

由最佳散焦影像序列測距

暨南大學資訊工程學系

指導教授 石勝文

研究生 郭維新

中華民國九十年六月二十八日

目錄

1	序言	9
2	簡介	11
3	影像的生成	16
3.1	凸透鏡的成像	16
3.2	光源的散佈函式	17
3.3	模糊半徑的介紹	24
4	散焦測距模組	26
4.1	散焦測距模組的誤差函式	26
4.2	搜尋最佳成像點的方法	30
5	決定最佳的相機參數	33
5.1	散焦測距估測誤差之 Cramér-Rao 下限	34
6	最佳散焦測距方法	39
6.1	Asymptotic Bayesian 遞迴式估測法	39

7	實驗結果與討論	42
7.1	測試平行平面資訊量測	45
7.2	測試傾斜平面資訊量測	48
7.3	測試複雜傾斜平面資訊量測	52
8	結論與未來展望	55

圖例目錄

1	凸透鏡的成像	16
2	光學轉換函式及影像感測效應在頻率領域的散佈情形	21
3	不同影像失焦在空間領域的散佈情形	22
4	不同影像失焦在頻率領域的散佈情形	23
5	以高斯函式近似光源散佈函式在頻率領域的散佈情形	23
6	透鏡的聚焦	24
7	最佳成像點較靠近 I_2 的情形	32
8	最佳成像點較靠近 I_1 的情形	32
9	誤差函式	32
10	散焦影像序列	33
11	估測相對模糊程度的 Fisher information matrix 值	37
12	估測相對模糊程度的 CRLB	38
13	Asymptotic Bayesian 遞迴式估測法示意圖	41
14	相對模糊程度低於下界限定值時之變異量變化圖	43

15	相對模糊程度低於下界限定值時，估測的相對模糊與實際的相對模糊之比較。	44
16	相對模糊程度高於下界限定值時，估測的相對模糊與實際的相對模糊之比較。	44
17	相對模糊程度高於下界限定值時之變異量變化圖	45
18	平行平面深度相關示意圖	46
19	測試平行平面貼圖影像	47
20	恢復平行平面深度影像	47
21	傾斜平面深度相關示意圖	48
22	測試傾斜平面貼圖影像	49
23	測試傾斜平面散焦影像	49
24	兩張影像：恢復傾斜平面深度影像	50
25	多張影像：恢復傾斜平面深度影像	50
26	複雜傾斜平面深度相關示意圖	52
27	測試複雜傾斜平面散焦影像	53
28	兩張影像：恢復複雜傾斜平面深度影像	53
29	多張影像：恢復複雜傾斜平面深度影像	54

中文摘要

在電腦視覺研究領域中散焦測距是非常重要的視覺模組。散焦測距的方法最早是在 1987 年由 Pentland 提出，它最少只需要兩張不同鏡頭設定底下所攝得的影像即可計算距離資訊。散焦測距系統大致上分為兩個處理的階段，第一個階段主要是攝影機的校正工作，第二階段為重建深度的方法，本論文中主要在探討第二階段重建深度方法的誤差分析及最佳化的方法。我們依據 Ens 和 Lawrence 提出的散焦測距演算法為基礎，加以改良成一個新的散焦測距方法，而新方法中也同時整合了散焦、對焦測距模組的優點。接著推導出散焦測距相對模糊程度間的 Cramér-Rao lower bound(CRLB)，並在使 CRLB 最小的準則下推導出最佳化相機參數值。由散焦測距相對模糊間的 CRLB 可知當散焦影像愈清晰則深度量測愈精準。但是在一般的場景中，物體總是分佈在不同的距離，而因為不同距離的物體會有不同的最佳成像點，所以其最佳化相機參數值也會不同，因此為了求得最佳深度估測值，我們必須以多張影像來進行散焦測距計算，並且我們也以 asymptotic Bayesian 遞迴式估測法整合多張影像的資訊。根據實驗結果顯示，使用最佳化的影像序列進行散焦測距將可大幅提昇三維估測的精準度。

關鍵詞：電腦視覺、散焦測距、Cramér-Rao 下界、最佳化相機參數

Abstract

Depth from defocus (DFD) is a very important visual module in the field of computer vision. It was first introduced by Pentland in 1987. The DFD technique requires only two images acquired at two different lens settings to compute depth estimates. In general, implementing a DFD system involves two major tasks, namely, calibration of the DFD system, and development of an accurate depth estimation method. This work is mainly related to the second task. Our depth estimation method is modified from the DFD algorithm proposed by Ens and Lawrence. In order to minimize the depth estimation error, we have analytically derive the Cramér-Rao lower bound of the blur estimation error. According to the error analysis result, we found that the depth estimation error can be minimized if one of the images is clearly focused. However, since objects in a scene usually distributed in a wide range of depth, different parts of the acquired image would become clearly focused with respect to a lens settings. Therefore, we have developed a method to compute a sequence of optimal lens settings so that the image sequence acquired with the optimal lens settings can be used to accurately estimate the corresponding depth information. We have also developed a recursive depth estimation method based on the asymptotic Bayesian computation framework. Experimental results have shown that the proposed method can dramatically improve

the depth estimation accuracy from the optimal defocused image sequence.

KeyWord: Computer Vision, Depth from Defocus, Cramér-Rao lower bound, Optimal Camera Parameter.

1 序言

在國立暨南國際大學資訊工程研究所兩年的學習生涯中，讓我在專業知識的提升，及人際關係的歷鍊都有非常大的收穫。初到這個有好山好水的學校沒多久就遭遇到罕有的大地震，歷經了天搖地動的劇變後，讓我們這些剛來學校報到的暨大資工所第一屆的學生，很快的就建立起彼此間信賴的關係，使得往後的求學日子裡，同學間都會互相勉勵打氣，共同討論學習。在地震過後克難的環境中，慶幸得到我的恩師石勝文教授的指導，石老師不但在日常生活中對學生照顧有加，在學業上更是孜孜不倦，不厭其煩的教導我，使我獲益良多，在這裡我要向石老師獻上我最大的敬意和謝意。同時我要感謝的還有吳育德老師，在求學的階段中，吳老師不但用心教導了我們專業上的知識，而且也很關心我們的日常生活狀況，實在也是位不可多得的良師。最後我要感謝的就是我最偉大的父親郭義雄先生、母親王富美女士，以及我最敬重的兄長郭雨新先生。當地震發生後寄讀於台大的日子，幸得兄長的關心和照顧使我可以專心致力於課業上的學習。匆匆研究所兩年的時光，箇中的甜、酸、苦、辣卻也不是三言兩語所能道盡，當我面臨許多挫折、失敗，感到沮喪、徬徨的時候，幸虧父母親都會在背後支持我，給我鼓勵，如果沒有您們就沒有今天的維新。經過了兩年的努力，我完成了這次畢業論文的相關研究，在往後的日子我一定會更加用心、努力以克服各種可能會面臨的挑

戰，不負父母，以及所有關心我的人的期望。

2 簡介

恢復影像中物件的三維資訊一直以來都是電腦視覺研究領域中非常重要的一環，有許多恢復影像中物件深度的方法(例如：立體的視覺 (stereo vision)[2]，由運動估測三維結構 (structure from motion)[15]，散焦測距模組 (depth from defocusing)[6][25]) 都已經在許多國際上的文獻被驗證過了，這些技術可以應用到三維空間的重建、影像物件的辨識等許多的層面，並且帶給了我們非常多有用的幫助。在 1987 年 Pentland[18] 首先提出散焦測距法 (Depth from Defocus)，由此又提供了重建三維資訊的一個新方法，並且帶動了往後許多散焦測距模組的相關研究。

過去在電腦視覺重建三維空間的研究領域中，都是由立體視覺的方法為主，但是利用此種方法重建三維空間時，如果物體距離攝影機太近，就會因為立體影像中物體的視差過大，而導致立體對應計算結果容易出錯。由於散焦測距模組在近距時優異的表現，並且演算法簡單、只需處理單眼影像 (Monocular Image)，所以散焦測距這個視覺模組一直是電腦視覺研究領域中非常受到重視的視覺模組。因為演算法簡單，所以運算速度比起其它同精確度的視覺模組還要快，而且只需要單眼影像，所以可以很容易地避開目前仍無有效解決方法的立體對應問題；同時也可以用在處理具有鏡面材質的物面重建工作上。雖然在處理遠距的物面重建工作時誤差會變得比較大，

但是這些遠距量測值也可提供立體視覺或其它的視覺模組作初始值用。另外在近距離時，由於攝影機鏡頭的景深會變得非常淺，因此僅能有部份的影像區域可以清晰地聚焦，此時散焦測距便可以用來提供非常精確的三維量測值（在某些特殊的應用中，散焦測距甚至可以提供解析度高至 $8\mu\text{m}$ 的距離資訊 [26]）。Pentland 是最早提出散焦測距演算法的研究者，他的散焦影像取樣方法是將攝影機鏡頭孔徑調到最小（針孔攝影機）取得第一張影像接著將孔徑放大攝取另一張不同模糊程度的影像，並且以 inverse filtering 的方法來計算近似為高斯函式的 PSF (point spread function，詳見 3.2 節)，並由此估測出物體深度 [18]。Subbarao 放寬了 Pentland 方法其中一張影像必須使用針孔相機取像的限制，並允許可以同時變動不同的鏡頭設定（對焦、光圈），仍能估算出深度資訊 [23]。Hwang 等人基於 PSF 為高斯函式的假設，在空間領域中發展估算模糊參數的演算法 [25]，他以微幅改變攝影機參數的方式，計算出攝影機參數的微分近似值，並推導出其與影像的 LOG (Laplacian of Gaussian) 影像模糊程度的關係式，從而求解出影像深度。Ens 和 Lawrence 提出了利用預先計算好的濾波器表，逐一驗證每一個濾波器，以找出最合適的濾波，再查出這個濾波器相對的深度。他們也驗證了這個方法估算深度的結果要比 inverse filtering 方法更為精準 [5]。Surya 和 Subbarao 以一個在空間領域經由 deconvolution (S 轉換) 的運算方式重建出散焦影像中物體的深度 [24]。Klarquist 等人提出了利用多張散焦影像作散焦測距的 Maximum-Likelihood 估

