒琳¹、陳怡良²

¹國立成功大學電機工程學系、電腦與通信工程研究所

²立德管理學院資訊工程學系

Email:pcchung@eembox.ee.ncku.edu.tw

PATTERN RECOGNITION、vol. 40、pp.1066-1077、2007

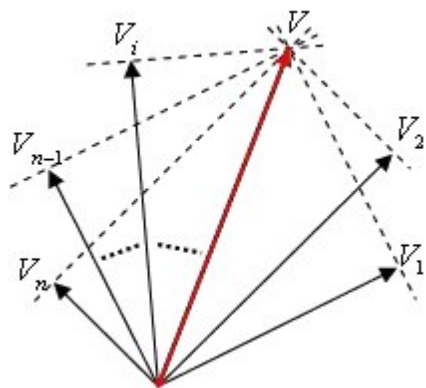
對於視訊影像分割，移動向量(motion vector)扮演著舉足輕重的角色。雖然移動向量已廣泛應用於視訊影像壓縮，但視訊影像壓縮所重視的是視覺上的最佳化而不是視訊影像中的物件真實的移動向量。因此若直接使用視訊影像壓縮中所使用的移動向量來當視訊影像分割或追蹤的特徵，將使得影像分割結果不如預期理想。基於此考量，我們提出了以光學流動倒投影(Optical Flow Back-Projection)技術來求取更為真實的影像移動向量，以利後續視訊影像分割與追蹤之用。

真實的三維空間中，動量(motion)描述了空間物件的移動特徵。在理想的情況上，真實三維空間轉換成二維的平面視訊影像；而動量也因此轉換成二維的移動向量(motion vector)。所謂移動向量的定義就是用來描述視訊影像中，影像資訊隨著時間的變化所產生相對應的移動特徵。一般來說，求取移動向量的演算法可概分為(1)區塊基礎無參數型(2)參數模型及(3)光學流動模型等三大類型。這些演算法以不同方式都為了找出視訊影像中連續兩張影像上對應點的轉換關係，藉此來描述移動向量。

區塊基礎無參數型就是所謂的區塊匹配(block-matching)方法，此方法是最簡單的方法，計算量也最小；然而此演算法所求取出來的移動向量偏重於視覺最佳化而不是視訊影像中物件真實的移動向量。因此區塊基礎無參數型演算法廣泛地應用於視訊影像壓縮，但卻不適用於視訊影像物件分割與追蹤之用。參數模型就是所謂仿射(affine)或透射(perspective)模型，此方法使用一些參數來描述是視訊影像中物件的平移、縮放與旋轉的關係。為了求取這些參數，必須先行找出更多視訊影像中連續兩張影像的對應點，這對於自動化影像處理將是一大挑戰，同時此方法的計算量也是非常驚人的大。光學流動模型早期由Horn-Schunck所提出的方法，主要是依據視訊影像中連續兩張影像的影像強度變化來求取其對應的移動向量。光學流動模型的優勢在於不用像參數模型必須先行找出對應點，而且也不像區塊基礎無參數型偏重於視覺最佳化，而是偏重於物件真實的移動向量。然而由於光學流動模型的限制，所求取出來的光學流動(optical flow)並非真實的移動向量，而是移動向量在影像梯度(Gradient)方向的分量。



