

使用區塊匹配之視差估計法則

林國祥
大葉大學資訊工程學系
副教授
khlin@mail.dyu.edu.tw

黃鈞凱
大葉大學資訊工程學系
研究生
r9706010@mail.dyu.edu.tw

摘要

本篇論文提出一個使用區塊匹配的視差估計法則。此提出方法由前處理、像素為基礎的視差估計以及視差修正所組成。在前處理的部份包含兩個程序，色彩轉換和邊緣偵測。為了達成視差估計，本論文提出一個基於正規化交互相關函數並且考慮色彩和空間特徵的成本函數。透過此成本函數可以計算獲得初始視差圖。當進行視差修正時，遮蔽區域內的視差值可以被修補獲得。根據實驗結果，本論文提出的法則可以達成良好地視差估計。

關鍵詞：立體匹配、視差估計。

Abstract

In this paper, we proposed a disparity estimation scheme using block matching. The proposed method is composed of pre-processing, pixel-based disparity estimation, and disparity refinement. In the pre-processing, two procedures, color transform and edge detection, are performed. To perform disparity estimation, a cost function based on normalized cross correlation as well as color and spatial features is developed and used as a measurement in the proposed disparity estimation method. In disparity refinement, the disparity values in occlusion regions are estimated. According to experimental results, our proposed scheme can perform disparity estimation well.

Keywords: Stereo Matching, disparity estimation.

1. 前言

隨著 3D 顯示器、視訊壓縮與網路傳輸技術的快速發展，具立體感之家用影音娛樂設備已經不再遙不可及。目前 3D 顯示器的輸入方式之一為兩個通道的彩色視訊（意即左眼與

右眼視訊）。這種立體視訊可以由兩架平行擺設的攝影機擷取獲得。為了達成有效的立體視訊資料壓縮，如何正確地取得視差資訊成為立體或多視角視訊壓縮技術內相當重要的問題之一。

然而，正確地估計出視差值是一件困難的工作。估計深度資訊可能的誤差，除了拍攝機器與人為操作的干擾外，主要影響還有場景光線、物件遮蔽、平滑區域等因素。此外，當拍攝時光線的變化會影響輸入影像的匹配困難度，而存在某張影像的遮蔽部份會讓另一張影像對應不到正確的地方變成多對一的情形，再者平滑的區域需要參考更大的面積才能取得足夠的紋理資訊。

要讓二維影像平面產生立體效果，需指派畫面各個區域適當的深度數值，以決定物件的前後關係。而計算深度數值的方法大致上可以分成局部性[3,9,10,13]與整體性[2,7,11,12,15,19]兩種方式。

局部性方法利用預先假設固定大小和形狀的搜尋視窗，對於視窗範圍內的像素點與其他影像中對應的區域內之像素點，計算局部相似度。它的優點是可以簡單地產生密集的視差圖，但是在視差中斷的部分可能會產生錯誤的配對。

整體性方法是以色彩分割的方法，將相似色彩或者平滑的區域分別分割出來，再以這些分割區域去產生配對，同時決定這些分割區域所有像素點的視差。雖然通常能取得較為正確的結果，卻需要許多難以決定的參數與昂貴的計算成本。

除了以上兩種方式，還有利用邊緣[8]、條件機率[14]、種子點[6]的視差計算方式，並且有數個對於遮蔽或不正確的初始視差修正方法[4,5,16,18]提出。

本篇論文使用局部性方法計算，配合數種限制條件(同軸線限制、相似性限制與連續性限制)，以成本函數計算出稠密的視差圖，再對某些不正確的位置進行處理以輸出最後的結果。

2. 本論文提出之視差估計法則

本論文提出之視差估計法則，可以分成三個部份：前處理、視差估計與視差修正。本論文提出 pixel-based 視差估計法則以正規化交互相關函數為基礎，配合顏色與空間特性，並透過區塊搜尋方式，來獲得視差資訊。視差估計之系統流程圖如圖 1 所示。後續，我們將逐一介紹本系統之每個部份。