測，他們所定義的問題與以往所有的散焦測距法的問題定義並不相同，可以說是一個全新的散焦測距法，可惜在他們的方法中不只要求最佳成像點的位置，同時也將整張清晰聚焦的影像當成是待求的未知數 [12]。這種作法不但會增加運算時間，同時也因為未知數變多而可能使估測值更容易受到雜訊的影響。Rajagopalan 和 Chaudhuri 在 [19] 將影像分成數個子影像，其中各鄰近子影像互有重疊，再以位移變動的區塊 (block shift variant blur model) 作計算，解決因為 PSF 在影像的鄰近區域間造成重疊的問題。另外他們在影像為 AR(autoressive)[10]process 的假設下以 CRLB[21] 計算出散焦測距模組最佳化相機參數值 [20]。由許許多多的研究者投入散焦測距模組的研究看來，更加顯示出這方面相關研究的重要性。

對焦測距模組和散焦測距模組同樣對近距離深度資訊量測有優異的表現。它們的不同處在於散焦測距模組只需要兩張不同模糊程度的散焦影像即可計算距離資訊；而對焦測距模組需要搜尋整個鏡頭設定值來求得距離資訊，這兩個視覺模組依循不同的演算法，前者快而準確度稍低；後者準確度較高，不過需要經過大量的計算，而且計算後所得的資料又大半被棄置不用。許多對焦測距研究者為避免以線性搜尋最清晰成像位置所造成的費時間問題，開始以非線形搜尋法減少搜尋的時間，他們的辦法是利用費式搜尋法 (Fibonacci Search) 將最佳成像點大略找出，而後假設在最佳成像點附近，影像清晰度相對於鏡頭的位置可以利用一個二次多項式來近似。而利用二次多

項式的內插就可求出最佳成像點的位置 [13, 28]。我們認為散焦、對焦測距應該可以被整合成一個又準確又快速的視覺模組。隨著科技的不斷演進，我們可以預期會有越來越快的電控鏡頭出現，從而縮短在不同鏡頭設定取得不同模糊程度影像的時間，因此我們可以利用多一點影像來提昇三維估測的準確度，不再限制自己只能取兩張影像作散焦測距。又如果我們可以事先知道最佳化的相機參數值，並且透過精密的校定程序而得到影像模糊度與鏡頭位置之間的關係，因此就能知道需要那幾張影像即可得到精確的最佳成像點，那就避免了類似對焦測距模組搜遍整個鏡頭設定空間來決定最佳成像點的問題。在這篇論文中，我們整合了散焦、對焦測距法的優點，並且以推導的最佳化相機參數值，提供一個最佳化的散焦測距模組。

本論文的架構如下：在第三章中我們介紹空間中的物體如何經由光學系統成像，並且討論影響物體影像模糊程度的光源散佈函式，及此光源散佈函式在研究的過程中如何以高斯函式取代之。第四章介紹散焦測距模組，我們提出一個新的計算方法，經由這個計算方法找出散焦影像間的關係，接著利用費式搜尋法及二次多項式內插法，搜尋最佳成像點的位置。第五章中我們以 CRLB 及受取樣定理限制之下 [16]，推導了散焦測距模組的最佳化相機參數值。第六章我們藉由最佳化相機參數值，決定了影像平面擺放位置，及擷取的散焦影像數量，經由這些最佳化的資訊使我們散焦測距模組能夠計算出最精確的結果。同時為避免一次需要處理所有影像資料的龐大計算量，並且

要求隨取隨用不同鏡頭設定的參數以加快運算時程，我們使用以 asymptotic Bayesian [9] 遞回式估測法 (recursive estimator) 爲主的整合式散焦測距模組，依據驗證的結果更進一步最佳化我們的散焦測距方法。第七章中我們以電腦模擬的方式對由最佳散焦影像序列測距做不同深度變化情形的驗證。第八章中我們對由最佳散焦影像序列測距做結論並且提出我們未來的研究方向。

3 影像的生成

影像的生成是因為物體上難以計數的點光源，散發出的光線穿過我們的相機鏡頭，落在影像平面上最後形成我們想要得到的影像。爲了知道這無數個點光源的成像原理，就必須要先知道一個點光源的成像情形，所以下一節我們將單就一個點光源經由相機鏡頭的成像情形作更深入的說明。

3.1 凸透鏡的成像

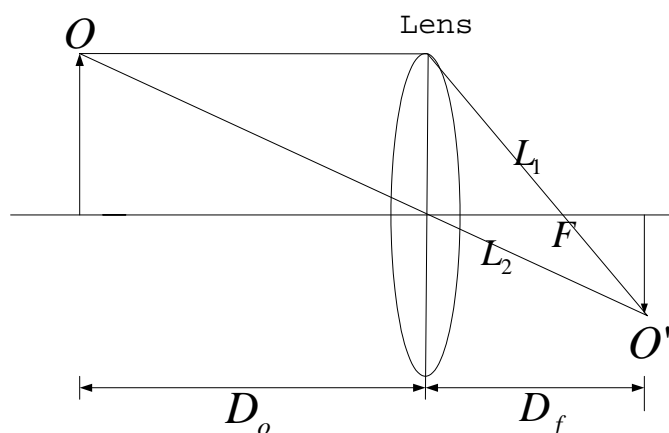


圖 1: 凸透鏡的成像

圖 1 是凸透鏡的成像示意圖， L_1 爲點光源 O 它所發出平行於透鏡軸心的光線穿過透鏡後折向焦點的光線， L_2 爲點光源穿過鏡心的光線， L_1 和 L_2 的相交點就是 O 的最清晰成像點 O' 。假設此一透鏡之焦距長度爲 F ，點光源距離透鏡中心 D_o (物距)，最清晰成像點 O' 距離透鏡中心爲 D_f (像距)，我們可以列出透鏡關係式：

$$\frac{1}{D_f} + \frac{1}{D_o} = \frac{1}{F} \quad (1)$$

在本論文中我們將像距 D_f 稱為“最佳成像點”，因為若將影像感測器放置於這個最佳成像點時就可以得到最清晰的點光源影像，如果相機焦距長度 F 與最佳成像點 D_f 的位置皆為已知，我們就可以求算出我們想要得到的空間深度值 D_o :

$$D_o = \frac{F \cdot D_f}{D_f - F} \quad (2)$$

3.2 光源的散佈函式

從 3.1 知道了一個點光源的成像的方式，可是在實際的情形中即使不考慮透鏡之球面像差 (spherical aberration)、色差 (chromatic aberration) 且在最佳成像點的地方取像，所得的影像卻還是會有些許模糊的現象，所以在這一小節裡我們討論會影響影像模糊的因素。

1. 光學轉換函式 (OTF): 一般我們稱較佳的光學系統指的是這個光學系統有較高的辨識率，然而大多數的透鏡都呈圓孔狀，影響到辨識率高低的主要成因也就是因為圓孔繞射的現象。物體上兩個距離相近的點光源 O_1 、 O_2 散發出的光線經過透鏡，產生圓孔繞射的現象，投射在影像平面後成 O'_1 、 O'_2 兩個繞射影像點，當 O_1 、 O_2 的距離越接近 (頻率越

高)，或透鏡孔徑越小的話， O'_1 、 O'_2 兩個點的範圍就會變大，進而產生重疊的現象，也就是說這時已經不能辨識出由透鏡所產生的影像 O_1 、 O_2 了。所以在辨識率允許的範圍下才能夠辨識出相近的不同點光源，我們通常以角度 θ_d 來表示辨識率的高低 [7]:

$$\theta_d = 1.22 \frac{\lambda}{L} \quad (3)$$

其中 λ 為光的波長， L 為鏡頭的直徑，由 (3) 式，可以知道辨識率和光波的波長成正比例關係，並且和孔徑的大小成反比例的關係，所以在辨識率的範圍限制之下，我們將點光源經由圓孔繞射後的成像關係式 (光學轉換函式) [14]，在頻率領域中以下式表示：

$$O(u, v; L, F) = \begin{cases} \left(\frac{L}{2F}\right)^2 (\gamma - \sin(\gamma)), & \sqrt{u^2 + v^2} \leq \frac{L}{\lambda F} \\ 0, & \sqrt{u^2 + v^2} > \frac{L}{\lambda F} \end{cases} \quad (4)$$

其中從空間領域的座標位置轉換成頻率領域中的座標以 (u, v) 表示， F 為鏡頭的焦距， L 為鏡頭的直徑，其數值可由鏡頭生產商所提供的 f-number 計算而得：

$$L = \frac{F}{f_{num}} \quad (5)$$

(4) 式 γ 表示成：

$$\gamma = 2 \cos^{-1} \left(\lambda \cdot f_{num} \cdot \sqrt{u^2 + v^2} \right) \quad (6)$$

2. 影像感測器之效應 (Image Sensor): 影像感測器最主要的功能在於將光線穿過鏡頭後的成像光子轉換成電子，可以供我們儲存、處理或是展示。現今最常用的影像感測器是 CCD(charge-coupled device)，而由於 CCD 感測器的解晰度有限，所以取得的影像會有些許的模糊情形。這是因為 CCD 像素面積並非無窮小，所有打在同一像素的光線效果互相疊加，這個效應在頻率領域中可以由下式描述：

$$S(u, v) = w_x \frac{\sin(\pi w_x u)}{\pi w_x u} \cdot w_y \frac{\sin(\pi w_y v)}{\pi w_y v} \quad (7)$$

其中 (u, v) 表示於頻率領域中的座標， w_x 、 w_y 為 CCD 照相機中每個影像點的長和寬。

3. 影像失焦 (Defocusing): 以上介紹的兩點使影像模糊的因素都是來自於硬體的設備，並且這就是為什麼即使我們在最佳成像點處取像，所得的影像仍有些許模糊的原因。而現在介紹的影像失焦的情形則是在取像的過程中由我們自己調整相機的參數值 (effective focal length)，也就是變動影像平面的擺放位置，這個散焦模糊函式於頻率領域中可表示為：

$$B(u, v, \delta D_i) = \frac{f_{num}}{\delta D_i \cdot \sqrt{u^2 + v^2} \cdot \pi} J_1 \left(\frac{\delta D_i \cdot \sqrt{u^2 + v^2} \cdot \pi}{f_{num}} \right) \quad (8)$$