圖一：區域分割結果(a)原圖 (b)分割結果



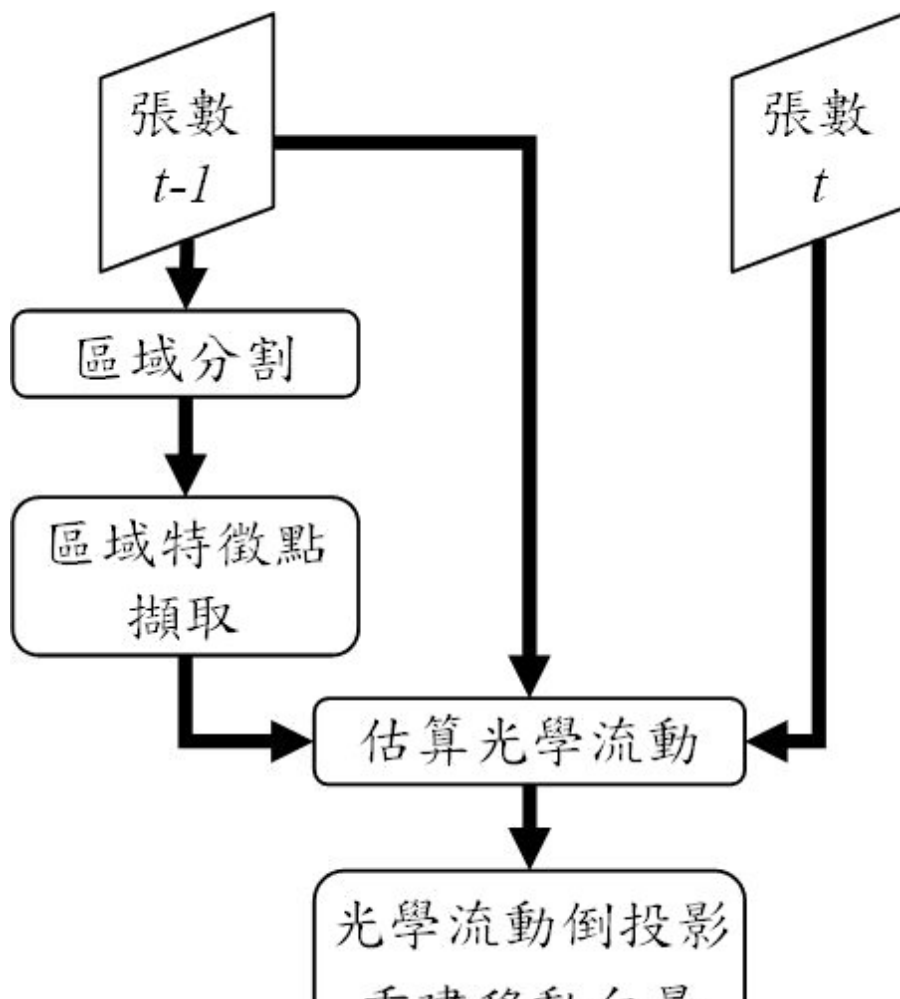
圖二：區域中光學流動分佈及其倒投影的移動向量
為了求取視訊影像中真實的移動向量，我們摒除了視覺最佳化的區塊基礎無參數型，也摒除了需前處理及

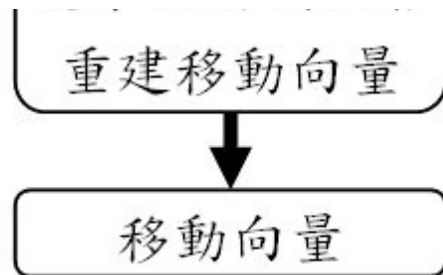
大量計算量的參數模型；而使用兼俱此兩種方法優勢的光學流動模型為基礎，藉由相同移動向量區域的光學流動倒投影出視訊影像中真實的移動向量。如前所述，光學流動模型所求取出來的光學流動並非真實的移動向量，而是移動向量在影像梯度(Gradient)方向的分量，基於上述考量，我們試圖找出影像中具相似影像強度的鄰近像素所組成區域，並假設每個區域內的像素具有相同的移動向量，如圖一所示，圖一(a)為原始影像，而圖一(b)為區域分割結果。接著我們選取區域中影像強度與週圍變異度較大的像素(pixel)為特徵點，對每個區域求取區域中每個特徵點的光學流動(optical flow)。由於區域中的每個特徵點的影像梯度不儘相同，因此若區域中的特徵點有 n 種不同的梯度，則我們必可找出 n 種不同方向的光學流動，如圖二中 V_1, V_2, \dots, V_n 所示。由於光學流動為移動向量在影像梯度方向的分量，加上我們假設每個區域內的像素具有相同的移動向量。因此我們只要倒投影區域中這些特徵點的光學流動(即各影像梯度的分量)，理論上便可求得區域的移動向量，如圖二中的 V 所示。然而由於雜訊的干擾及區域分割精確性的問題，致使這些光學流動(分量)可能無法倒投影出唯一的移動向量(主向量)。因此我們使用最佳化的方法，試圖對一區域找出最適合的移動向量，使得其在該區域各特徵點梯度方向之分量與求取之光學流動誤差最小。即：