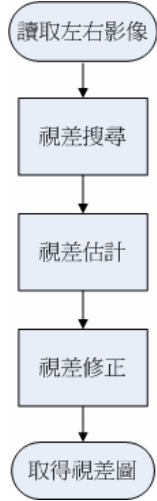


圖 1、本論文提出之視差估計系統流程圖

2.1 前處理

為了視差估計與視差修正，本論文前處理部份包含色彩轉換與邊緣偵測。色彩轉換部份，將輸入彩色影像轉換或的灰階影像。根據文獻[1]，色彩轉換公式如下：

$$Y = 0.299R + 0.587G + 0.114B \quad (1)$$

其中，R、G、B 分別表示 RGB 色彩空間中三原色，R、G、B 之資訊。

由於視差修正時需要邊緣資訊，本論文使用 Sobel 運算子[1]，獲得影像中邊緣的資訊。

2.2 視差估計

在本論文中，我們假設輸入之左右眼影像已經完成影像校正的處理，意即左右眼影像內之同軸線(epipolar line)已經校正成水平線。故視差估計僅僅需要進行水平搜尋即可。

為了衡量是否配對，本論文基於正規化交互相關函數 (Normalized cross correlation, NCC)、顏色與空間相關性，設計一個成本函

數。此成本函數(Cost)定義如下：

$$Cost(i, j, d) = \alpha_1 \times f_1(i, j, d) + \alpha_2 \times f_2(i, j, d) + \alpha_3 \times f_3(i, j, d) \quad (2)$$

其中， $f_1(i, j, d)$ ， $f_2(i, j, d)$ ， $f_3(i, j, d)$ 分別表示正規化交互相關函數、顏色與空間相關性特性。 α_1 、 α_2 和 α_3 為權重係數。

基於公式(2)，則初始視差值可以求得如下：

$$d^l(i, j) = \min_d Cost(i, j, d) \quad (3)$$

以下，我們將逐一說明 f_1 、 f_2 和 f_3 。

2.2.1 正規化交互相關函數

正規化交互相關函數已經廣泛地用於評估兩張影像或兩個區塊間的相似度。其優點為不易受到影像間光線強度線性變化的影響。故本論文利用正規化交互相關函數評估兩個區塊內容之相似度。正規化交互相關函數可以表示如下：

$$NCC(i, j, d) = \frac{\sum_{\mu=-w}^w \sum_{\nu=-w}^w (I_l(i+\mu, j+\nu) - \bar{I}_l) \cdot (I_r(i+\mu-d, j+\nu) - \bar{I}_r)}{\sqrt{\sum_{\mu=-w}^w \sum_{\nu=-w}^w (I_l(i+\mu, j+\nu) - \bar{I}_l)^2} \sqrt{\sum_{\mu=-w}^w \sum_{\nu=-w}^w (I_r(i+\mu-d, j+\nu) - \bar{I}_r)^2}} \quad (4)$$

其中， (i, j) 為影像座標， (u, v) 為樣板視窗座標， d 是視差， I_l 、 I_r 分別表示左、右眼影像， $\bar{I}_l$ 、 $\bar{I}_r$ 是在樣版視窗中的灰階強度平均值。 w 為樣板視窗大小，本論文中 w 設定為 2。特徵 f_1 可以計算如下：

$$f_1(i, j, d) = 1 - NCC(i, j, d) \quad (5)$$

2.2.2 顏色差值

由於 NCC 僅利用亮度(luminance)資訊進行區塊匹配搜尋視差值，可能會於平滑但不同顏色之區域內產生錯誤。為了避免這種可能的錯誤，本論文利用顏色資訊解決此問題。