其中 u 、 v 為頻率領域中的座標， δD_i 代表變動過後影像平面位置和最佳成像點之間的距離， J_1 代表一階的 Bessel function[1]。

匯集以上三點因素所形成的模糊函式就是影響光源散佈的函式 (Point Spread Function 簡稱 PSF)，即點光源在影像平面上呈像函式，可表示為：

$$PSF(u, v, s) = O(u, v; L, F) \cdot S(u, v) \cdot B(u, v, \delta D_i) \quad (9)$$

其中 s 為模糊參數。而在感測器上成像可以表示成原始點光源像與 PSF 函式的卷積運算 (convolution integral 本論文中 $*$ 符號亦同此意)：

$$I_i(x, y) = I_o(x, y) * PSF(x, y, s) \quad (10)$$

其中 I_i 為影像感測器測量到光源的光強度， I_o 為物體發出的光強度。下圖 2 為光學轉換函式及影像感測效應在頻率領域所造成的影像光亮強度的變化情形，圖 2 上方的是光學轉換函式波形 (我們模擬綠色光波，波長為 546.1nm，並且假設 f-number 為 5.6) 取其垂直方向的對數值後觀察於 $-200mm^{-1}$ 至 $200mm^{-1}$ 間的散佈變化情形；圖 2 下方是影像感測效應的波形 (我們假設空間中點的長和寬都是 $10^{-2}mm$)，同樣也是取其垂直方向的對數值後觀察 $-200mm^{-1}$ 至 $200mm^{-1}$ 間的散佈變化情形。由此可以看出影像感測效應影響影像模糊的程度會比光學轉換函式的影響來的高。

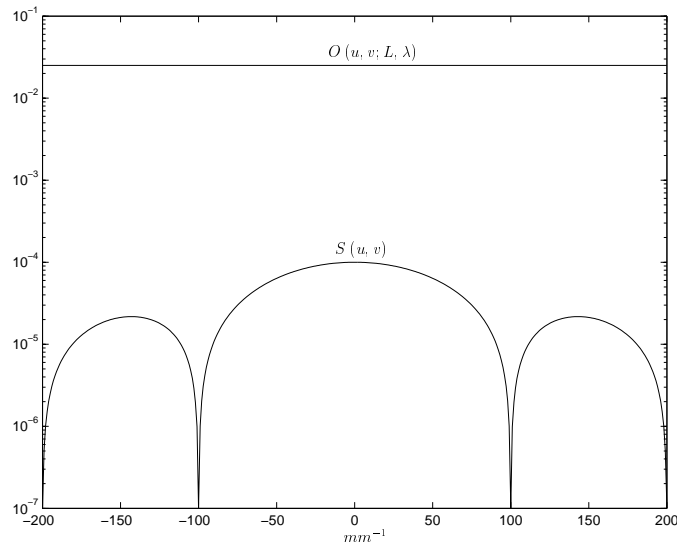


圖 2: 光學轉換函式及影像感測效應在頻率領域的散佈情形

圖 3 是將影像平面擺放在距離最佳成像點不同位置時 (假設影像平面擺放位置分別為距離最佳成像點 0.1mm 及 0.5mm) , 點光源在影像中 (空間領域) 變模糊之後展開的波形切面圖。當影像平面擺放距離最佳成像點愈近, 則光源成像較清晰, 點光源發散的亮度及分佈範圍像一個細長的圓柱; 反之當影像平面擺放距離最佳成像點愈遠, 則光源成像較模糊, 點光源發散的亮度及分佈範圍像一個寬扁的圓柱。圖 4 則是圖 3 在頻率領域的對應, 我們可以看到在空間領域中較寬扁的模糊函式, 在頻率領域中主要的波段變得較窄, 可通過的高頻訊號有限; 反之在空間領域中較細長的模糊函式, 在頻率領域中主要的波段變得較寬, 大多數的訊號都可以通過。當光學轉換函式, 影像感測效應及影像失焦情形這三個在頻率領域中的函式相乘後既是我們的光源散佈函式, 依據中央極限定理此光源散佈函式會趨近於高斯函式, 而且

前人在散焦測距的相關研究上也多是採用高斯函式，所以我們這次的研究當中也採用高斯函式作為透鏡的 PSF，PSF 在頻率領域中以高斯函式來近似：

$$PSF(u, v, s) = E^{-\frac{1}{2}s^2(u^2+v^2)} \quad (11)$$

其中 (u, v) 表示於頻率領域中的座標， s 代表模糊程度的參數 (4.1 節會有更詳細的介紹)。圖 5 為一個 PSF 的理論曲線與其近似表為高斯函式之曲線，我們可以看出有較清晰光源成像的函式比較模糊光源成像的函式更近似高斯函式的曲線。

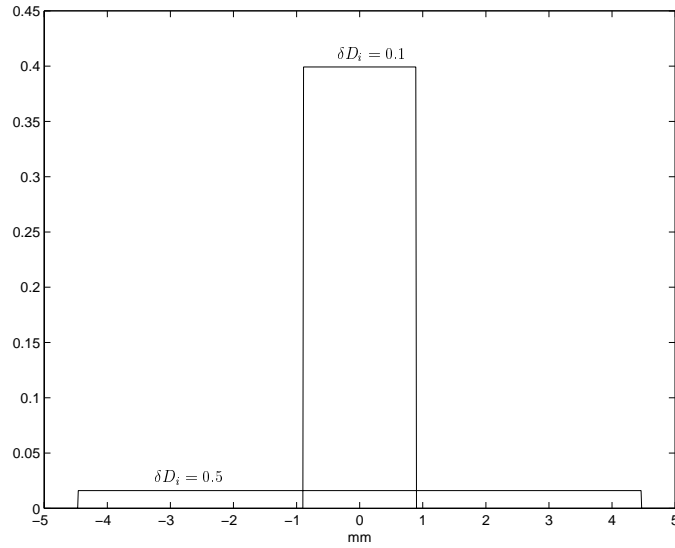


圖 3: 不同影像失焦在空間領域的散佈情形

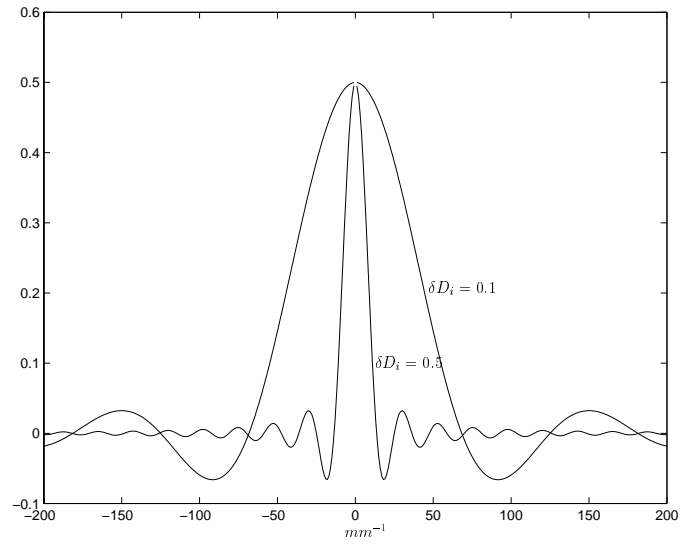


圖 4: 不同影像失焦在頻率領域的散佈情形

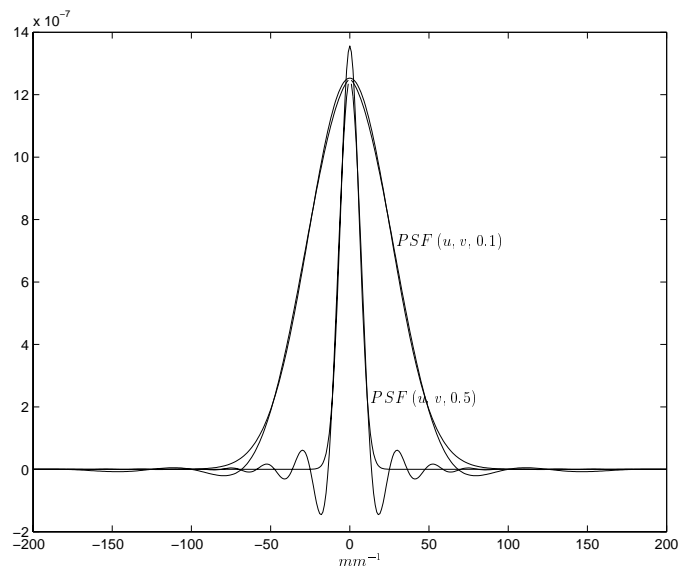


圖 5: 以高斯函式近似光源散佈函式在頻率領域的散佈情形

3.3 模糊半徑的介紹

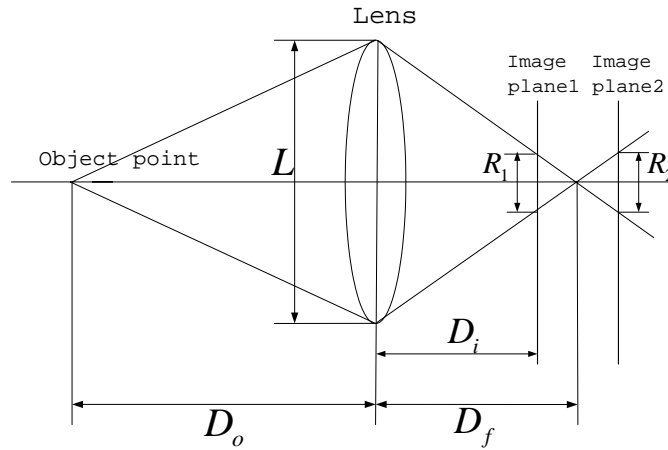


圖 6: 透鏡的聚焦

圖 6 為透鏡的聚焦示意圖。如果我們將影像感測器置於 D_f 之位置即可得到此一點光源最清晰的影像。但是如果影像感測器被置於 D_i 之位置，則獲得的點光源影像將會變得較為模糊，在不考慮其他光學因素的影響下，此一點光源將在影像平面上均勻地投射出一個圓形光亮區，而這個圓形的半徑 R_i 我們稱之為“模糊半徑”，可由下式決定：

$$R_i = \frac{L \cdot |D_f - D_i|}{2 \cdot D_f} \quad (12)$$

影像感測器只要是被置放在離最佳成像點同等距離的光軸上，則此兩張影像的模糊程度會是一樣的，所以圖 6 中 Image plane 1 和 Image plane 2 所對應的影像是模糊程度完全一樣的影像，也就是它們有相同的模糊半徑 $R_1 = R_2$ 。