若 $V_1 \angle \theta_1, V_2 \angle \theta_2, \dots$ ，及 $V_n \angle \theta_n$ 表示區域中各特徵點之光學流動，則依我們提出之方法可求得區域最

$$\text{佳移動向量 } V \angle \theta, \text{ 其中 } \theta = \tan^{-1} \left(\frac{\left(\sum_{i=1}^n V_i \sin \theta_i \right) \left(\sum_{i=1}^n \cos^2 \theta_i \right) - \left(\sum_{i=1}^n V_i \cos \theta_i \right) \left(\sum_{i=1}^n \cos \theta_i \sin \theta_i \right)}{\left(\sum_{i=1}^n V_i \cos \theta_i \right) \left(\sum_{i=1}^n \sin^2 \theta_i \right) - \left(\sum_{i=1}^n V_i \sin \theta_i \right) \left(\sum_{i=1}^n \cos \theta_i \sin \theta_i \right)} \right);$$

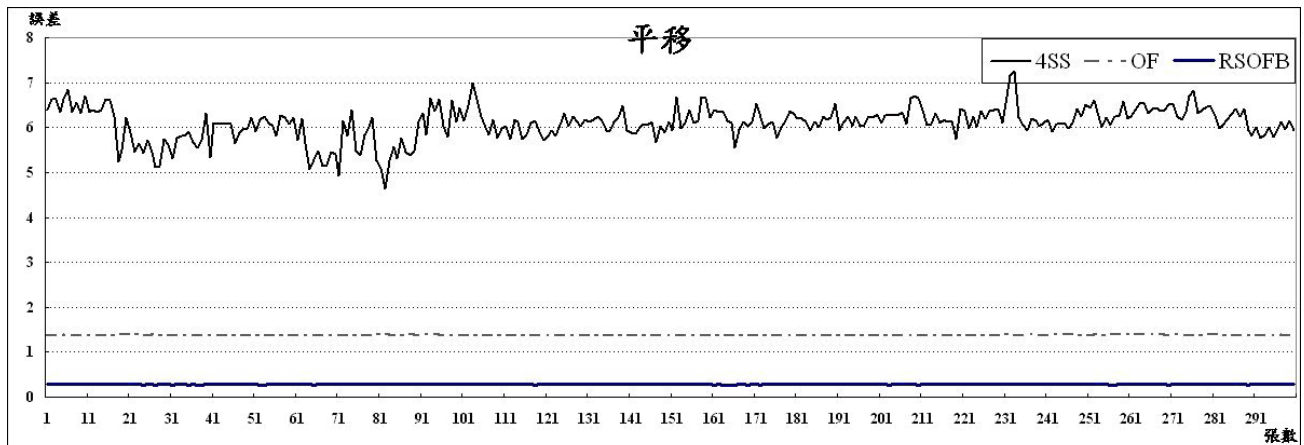
而 $V = \frac{\sum_{i=1}^n (V_i \cos(\theta - \theta_i))}{\sum_{i=1}^n \cos^2(\theta - \theta_i)}$ 。綜上所述，我們提出的區域基礎的光學流動倒投影演算法流程如圖三所示。