本論文提出之顏色特徵如下：

$$f_2(i, j, d) = \frac{MAD(i, j, d)}{255} \quad (6)$$

其中，

$$MAD(i, j, d) = \begin{cases} \frac{1}{3}(|\overline{R_l} - \overline{R_r}| + |\overline{G_l} - \overline{G_r}| + |\overline{B_l} - \overline{B_r}|) & \text{if } |\overline{R_l} - \overline{R_r}| < Thd1 \ \& \ |\overline{G_l} - \overline{G_r}| < Thd1 \ \& \ |\overline{B_l} - \overline{B_r}| < Thd1 \\ 255 & \text{otherwise} \end{cases} \quad (7)$$

$\overline{R_l}, \overline{R_r}$ 為區塊內紅色資訊的平均值，

$\overline{G_l}, \overline{G_r}$ 為區塊內綠色資訊的平均值，

$\overline{B_l}, \overline{B_r}$ 為區塊內藍色資訊的平均值。

2.2.3 空間相關性

考慮到視差圖的特性，除了物件邊緣之外，同一個物件內的深度資訊應該具有連續性，不會發生突然的變化，故本論文也考慮此空間相關性進行視差估計。

考慮視差估計之順序，本論文之空間相關性定義如下：

$$f_3(i, j, d) = 1 - \frac{1}{4} \sum_{m, n \in \{0, -1, 1\}} U(|d(i, j) - d(i+m, j+n)| < Thd2) \quad (8)$$

其中， $U(\cdot)$ 表示 unit step function。

2.3 視差修正

為了確定成本函數所取得的深度圖不被遮蔽所影響，所以可以利用左右一致性的幾何方式來幫助判斷視差的可靠度。其方式是以計算出來的視差對應到另一張圖片的對應位置，如果這組對應視差相同，則保留這組視差，若不同則捨棄，標示為空洞點(遮蔽區域)。

經過左右一致性的處理後，左右深度圖都會產生許多判定為遮蔽點，而捨棄視差造成的空洞點，為了補上這些空洞點，此處考慮計算產生的些微誤差，將左右一致性的視差誤差調整為 3，當誤差大於 3 才真正捨棄視差值。

若不考慮視差中斷的位置，當遮蔽在已知點中間，採用內插的方式計算。若遮蔽點不在中間，則採取先前計算出的已知點作外插。

當空洞點發生時，這裡參考邊緣偵測演算法(Sobel)所得到的邊緣圖為視差中斷處，並分別依據以下幾種情況，來去做對應的運算：

情況 1. 若前兩點都不是邊界，則使用前兩點的像素資訊來外插。

情況 2. 若只有前一點不在邊界，往下搜尋有左右一致的點，則作內插。

情況 3. 如果只有前一點不在邊界，往下搜尋直到邊界點都沒有參考點，則直接取前一點的視差值。

情況 4. 當前一點位於邊界，則往下搜尋非空洞點來當視差。

情況 5. 如果前一點在邊界上，而往下搜尋到邊界點之前都沒有視差值可參考的話，則直接拿正上方上一條掃描線上的視差值。

最終，可取得處理遮蔽後的左右視差圖。

3. 實驗結果

本篇論文以 middlebury 網站上所提供的 2 組左、右影像為測試資料，影像尺寸均為 450 × 375 個像素點。本實驗中成本函數的參數權重 α_1 、 α_2 和 α_3 ，分別設定為 0.93、0.05 以及 0.02。Thd1 和 Thd2 設定為 25 和 5。Cone 與 Teddy 的視差搜尋範圍均設定為 60。為了得知計算結果的正確性，根據文獻 [17]，本論文採用之評估因子如下：

均方根誤差 (root mean squared error, RMSE)

$$RMSE = \left(\frac{1}{N} \sum_{(x,y)} |d_c(x,y) - d_T(x,y)|^2 \right)^{\frac{1}{2}} \quad (9)$$