並且我們也發現影像的模糊程度 (即 PSF 以高斯函式近似 (11) 中的 s) 和模糊半徑存在一個正比的關係：

$$s = M \cdot R_i \quad (13)$$

其中 R_i 代表第 i 張影像的模糊半徑。

4 散焦測距模組

在 3.1 節 (2) 式提到如果知道最佳成像點 (D_f) 的位置，則可以推算出想要知道的深度量測值 (D_o)，在這一章節裡我們將介紹用來估算深度量測值的散焦測距模組及搜尋最佳成像點的方法。

4.1 散焦測距模組的誤差函式

我們將 (5) 式代入 (12) 式中，可得：

$$R_i = \frac{|F \cdot (D_f - D_i)|}{2 \cdot f_{num} \cdot D_f} \quad (14)$$

爲了精確計算深度資訊，我們必須控制取像系統的某些跟聚焦相關的參數，並取得在不同參數值時的影像。在本論文中，我們採用了控制 D_i 即改變影像平面的位置的作法。Nayar 在 [27] 中提出了一個簡易的方法可將一般鏡頭轉成 Telecentric Lens，所以我們使用的鏡頭是 Telecentric Lens，它的好處是當改變影像平面位置時，只會影響到影像清晰度而不會影響到其它的幾何成像以及亮度值。假設影像平面位置在 D_i 時我們取得的影像記爲 I_i 。因此第 i 張和 $i + 1$ 張影像分別是在影像平面位於 D_i 與 D_{i+1} 時所取得，由於影像位置不同，因此它們的模糊半徑也不同。我們可以從 (14) 式推導出這兩張影像的模糊半徑的差：

$$\begin{aligned}
\delta R_{i,i+1} &= R_i - R_{i+1} \\
&= F \cdot \frac{|D_f - D_{i+1}| - |D_f - D_i|}{2 \cdot f_{num} \cdot D_f}
\end{aligned} \tag{15}$$

其中 $\delta R_{i,i+1}$ 代表兩張影像模糊半徑的差，如果 $\delta R_{i,i+1} > 0$ 表示第 $i+1$ 張影像比第 i 張影像清晰，反之如果 $\delta R_{i,i+1} < 0$ 表示第 i 張影像比第 $i+1$ 張影像清晰。我們歸納 (15) 式就可以得到：

$$D_f = \left| \frac{(D_i - D_{i+1}) \cdot F}{2 \cdot \delta R_{i,i+1} \cdot f_{num}} \right| \tag{16}$$

由 (16) 式知道由兩張不同模糊程度的散焦影像就可以用來估測最佳成像點 [5, 18]，再將最佳成像點代入 (2) 就能計算出我們想知道的深度資訊。而我們的散焦測距模組是依據兩張散焦影像的相對模糊關係找出最佳成像點，我們讓較清晰的影像對相對模糊關係函式作卷積運算，使較清晰的影像變成較模糊的影像。

每一個透鏡都有其特殊的 PSF，其正確的數值應由實驗來決定，本論文中我們以 3.2 節 (11) 提及的高斯函式 [25] 作為 PSF 來定立影像模糊的程度，下式為高斯函式在空間領域中的表示式：

$$G(x, y, s^2) = \frac{E^{-\frac{x^2+y^2}{2 \cdot s^2}}}{2 \cdot \pi \cdot s^2} \tag{17}$$

其中 s 可以看成模糊程度的參數，在本論文裡我們稱爲“絕對模糊參數”， s 和模糊半徑存在一個常數的正比例關係 (13)。解決了 PSF 的形式之後，我們就可以藉由 PSF 描述之前 I_i 及 I_{i+1} ：

$$\begin{aligned} I_i &= I_o * G(x, y, s_i^2) \\ I_{i+1} &= I_o * G(x, y, s_{i+1}^2) \end{aligned} \quad (18)$$

其中 I_o 代表物體於空間中的光強度， s_i 和 s_{i+1} 分別表示爲第 i 張和第 $i+1$ 張影像的絕對模糊參數值。假設 I_i 較 I_{i+1} 清晰，則它們之間存在一個相對模糊的關係：

$$I_i * G(x, y, \theta^2) - I_{i+1} = I_o * [G(x, y, s_i^2) * G(x, y, \theta^2) - G(x, y, s_{i+1}^2)] \quad (19)$$

其中 θ 爲兩張影像間的相對模糊參數值，經由 (19) 式及高斯函式的特性，我們可以將代表兩張影像的絕對模糊參數 s_i 、 s_{i+1} 和兩張影像間的相對模糊參數 θ 表示成：

$$s_i^2 + \theta^2 = s_{i+1}^2 \quad (20)$$

值得注意的，在我們拍攝影像的過程中或多或少都會有雜訊，所以研究的過程中我們假設加入的影像雜訊是呈高斯分佈的白雜訊，(18) 式可以表示成：

$$\begin{aligned} I_i &= I_o * G(x, y, s_i^2) + \varepsilon_i \\ I_{i+1} &= I_o * G(x, y, s_{i+1}^2) + \varepsilon_{i+1} \end{aligned} \quad (21)$$

其中 ε_i 、 ε_{i+1} 分別表示為第 i 張和第 $i+1$ 張影像的雜訊。了解到兩張影像間的關聯性後，接下來我們即可定義誤差函數：

$$e_{i,i+1}(D_f) = \begin{cases} \|(I_i * G(x, y, \theta^2) - I_{i+1}) \cdot W\|_F^2, & \delta R_{i,i+1} < 0, \\ \|(I_{i+1} * G(x, y, \theta^2) - I_i) \cdot W\|_F^2, & \delta R_{i,i+1} > 0, \end{cases} \quad (22)$$

在 (22) 式裡 θ 為相對模糊參數， W 為一視窗函式（我們採用 Hamming Window）。而引入此一視窗函式的主要目的是為了解決在執行卷積運算時於影像邊緣作 zero padding 所衍生的問題。最佳成像點 D_f 就可以透過對 (22) 式最佳化而求得，我們以下面介紹的流程作更詳細的說明：

1. 進行散焦測距計算的過程中，我們控制影像平面擺放位置 (D_i)，並利用

(15) 式計算 $\delta R_{i,i+1}$ 。

2. 依據 $\delta R_{i,i+1}$ 之正負號及其數值大小決定影像模糊的 PSF 即 $G(x, y, \theta^2)$ [於

後我們以 $\text{PSF}(\theta^2)$ 表示之]，此 $\text{PSF}(\theta^2)$ 的相對模糊參數 $\theta = k\sqrt{|R_i^2 - R_{i+1}^2|}$

(k 為一常數，參考 (13) 式)。

- 3.

$$\begin{cases} I_{b1} = I_{i+1} * G(x, y, \theta^2), \delta R_{i,i+1} > 0, \\ I_{b2} = I_i * G(x, y, \theta^2), \delta R_{i,i+1} < 0, \end{cases} \quad (23)$$

(23) 式若 $\delta R_{i,i+1} > 0$ 則將 I_{i+1} 以先前決定的 $\text{PSF}(\theta^2)$ 作卷積運算而得到一

張較模糊的影像 I_{b1} 。反之若 $\delta R_{i,i+1} < 0$ 則將 I_i 以先前決定的 $\text{PSF}(\theta^2)$ 作卷積運算而得到一張較模糊的影像 I_{b2} 。

4. 將作完卷積運算的模糊影像與另一張影像相減 (若 $\delta R_{i,i+1} > 0$ 則 $I_{b1} - I_i$ ；反之 $\delta R_{i,i+1} < 0$ 則 $I_{b2} - I_{i+1}$)，並計算誤差大小。如果誤差很小，則代表這個 D_f 數值是正確的，否則這個數值是有問題的。

由 (16) 式可知 $\delta R_{i,i+1}$ 相對於 D_f 是一遞減函數，因此當 D_f 增加時影像 I_i 與 I_{i+1} 之模糊程度差異逐次減少。若正確的最佳成像位置在 D_f^* ，而且 δR^* 為其相對應的相對模糊度。則當我們逐漸將 D_f 由 D_{fmin} 往 D_{fmax} 方向增加時，其對應的 δR 在逼近 δR^* 時，(22) 式的誤差值就會變小； D_f 遠離 D_f^* 時誤差就會變大。所以 (22) 式大致上符合一個 unimodal function 的要求 (圖 9)。

4.2 搜尋最佳成像點的方法

在上一節中我們定義了 D_f 的誤差函式，估測 D_f 的方法則是搜尋 $e(D_f)$ 的最小值。在不失一般性的情形之下，我們假設有兩張不同模糊程度的散焦影像 I_1 、 I_2 它們的模糊半徑分別為 R_1 、 R_2 ，在搜尋最佳成像點的過程中，我們將影像分割成許多小區塊來搜尋，並假設這些區塊都平行於我們影像平面，如果 D_f 比較靠近 I_2 (如圖 7) 則 $R_1 > R_2$ ，所以將 I_2 對相對模糊參數為 θ 的 $\text{P}(\theta^2)$ 作卷積運算然後再將經過模糊後的影像和 I_1 相減即為 $e(D_f)$ ；若搜

尋的過程中假使 D_f 比較靠近 I_1 (如圖 8) 則 $R_2 > R_1$ ，所以將 I_1 對 $P(\theta^2)$ 作卷積運算然後再將經過模糊後的影像和 I_2 相減即為 $e(D_f)$ 。因為 (22) 式在沒有雜訊的情況下為 unimodal function (圖 9)，所以我們採用了 Fibonacci search 方法 [17] 較有效率地搜尋最佳的 D_f 值。費式搜尋法類似於二元搜尋法，差別在於費式搜尋不需經過事前的排序，以及劃分搜尋的範圍的方式是以費式數列 (Fibonacci numbers) 來劃分：

$$F_0 = 1, F_1 = 1 \quad (24)$$

$$F_n = F_{n-1} + F_{n-2}, n \geq 2$$

上式為費式數列的產生方式，當 $n \geq 2$ 時，其費式數字為前面兩項費式數字之和。如果搜尋的區間為 d_0 則當搜尋的次數為 i 次的時候，最佳解將被我們收斂在 d_i 之間：

$$d_i = \left(\frac{F_{n-i+1}}{F_n} \right) d_0 \quad (25)$$

由上式可以知道當我們以費式搜尋 n 次時，則最佳解可以收斂在原本搜尋區間的 $\frac{1}{F_n}$ 範圍之內。當我們以費氏搜尋法 (Fibonacci search) 大略找出最佳成像點，而後在最佳成像點附近，利用一個二次多項式來近似，再利用二次多項式的內插，求出最佳成像點的位置。