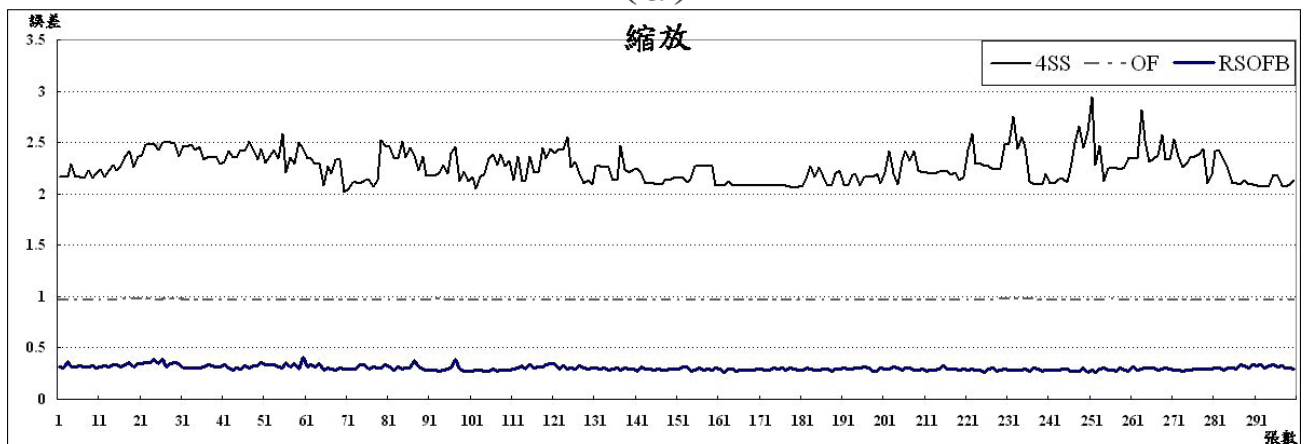


圖三:區域基礎的光學流動倒投影演算法流程圖

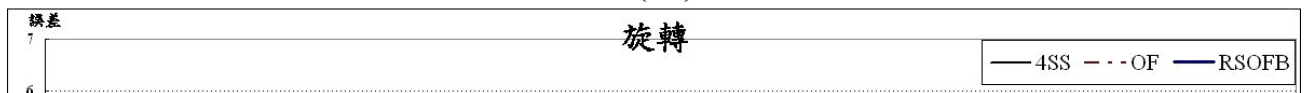
為了驗證區域基礎的光學流動倒投影演算法所求移動向量的準確性，我們將我們提出的區域基礎的光學流動倒投影(RSOFB)方法跟區塊基礎無參數型的四步搜尋(4SS)演算法及原始的光學流動模型的Horn-Schunck(OF)方法做實驗比較。我們對測試視訊影像分別模擬平移、縮放與旋轉，使用上述方法求取移動向量，所得誤差分佈圖如圖四所示。其中水平軸為視訊影像張數，而垂直軸表示平均移動誤差大小。由圖四結果可明顯看出我們的方法可求得較為精確的移動向量。同時我們也考慮到光學流動模型易受外在光源變化的影響，因此我們也模擬不同光線(調整影像中的gamma值)變化，如圖五所示。圖六顯示不同gamma值變化下的比較結果，也可看出雖然我們提出的方法也會受光線變化影響準確性，但整體而言仍有不錯的準確性。最後我們用三維虛擬實境的Yosemite峽谷所映射的二維視訊影像當成實驗視訊影像，如圖七(a)所示，實驗結果如圖七(b)。實驗結果顯示我們的方法(RSOFB)所得的移動向量誤差都在2個像素以下，而四步搜尋演算法(4SS)及Horn-Schunck方法(OF)的誤差卻大都高達2個像素以上。由上述的實驗結果顯示，我們提出的區域基礎的光學流動倒投影技術確實能比四步搜尋演算法及Horn-Schunck方法求取更真實的影像移動向量。

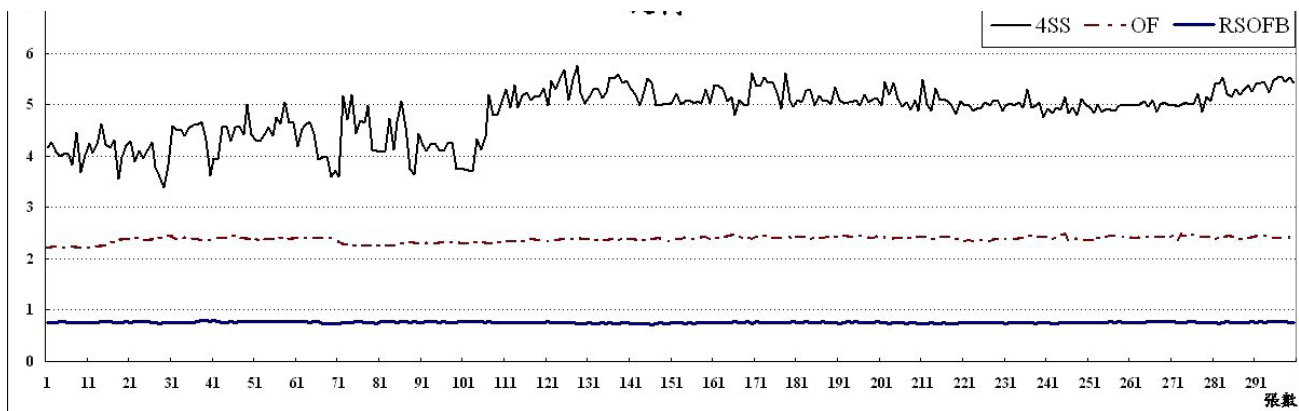


(a)



(b)