其中， N 表示像素點的數目， d_c 、 d_T 分別代表計算求得的視差圖以及真實的視差圖(ground truth)。

錯誤配對像素百分比 (Percentage of bad matching pixel, PBM)

$$PBM = \frac{1}{N} \sum_{(x,y)} (|d_c(x,y) - d_T(x,y)| > \delta_d) \quad (10)$$

其中， δ_d 是一個門檻值，後面實驗將它設為 1。

首先，我們評估重建影像之視覺品質，並且與其他方法做比較。圖 2 和 3 分別顯示 Cones 和 Taddy 之重建影像。由圖 2(c)(d) 和圖 3(c)(d) 可以看出，基於 NCC 之區塊匹配法則獲得之視差值，易產生錯誤並且造成重建影像視覺品質不佳。由圖 2 和 3 可以觀察出，本論文之重建影像的視覺品質優於其他方法之重建影像的品質。

表 2 列出非遮蔽區域內 PSNR 和 RMSE。表 3 列出非遮蔽區域內 PBM 之結果。由表 2 可以看出，雖然本論文之 PSNR 略低於 SAD 之 PSNR，但是本論文之 RMSE 高於 SAD 之 RMSE。另外，在 $\delta = \pm 1$ 之條件下，本論文之 PBM 結果優於 SAD 和 NCC 之 PBM 結果。這些結果顯示，本論文提出之視差估計法則可以獲得良好的視差圖。

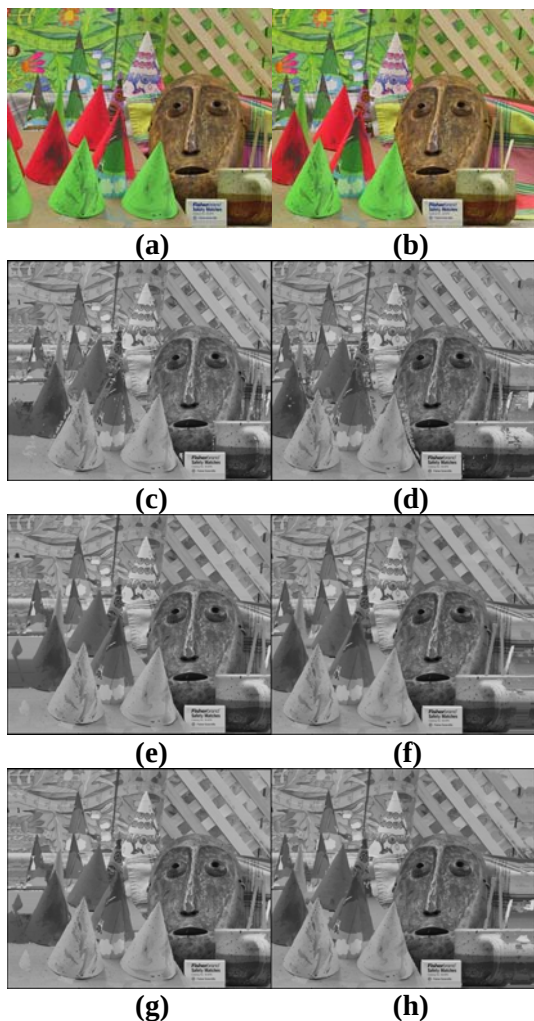


圖 2、(a)原始左圖，(b)原始右圖，(c) NCC 重建之左圖，(d) NCC 重建之右圖，(e) SAD 重建之左圖，(f) SAD 重建之右圖，(g)本論文重建之左圖，(h) 本論文重建之右圖