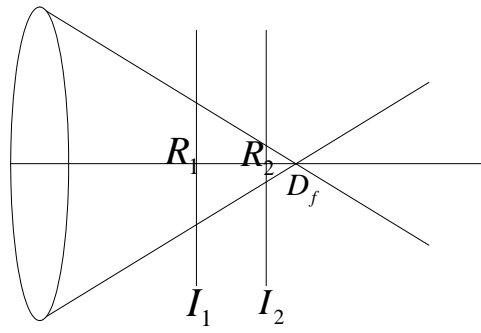
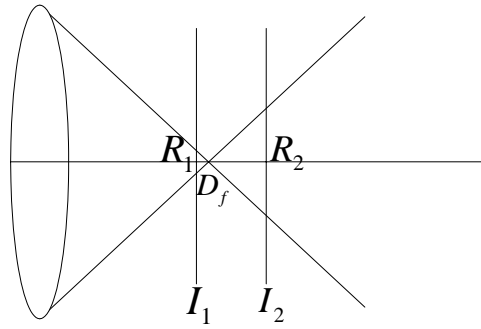
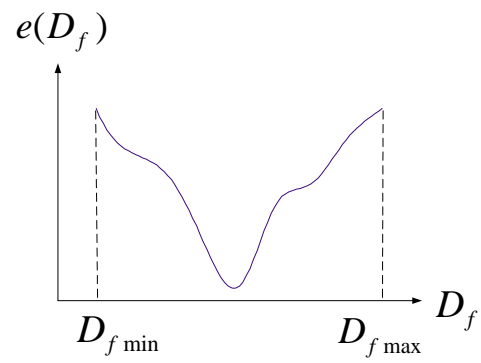
圖 7: 最佳成像點較靠近 I_2 的情形圖 8: 最佳成像點較靠近 I_1 的情形

圖 9: 誤差函式

5 決定最佳的相機參數

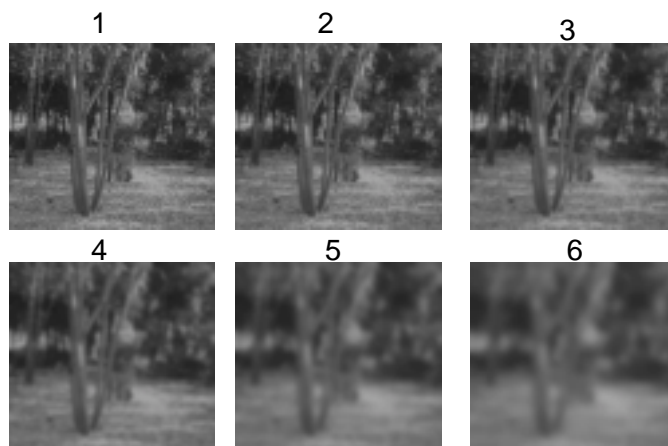


圖 10: 散焦影像序列

經由第 4 章的介紹我們知道只要有兩張不同模糊程度的散焦影像，就可以定出最佳成像點 D_f 的誤差函式，當找到最精準的 D_f 後就可以求得空間中深度的資訊。圖 10 中有 6 張不同模糊程度的散焦影像，6 張影像中任取兩張可以有 15 種可能的組合方式 (若有 n 張影像則有 C_2^n 種組合) 供散焦模組作運算，在研究過程中我們發現選擇不同組合的影像會影響到估算出來深度的精確值，但是我們應該如何選取兩張散焦影像？這將是我們這個章節討論的重點。

5.1 散焦測距估測誤差之 Cramér-Rao 下限

對任意一個無偏差 (unbiased) 的估測器 $\hat{\theta}$ 而言，其變異量之 CRLB 可由下式決定：

$$\text{Var} [\hat{\theta}] \geq J^{-1}, \quad (26)$$

其中 J 為 Fisher information matrix[21]:

$$J = E \left\{ \left[\frac{\partial}{\partial \theta} \ln P(\epsilon) \right] \left[\frac{\partial}{\partial \theta} \ln P(\epsilon) \right]' \right\} \quad (27)$$

其中 ' 的符號代表轉置運算 (transpose)，在不失一般性的情況下，我們假設 I_1 、 I_2 分別為兩張不同模糊程度的散焦影像，並且 I_1 比 I_2 模糊，若影像雜訊為呈 Gaussian 分佈的白雜訊則 $I_1 = I_2 * G(x, y, \theta^2) + \epsilon(x, y)$ ， $\epsilon(x, y)$ 就是我們量測影像的雜訊值，為了方便推導，我們只針對量測的影像加入雜訊值。

(27) 式中 $P(\epsilon)$ (本論文中量測影像雜訊的機率分佈函式) 就可以表示成：

$$P(\epsilon) = \frac{1}{2\pi\sigma^2} \cdot \exp \frac{-\sum_{x,y} \epsilon(x,y)^2}{2\sigma^2} \quad (28)$$

(28) 式中， $\epsilon(x, y) = I_1 - I_2 * G(x, y, \theta^2)$ ，我們對量測影像雜訊的機率分佈函式取自然對數後再對 θ 偏微分，就可以得到：

$$\frac{\partial}{\partial \theta} \ln P(\epsilon) = \frac{1}{\sigma^2} \sum_{x,y} \epsilon(x, y) \cdot (I_2 * \frac{\partial G(x, y, \theta^2)}{\partial \theta}) \quad (29)$$

將 (29) 式代入 (27) 式：

$$\begin{aligned}
 J &= E \left[\frac{1}{\sigma^4} \left(\sum_{x,y} \epsilon(x,y) \cdot I_2 * \frac{\partial G(x,y,\theta^2)}{\partial \theta} \right) \left(\sum_{x_1,y_1} \epsilon(x_1,y_1) \cdot I_2 * \frac{\partial G(x_1,y_1,\theta^2)}{\partial \theta} \right) \right] \\
 &= \frac{1}{\sigma^4} \sum_{x,y} \sum_{x_1,y_1} \left(I_2 * \frac{\partial G(x,y,\theta^2)}{\partial \theta} \right) (x,y) \left(I_2 * \frac{\partial G(x_1,y_1,\theta^2)}{\partial \theta} \right) (x_1,y_1) E[\epsilon(x,y) \cdot \epsilon(x_1,y_1)]
 \end{aligned} \tag{30}$$

因爲假設影像雜訊爲白雜訊，所以 (30) 式中 $(x,y) \neq (x_1,y_1)$ 則 $E[\epsilon(x,y) \cdot \epsilon(x_1,y_1)] =$

0，若 $(x,y) = (x_1,y_1)$ 則 $E[\epsilon(x,y) \cdot \epsilon(x_1,y_1)] = \sigma^2$ ，因此 (30) 式可以表示成：

$$J = \frac{1}{\sigma^2} \sum_{x,y} \left(I_2 * \frac{\partial G(x,y,\theta^2)}{\partial \theta} \right)^2 (x,y) \tag{31}$$

因爲我們是對影像所包含的範圍計算，若當計算的位置超出影像的範圍則

沒有任何影像資訊 ($J = 0$)，所以我們可以將 (31) 式中對 (x,y) 的累加改成用

(x,y) 方向的積分來近似：

$$J \cong \frac{1}{\sigma^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(I_2 * \frac{\partial G(x,y,\theta^2)}{\partial \theta} \right)^2 dx dy \tag{32}$$

經由 (19) 式高斯函式的特性，我們可以將 (32) 的 I_2 換成以 I_o 表示：

$$J \cong \frac{1}{\sigma^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(I_o * \frac{\partial G(x,y,s_2^2 + \theta^2)}{\partial \theta} \right)^2 dx dy \tag{33}$$

爲了運算上的方便我們將 (33) 式經由 Parseval's theorem 從空間領域轉置成頻

率領域，所以 (33) 式中卷積運算就可以置換成較容易的乘法計算：

$$J \cong \frac{1}{2\pi\sigma^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\hat{I}_o|^2 \cdot \left| \frac{\partial \hat{G}(x,y,s_2^2 + \theta^2)}{\partial \theta} \right|^2 du dv \tag{34}$$

其中 \hat{I}_o 與 $\frac{\partial \hat{G}(u, v, s_2^2 + \theta^2)}{\partial \theta}$ 分別為 I_o 與 $\frac{\partial G(x, y, s_2^2 + \theta^2)}{\partial \theta}$ 在頻率領域中的表示式。又因為 (34) 式已經轉換到和 (32) 式不同的領域，所以 (32) 式中對 x, y 方向的積分在 (34) 式中置換成對頻率 u, v 方向的積分。我們可以令在 (34) 式中可以找到最大的 $|\hat{I}_o|^2$ 值為 M 將之代入 (34) 式即可得下列不等式：

$$J \leq \frac{M}{2\pi\sigma^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left| \frac{\partial \hat{G}(u, v, s_2^2 + \theta^2)}{\partial \theta} \right|^2 dudv \quad (35)$$

(35) 式中 $\frac{M}{2\pi\sigma^2}$ 為一常數可以用 K 來取代：

$$J \leq K \cdot \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[\frac{\partial G(u, v, s_2^2 + \theta^2)}{\partial \theta} \right]^2 dudv \quad (36)$$

(36) 式經過積分後我們就可以得到下式：

$$J \leq K \cdot \frac{\theta^2}{(s_2^2 + \theta^2)^3} \quad (37)$$

由 (37) 式中我們可以知道當 $\theta = \frac{s_2}{\sqrt{2}}$ 的時候 J 會有最大值。圖 11 為假設絕對模糊參數 $s_2 = 0.1$ 實驗中 J 的表示圖，水平方向為相對模糊程度的大小，垂直方向為對應的 J 值。而 CRLB 是 J^{-1} (圖 12)。所以最佳相機參數值應該要設定為可以使得相對模糊程度為 $\frac{s_i}{\sqrt{2}}$ (s_i 為第 i 張影像的絕對模糊參數)。

在 (37) 式中我們發現要是絕對模糊程度 s 愈接近 0 則 J 愈大。而最佳相機參數值 $\theta = \frac{s}{\sqrt{2}}$ ，所以要是 s 太小的話，(19) 式中的高斯函式因為 Aliasing 的關係造成影像失真，此時 (19) 式 discrete convolution 近似將不適用，由此可知