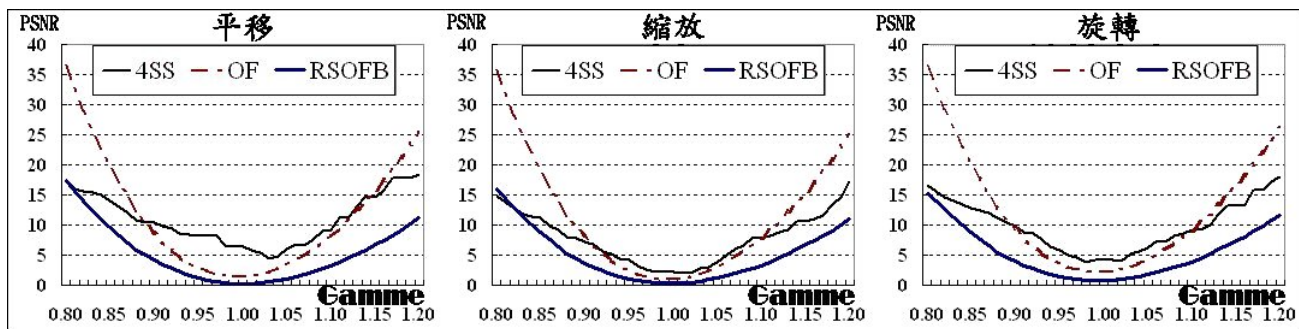
(c)

圖四: 四步搜尋演算法(4SS)、Horn-Schunck方法(OF)及區域基礎的光學流動倒投影演算法(RSOFB)實驗比較(a)平移 (b)縮放 (c)旋轉。

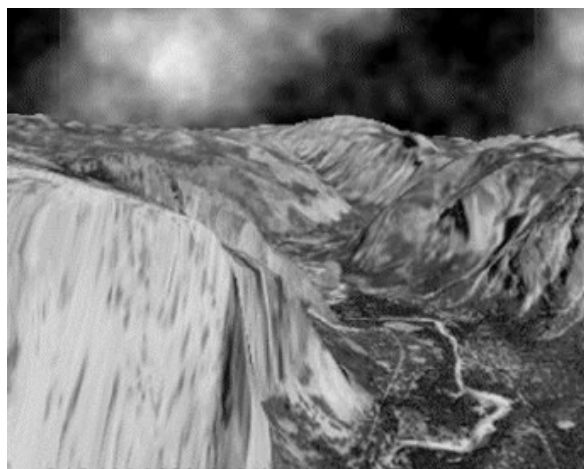


(a)Gamma=0. 1(b)Gamma=0. 5(c)Gamma=1(d)Gamma=1. 5(e)Gamma=2

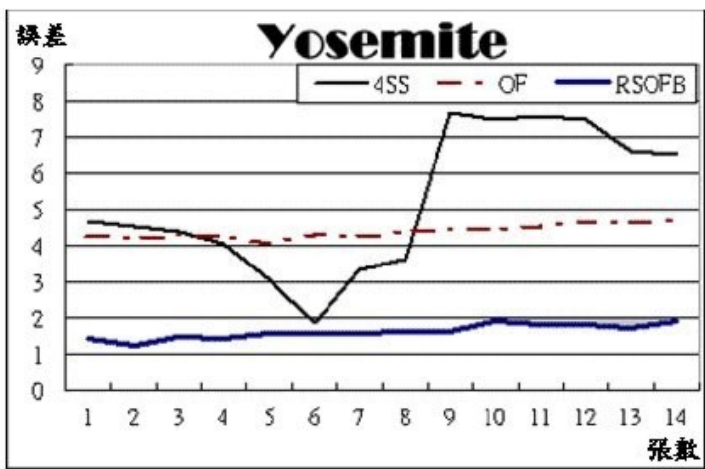
圖五：不同gamma值對影像亮度的變化



圖六：不同gamma值變化的實驗比較結果



(a)



(b)

圖七：Yosemite 視訊影像實驗結果(a)Yosemite視訊影像 (b)實驗結果

在此研究主題中，我們驗證了使用區域基礎的光學流動倒投影技術確實能比傳統應用於視訊影像壓縮的區塊基礎無參數型方法及傳統的光學流動模型方法取得真實的移動向量。雖然區塊基礎無參數型方法能獲得視覺上較佳的結果，但卻不是真實的移動向量，這對有真實的移動向量需求的應用極為不適合。我們的方法改良傳統光學流動模型方法並輔以區域性的特徵，獲得比傳統光學流動模型方法更為真實的移動向量。這對後續需要真實移動向量的應用如MPEG-4的移動物件平面(VOP)分割及視訊影像物件追蹤等，我們提出的區域基礎的光學流動倒投影技術將是更好的選擇。然而，由我們的實驗經驗得知，當影像的移動向量太大時，光學流動模型將會失去準確性。因此，光學流動模型，包含我們提出的方法，是比較適用於一般低的移動向量的視訊影像。

Copyright 2009 National Cheng Kung University