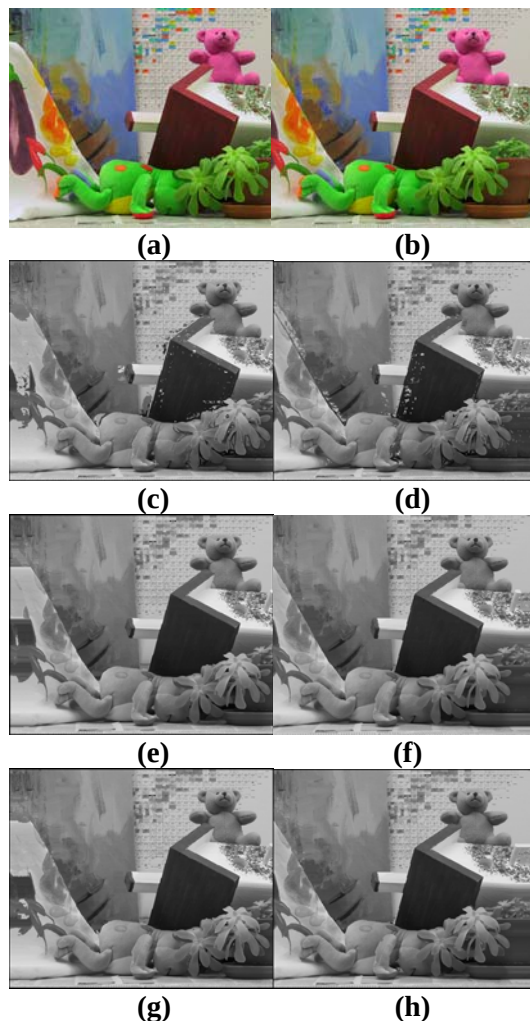


圖 3、(a)原始左圖，(b)原始右圖，(c) NCC 重建之左圖，(d) NCC 重建之右圖，(e) SAD 重建之左圖，(f) SAD 重建之右圖，(g) 本論文重建之左圖，(h) 本論文重建之右圖

4. 結論

本論文提出一個基於區塊匹配之視差估計法則。此視差估計法則包含三個部份：前處理、視差估計、視差修正。前處理包含色彩轉換和邊緣偵測。為了估計視差值，本論文基於正規化交互相關函數，顏色特徵和空間特徵設計一個成本函數。透過此成本函數，一張初始視差圖可以獲得。為了獲得遮蔽區域內之視差值，本論文提出一個視差修正法則。

根據實驗結果，對於非遮蔽區域，本論文提出之視差估計法則可以有效獲得視差值。相較於傳統 NCC 和 SAD 區塊匹配法則，本論文提出之法則的效能較佳。

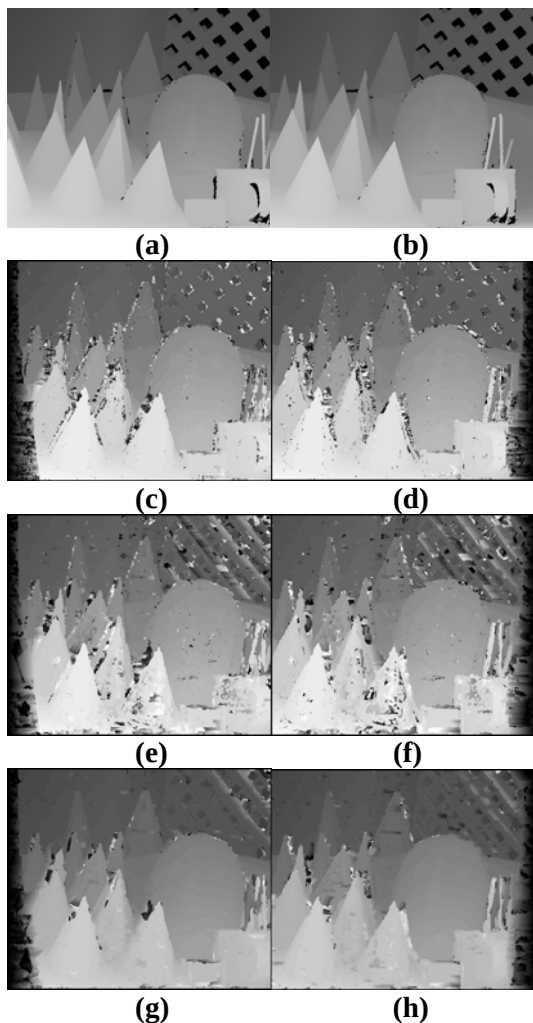


圖 4、Cones 實驗結果：(a)左真實視差圖，(b)右真實視差圖，(c) NCC 左視差圖，(d) NCC 右視差圖，(e) SAD 左視差圖，(f) SAD 右視差圖，(g) 本論文左視差圖，(h) 本論文右視差圖