我們的 s 不能小到使得 $\text{PSF}(\theta^2)$ 的 spectrum 超出 Nyquist 頻率。而我們取樣的大小為影像點的點距，我們以 dp 表示，所以 $\text{PSF}(\theta^2)$ 的 Nyquist 頻率可以表示為 $\frac{\pi}{dp}$ ， θ 的下界就被定義為：

$$\theta \geq \frac{c \cdot dp}{\pi} \quad (38)$$

其中 c 為一常數可由實驗求得。

在推導量測影像雜訊的 CRLB 時我們並未計入 I_2 之雜訊，而在實際取像時這兩張影像皆會有雜訊的存在，由 (38) 式我們知道 θ 會有個下界，當 θ 大過下界的值時，即使 I_2 加入雜訊值 ϵ_2 ，如 (39) 式，雜訊值也會因為經過模糊參數較大的卷積運算而被抑制。

$$(I_2 + \epsilon_2) * G(x, y, \theta^2) = I_2 * G(x, y, \theta^2) + \epsilon_2 * G(x, y, \theta^2) \quad (39)$$

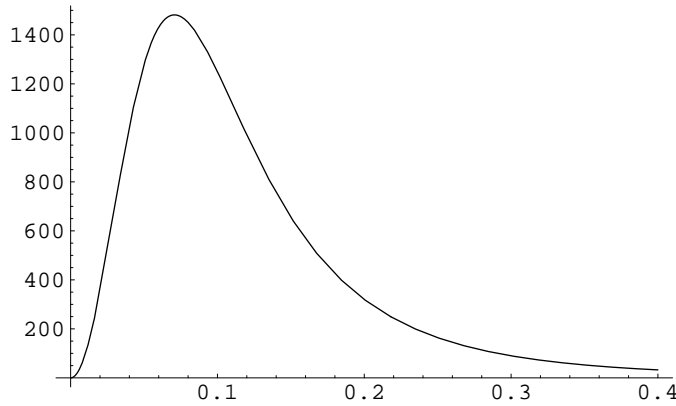


圖 11: 估測相對模糊程度的 Fisher information matrix 值

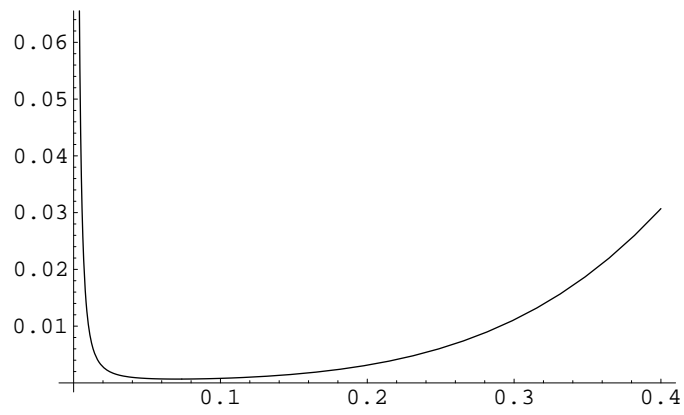


圖 12: 估測相對模糊程度的 CRLB

6 最佳散焦測距方法

由 5.1 節 (37) 式我們可以得到：

$$J \propto K \frac{\theta^2}{s^2 + \theta^2} \quad (40)$$

由上式可以知道當 s 愈小 J 愈大，也就是說如果我們決定使用的散焦影像能夠愈清晰的話，則估算出來的深度值會更精確。在現實生活中我們欲重建的三維深度資訊大多都有著各種不同的深度變化，很少會是一個單一深度值的平行平面，所以藉由 (40) 式我們必須有多張在不同深度位置清晰影像的資訊才能估算更精準的深度量測值。而整合多張影像的誤差函式就可以重新定成：

$$e_{n-1} = \sum_{i=1}^{n-1} \|(I_i * G(x, y, s^2) - I_{i+1}) \cdot W\|_F^2 \quad (41)$$

整合多張散焦影像來估算的深度值雖然會更加的精準，但是由整合多張影像所帶來的龐大計算量以及系統記憶體의 負載問題也是不能夠忽略的。所以我們採用 asymptotic Bayesian 遞迴式估測法來解決這個問題。

6.1 Asymptotic Bayesian 遞迴式估測法

爲了能求得更精準的 D_o 我們使用 asymptotic Bayesian 遞迴式估測法以整合多張影像所提供的資訊。爲方便說明 asymptotic Bayesian 方法，我們以圖 13 作

為說明例。在處理第一對影像時我們並無任何資訊，因此我們只能信賴由這兩張影像所提供的解，在最小化 $e_{1,2}$ 之後即可決定最佳 D_{f1} ，然後在這個最佳解附近將 $e_{1,2}$ 化簡為一個二次式：

$$e_{1,2}(D_f) \approx \hat{e}_{1,2}(D_f) = \Phi_1 \cdot (D_f - D_{f1})^2 + c, \quad (42)$$

圖 13(a) 中可見 $e_{1,2}$ 與 $\hat{e}_{1,2}$ 在 D_{f1} 附近非常近似。而後在第三張影像進來時，我們可針對 $e_{1,2} + e_{2,3}$ （如圖 13(b)）同時求解，但是當影像數量繼續增加時這樣的運算量將大幅上昇，因此在這個步驟中我們是針對 $\hat{e}_{1,2} + e_{2,3}$ 來求解（如圖 13(c)）。圖 13(d) 中，我們將 $e_{1,2} + e_{2,3}$ 與 $\hat{e}_{1,2} + e_{2,3}$ 的函數圖形放在一起供比較用。由這個圖我們可發現在最佳解的附近這兩種方式都能提供非常好的解。如此我們即可將以前的所有影像所提供的資訊都簡化為一個二次多項式，而這個多項式的二階係數 Φ_i 則是對先前量測值的信賴度。 Φ_i 值愈高代表我們對 D_f^* （正確的最佳成像位置）的估測愈可相信，而事實上在 Φ_i 的數值夠高時這個二次式將會主導 $\hat{e}_{1,2} + e_{2,3}$ 的最小化。在完成 $\hat{e}_{1,2} + e_{2,3}$ 的最小化之後，我們即可再對 $e_{2,3}$ 求取近似的二次式，然後再將這個二次式的係數值加入原本的係數值中，使得 Φ_i 在遞回求解的過程中不斷增加。我們將 asymptotic Bayesian 遞迴式估測法整合多張影像的數式表示如下：

$$\begin{aligned} \hat{e}(D_f) &= \sum_{i=1}^{n-1} \hat{e}_{i,i+1}(D_f) + e_{n-1,n} \\ \Phi_{n+1} &= \Phi_n + \frac{\partial^2 e_{n-1,n}}{2\partial D_f^2}(D_f) \end{aligned} \quad (43)$$

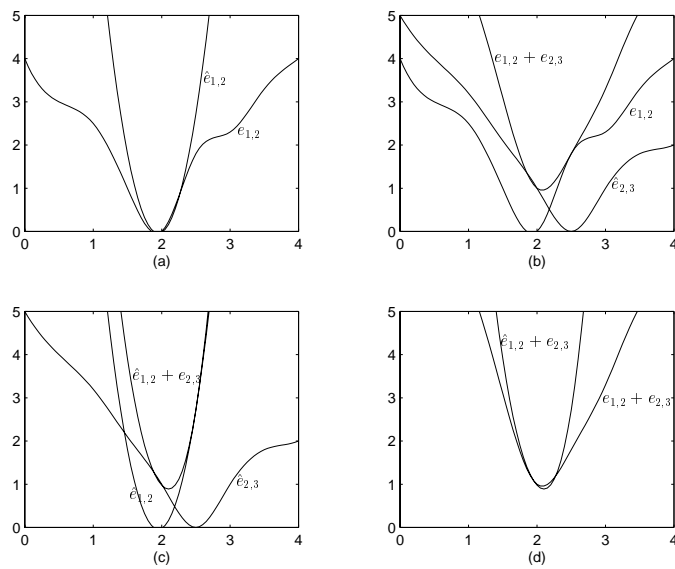


圖 13: Asymptotic Bayesian 遞迴式估測法示意圖

7 實驗結果與討論

以下將討論我們的實驗過程及結果：於電腦模擬實驗中首先模擬一臺相機鏡頭 (Cosmicar B2514D-TH), 它的焦距為 25mm, f-number 為 1.4 至 22 之間。影像合成的方法是把 D_i 和自訂的深度 D_0 代入 (14) 式, 再以得到的模糊半徑 R 代入 (17) 式以得到光源散佈函式, 並和貼圖影像做卷積運算就可以合成出實驗需要的影像。

首先必須要驗證我們散焦測距模組的最佳相機參數值, 實驗的過程中我們假設影像的深度為 300mm 的平行平面, 貼圖影像為 64×64 點, 顏色為灰階色彩。第 1 張影像的影像平面擺放位置為距離透鏡中心 $27.4227mm$, 每遞增 $0.0034mm$ 取像一張, 共計 20 張的測試影像序列, 其中第 11 張影像相對於第一張影像為最接近最佳化相機參數值, 其影像平面擺放位置距離透鏡中心 $27.4564mm$ 。於這 20 張影像中各自加入標準差為 4 個灰階值的雜訊, 並以第 1 張影像分別對其餘 19 張影像進行我們散焦測距的深度估算, 並觀察各對影像的變異量變化。實驗的結果如圖 14 所示, 變異量的變化並不如在第 5 章所推導的 CRLB (參考 (37) 式及圖 12), 這也就是我們之前 5.1 節提及當相對模糊程度太小, 因為 Aliasing 造成影像失真, 以致我們推導出散焦測距相對模糊程度間的 CRLB 不適用。如圖 15, 水平和垂直方向代表相對模糊程度 θ , 而 L_1 為我們量測的相對模糊曲線, L_2 為實際的相對模糊曲線, 此時的 L_1 都