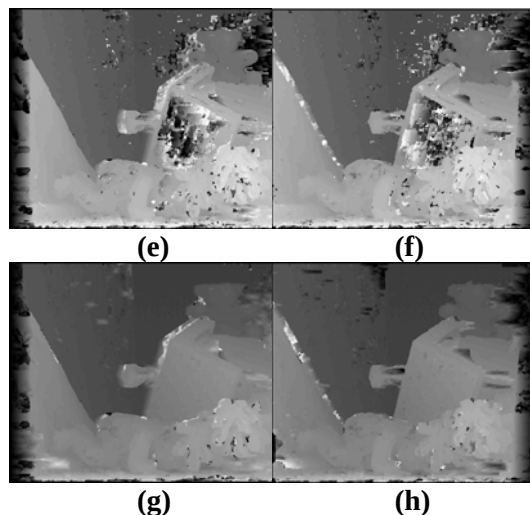
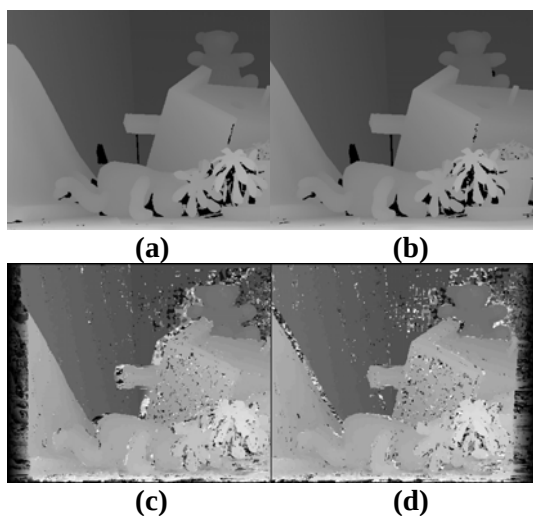


圖 5、Taddy 實驗結果：(a)左真實視差圖，(b)右真實視差圖，(c) NCC 左視差圖，(d) NCC 右視差圖，(e) SAD 左視差圖，(f) SAD 右視差圖，(g) 本論文左視差圖，(h) 本論文右視差圖

參考文獻

- [1] Rafael C. Gonzalez and Richard E. Woods, **Digital Image Processing, Third Edition**, Pearson Prentice Hall, 2010.
- [2] Asmaa Hosni, Michael Bleyer, Margrit Gelautz and Christoph Rhemann, "Local stereo matching using geodesic support weights," **IEEE International Conference on Image Processing**, pp.2093-2096, 2009.
- [3] Mohammadjavad Abdollahifard, Karim Faez and Mohammadreza Pourfard, "Fast stereo matching using two stage color-based segmentation and dynamic programming," **International Symposium on Mechatronics and its Applications**, pp.1-6, 2009.
- [4] A.K. Riemens, O.P. Gangwal, B. Barenbrug and R.-P.M. Berretty, "Multi-step joint bilateral depth upsampling," **Visual Communications and Image Processing**, pp.1-12, 2009.
- [5] An-Chun Luo, Wen-Chao Chen, De-Jin Shau and Chung-Wei Lin, "Occlusion size aware multi-viewpoint images generation from 2D plus depth images," **The International Society for Optical Engineering**, pp1-11, 2010.

- [6] Jan Cech and Radim Sara, "Efficient Sampling of Disparity Space for Fast and Accurate Matching," **IEEE Conference on Computer Vision and Pattern Recognition**, pp.1- 8, 2007.
- [7] Andrea Fusiello, Vito Roberto and Emanuele Trucco, "Efficient Stereo with Multiple Windowing, " **IEEE Computer Society Conference on Computer Vision and Pattern Recognition**, pp.858-863, 1997.
- [8] Heiko Hirschmüller, "Improvements in Real-Time Correlation-Based Stereo Vision," **Stereo and Multi-Baseline Vision**, pp.141-148, 2001.
- [9] Ilkwon Park and Hyeran Byun, "Depth map refinement using multiple patch-based depth image completion via local stereo warping," **Optical Engineering**, pp.1-8, 2010.
- [10] Ju Yong Chang, Kyoung Mu Lee and Sang Uk Lee, "Stereo matching using iterative reliable disparity map expansion in the color-spatial-disparity space, " **Pattern Recognition**, Vol.40, pp.3705-3713, 2007.
- [11] Jie Zhao, Jiali Chai and Guozun Men, "A Fast Quasi-Dense Matching Algorithm with an Adaptive Window," **Chinese Control and Decision Conference**, pp.4375-4379, 2009.
- [12] Kun Wang, "Adaptive stereo matching algorithm based on edge detection, " **International Conference on Image Processing**, pp.1345-1348, 2004.
- [13] Andres Klaus, Mairo Sormann and Konrad Karner, "Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure, " **International Conference on Pattern Recognition**, pp.15-18, 2006.
- [14] Liang Zhang, "Fast Stereo Matching Algorithm for Intermediate View Reconstruction of Stereoscopic Television Images," **IEEE Transactions on Circuits and Systems for Video Technology**, Vol. 16, pp.1259-1270, 2006.
- [15] Li Tang, Chengke Wu and Zezhi Chen, "Image dense matching based on region growth with adaptive window," **Pattern Recognition Letters**, Vol.23, pp.1169-1178, 2002.
- [16] Marcus Mueller, Frederik Zilly and Peter Kauff, "Adaptive cross-trilateral depth map filtering," **3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video**, pp. 1-4, 2010.
- [17] Pedro F. Felzenszwalb and Daniel P. Huttenlocher, "Efficient belief propagation for early vision," **International Journal of Computer Vision**, Vol.70, pp.55-79, 2006.
- [18] Xiaoyong Lin, Yu Liu and Wenzhan Dai, "Study of Occlusions Problem in Stereo Vision," **World Congress on Intelligent Control and Automation**, pp.5062-5067, 2008.
- [19] Zheng Gua, Xianyu Su, Yuankun Liu and Qican Zhang, "Local stereo matching with adaptive support-weight, rank transform and disparity calibration, " **Pattern Recognition Letters**, Vol.29, pp.1230-1235, 2008.

表 1 實驗平台

CPU	Intel Core i5 650 Turbo3.46GHz 3.2GHz
RAM	DDR3 4.0 GB
O.S.	Microsoft Windows XP SP3
Compiler	Borland C++ Builder 6 Enterprise Suite
Screen	22" wide iZ3D monitor

表 2 左圖非遮蔽區的實驗結果

影像名稱	方法	合成 PSNR(dB)	RMSE(pixel)
Cones	NCC	27.07	5.40
	SAD	30.98	6.18
	Proposed	30.03	4.09
Taddy	NCC	27.68	6.15
	SAD	33.27	6.71
	Proposed	32.22	4.44

表 3 左圖非遮蔽區之 PBM 分析

影像名稱	方法	± 1	± 2	± 3	± 4	± 5
Cones	NCC	85.12	49.29	14.16	7.45	6.84
	SAD	84.06	51.79	21.36	14.49	12.58
	Proposed	83.63	49.91	16.36	8.31	6.46
Taddy	NCC	76.57	31.59	14.11	11.86	10.62
	SAD	75.75	35.15	20.60	17.20	15.23
	Proposed	74.68	31.31	13.75	9.62	7.33