呈現高估的情形。接著我們以 (38) 式，對 θ 的下界作限定，依據實驗的數值， θ 的下界被定為 0.0095 ，藉著被設定的 θ 下界及之前的進行的實驗方式，我們將影像平面擺放在距離透鏡中心 $27.8727mm$ 處攝取第一張影像，每遞增 $0.0135mm$ 取像一張，第 11 張影像平面擺放位置距離透鏡中心 $28.0076mm$ ，最接近最佳相機參數值。此時因為 θ 變大，Aliasing 的情形不會發生，因此符合了我們所推導的散焦測距相對模糊程度間的 CRLB。圖 16 為我們估測的相對模糊曲線 L_1 相較於實際的相對模糊曲線 L_2 的表示圖，而變異量的變化 (圖 17) 大致上也如我們所推導的結果。當相對模糊程度過大時，光源散佈函式的含概範圍也跟著變大，超出了我們實驗影像 64×64 點的範圍，導致圖 16 及圖 17 中相對模糊程度大於 0.045 以後估測不準和變異量變化不如預期的情形，由此我們也發現 s 受到影像大小的影響，也會有上界的限定。

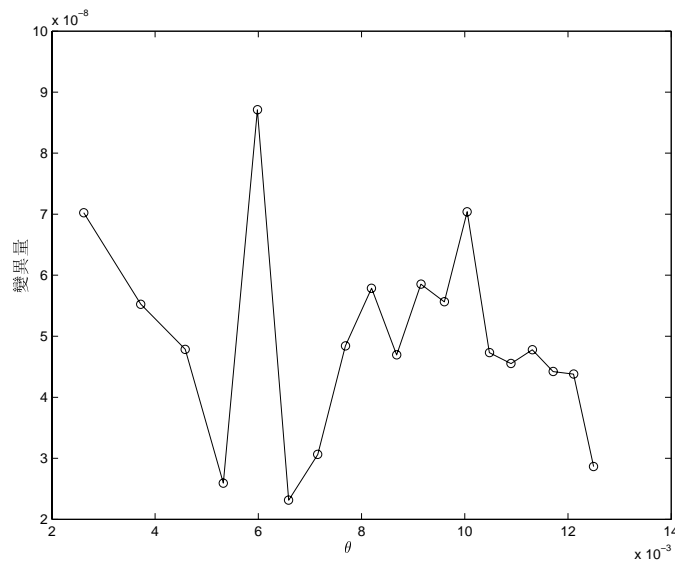


圖 14: 相對模糊程度低於下界限定值時之變異量變化圖

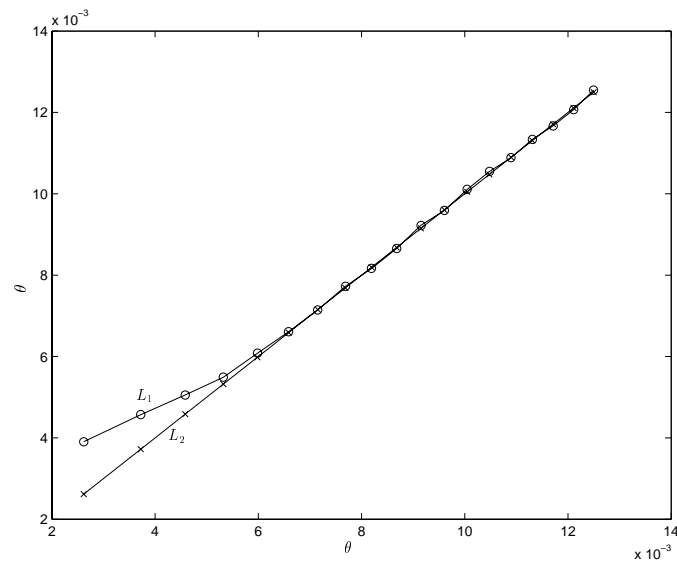


圖 15: 相對模糊程度低於下界限定值時，估測的相對模糊與實際的相對模糊之比較。

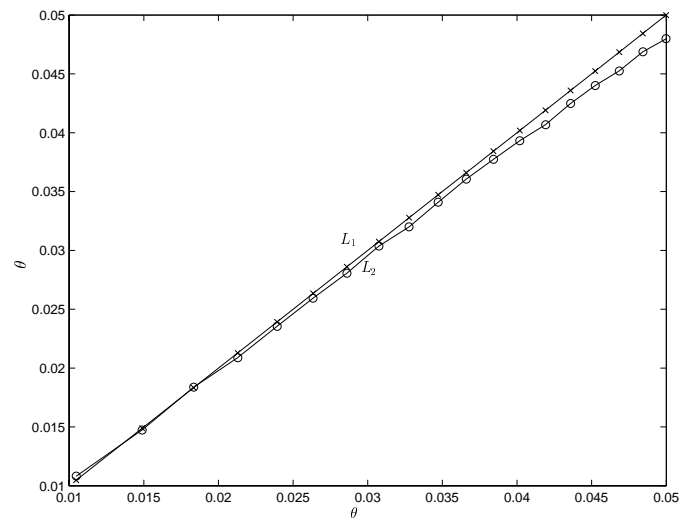


圖 16: 相對模糊程度高於下界限定值時，估測的相對模糊與實際的相對模糊之比較。

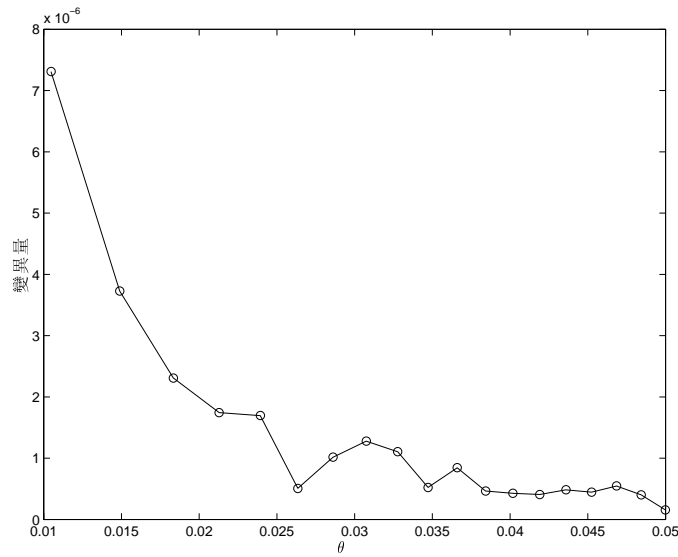


圖 17: 相對模糊程度高於下界限定值時之變異量變化圖

依據最佳相機參數值的設定，在散焦測距系統的電腦模擬實驗中我們測試了三種不同的深度變化，一種是整張影像深度完全相同的平行平面，第二種是假設對一張傾斜的平面取像，第三種是假設對有較複雜變化的傾斜平面取像。實驗所用之影像大小為 256×256 點，顏色為灰階色彩，在影像中我們加入了標準差為 4 個灰階值的雜訊，並且於實驗中不考慮量子化的誤差 (qantization error) 所產生的問題。

7.1 測試平行平面資訊量測

在模擬影像的過程裡我們影像合成的方式是先自訂影像深度，並假設我們的物體深度都介於 300mm 到 1000mm 之間，所以對應的 D_i 範圍為 27.3018mm 到 25.7961mm，依據最佳化像機參數的設定值於每兩張模糊程度相差 0.0095 便

合成一張影像。並將影像分成 16×16 塊部份重疊的區塊，每塊大小為 32×32 個像素。接著假設目標物的位置在 100mm 至無窮遠處的範圍，因此我們在其對應的最佳成像點 (25mm 至 33.3333mm) 之範圍中搜尋 D_f ，再配合 asymptotic Bayesian 遞迴式估測法得到精確的 D_f ，並用這個數值重建出實際的深度。

在測試平行平面的資訊量測的模擬實驗中，我們將物體深度設成 300mm (圖 18)，我們有兩種實驗版本，一種是沒有加入雜訊的，另一種是有加入雜訊的版本。在這個實驗中，我們使用如圖 19 的影像，貼在圖 18 的平面上以合成實驗所需的影像。圖 20 左上圖為沒有加入雜訊所估算出來的三維網狀圖，

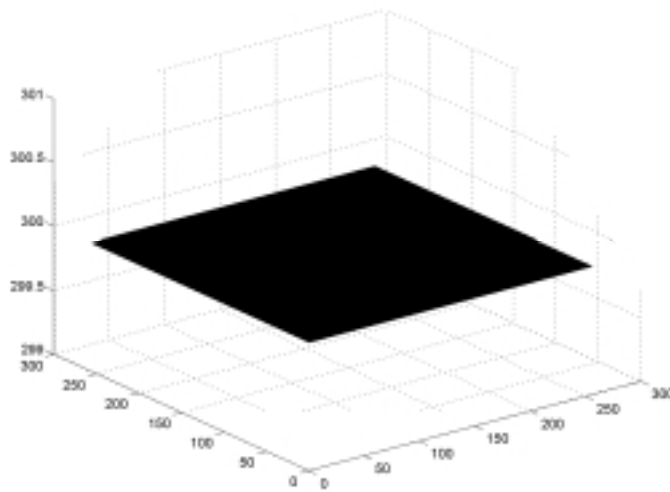


圖 18: 平行平面深度相關示意圖

圖 20 右上圖為從側面看的深度值表示圖。由實驗的結果發現在沒加入雜訊的平面影像裡，比較估算出來的深度和實際深度其平均誤差小於 0.044mm 。

圖 20 左下圖為加入誤差的平面的三維估算網狀圖，圖 20 右下圖為從側面看的



圖 19: 測試平行平面貼圖影像

深度值表示圖。在加入雜訊的實驗中裡，深度估測結果依然可以將誤差控制在 $0.075mm$ 之內。

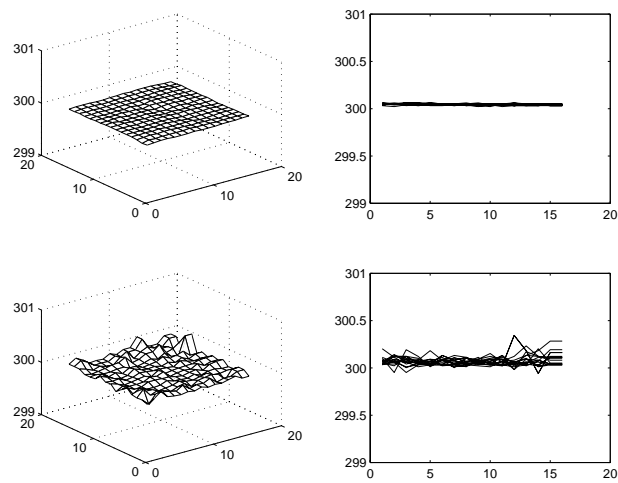


圖 20: 恢復平行平面深度影像

7.2 測試傾斜平面資訊量測

在測試傾斜平面的資訊量測的模擬實驗中，我們模擬有一個平面斜放在攝影機前，其深度範圍在 $800mm$ 到 $950mm$ 之間 (圖 21)，我們有四種實驗版本，一種是只使用兩張散焦影像量測深度資訊，其中又將之分成未加入雜訊及加入雜訊的情形；另一種是以 asymptotic Bayesian 整合多張影像並再分成沒有加入雜訊，及有加入雜訊的版本。在這個實驗中，我們使用如圖 22 的影像，貼在圖 21 的平面上以合成實驗所需的影像 (如圖 23)。圖 24 左上圖為只有使

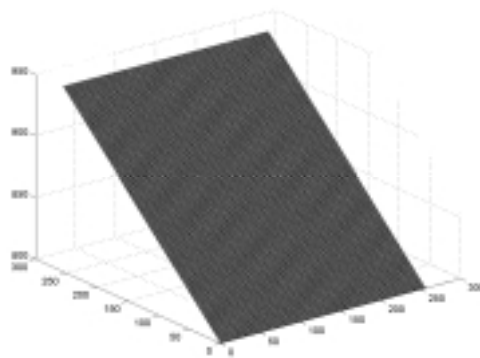


圖 21: 傾斜平面深度相關示意圖

用兩張影像，沒有加入雜訊估算出來的三維網狀圖，圖 24 右上圖為從側面看的深度值表示圖，因為我們將影像切割成許多平行於影像平面的區塊進行計算，所以會導致估算出來的結果會有誤差。由實驗的結果發現在沒加入雜訊的傾斜平面影像裡，比較估算出來的深度和實際深度其誤差為 $4.244mm$ 。圖 24 下排則為加上雜訊之後的估測結果平均誤差為 $6.486mm$ 。



圖 22: 測試傾斜平面貼圖影像



圖 23: 測試傾斜平面散焦影像

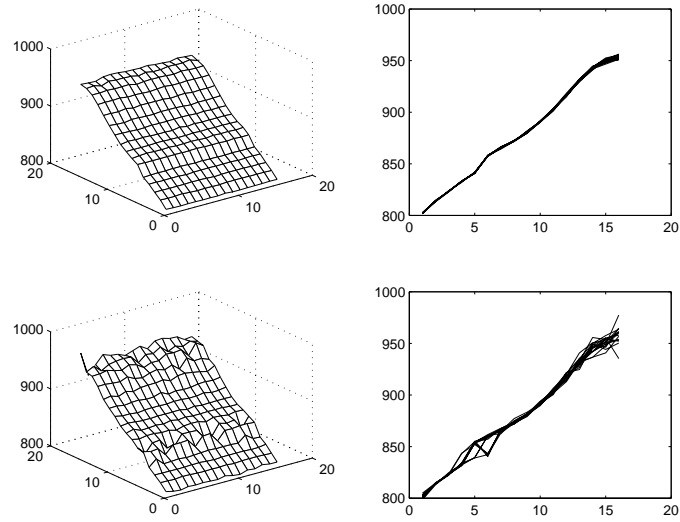


圖 24: 兩張影像：恢復傾斜平面深度影像

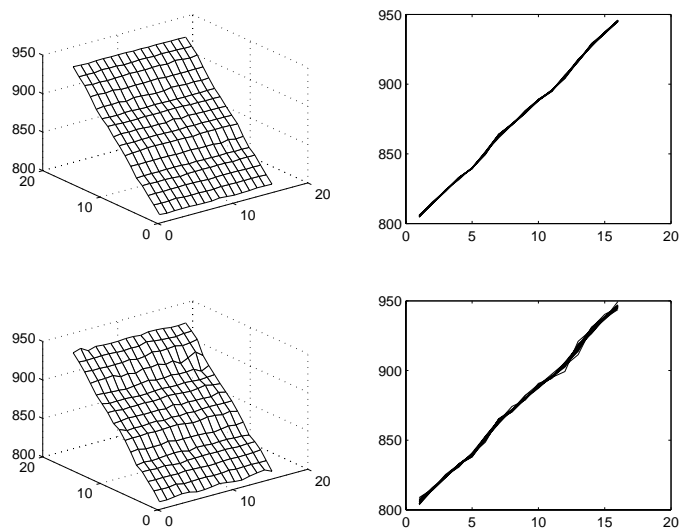


圖 25: 多張影像：恢復傾斜平面深度影像

圖 25 左上圖為使用 asymptotic Bayesian 遞迴式估測法整合多張影像，沒有加入雜訊估算出來的三維網狀圖，圖 25 右上圖為從側面看的深度值表示圖。由實驗的結果發現在沒加入雜訊的傾斜平面影像裡，比較估算出來的深度和實際深度其誤差為 $1.368mm$ 。圖 25 下排則為加上雜訊之後的估測結果平均誤差為 $1.693mm$ 。由圖 25 和圖 24，以及我們實驗的數據發現到整合多張影像估算而得的深度量測值確實比只用兩張影像計算出的深度量測值來的好。

7.3 測試複雜傾斜平面資訊量測

爲了知道我們的散焦測距模組在對較複雜的深度變化是否也有一定的精確性，所以在這個實驗中，我們模擬在攝影機前有個複雜的傾斜平面，其左半面的深度範圍在 475mm 到 500mm 之間，右半面的範圍在 525mm 到 500mm 之間，呈一“X”形狀的深度變化 (圖 26)。我們實驗的方式和測試傾斜平面資訊量測 (7.2) 雷同，我們使用如圖 22 的影像，貼在圖 26 的平面上以合成實驗所需的影像 (如圖 27)。

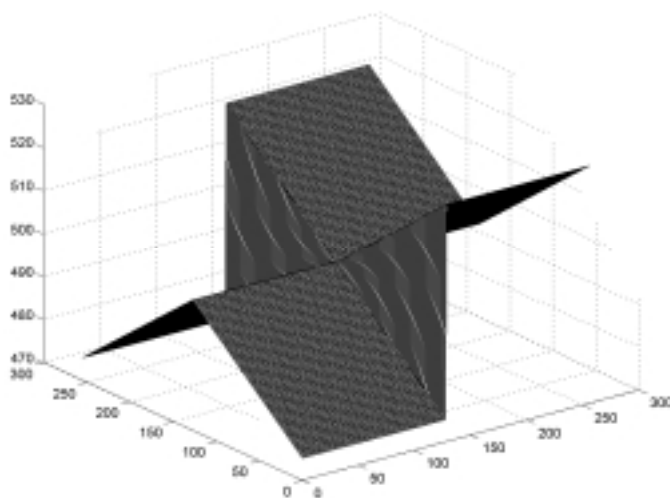


圖 26: 複雜傾斜平面深度相關示意圖

圖 28 左上圖爲只有使用兩張影像，沒有加入雜訊估算出來的三維網狀圖，圖 28 右上圖爲從側面看的深度值表示圖。由實驗的結果發現在沒加入雜訊的傾斜平面影像裡，比較估算出來的深度和實際深度其誤差爲 1.434 mm 。

圖 28 下排則爲加上雜訊之後的估測結果平均誤差爲 2.115 mm 。圖 29 左上圖



圖 27: 測試複雜傾斜平面散焦影像

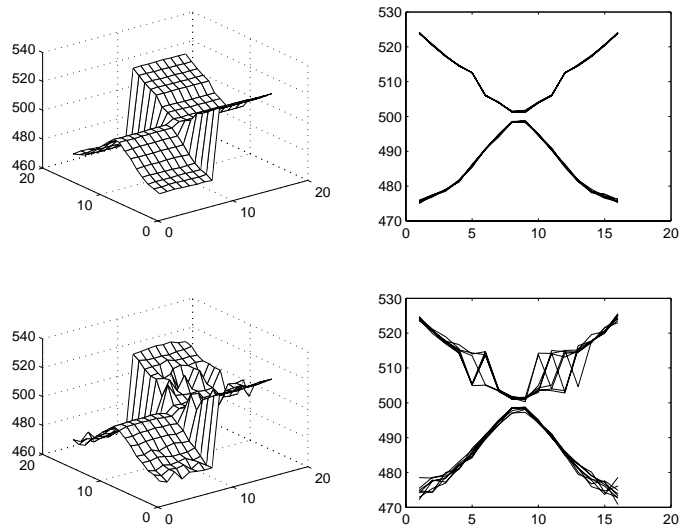


圖 28: 兩張影像：恢復複雜傾斜平面深度影像

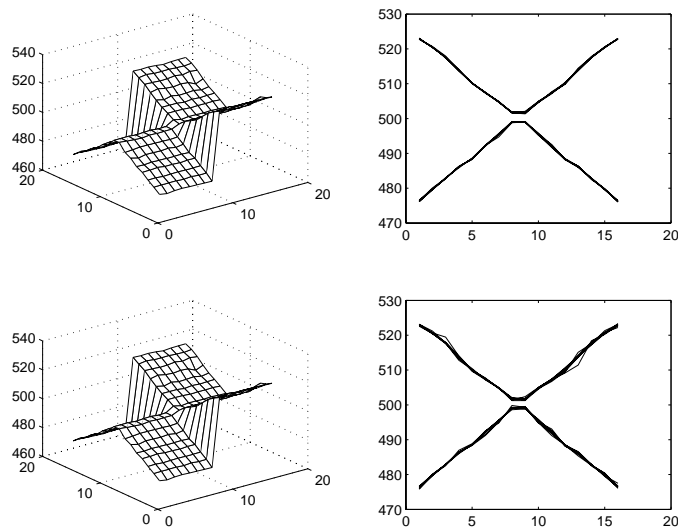


圖 29: 多張影像：恢復複雜傾斜平面深度影像

為使用 asymptotic Bayesian 遞迴式估測法整合多張影像，沒有加入雜訊估算出來的三維網狀圖，圖 29 右上圖為從側面看的深度值表示圖。由實驗的結果發現在沒加入雜訊的傾斜平面影像裡，比較估算出來的深度和實際深度其誤差為 $0.324mm$ 。圖 29 下排則為加上雜訊之後的估測結果平均誤差為 $0.429mm$ 。比較圖 29 和圖 28 及我們實驗的數據發現到在較複雜的深度變化情形，整合多張影像估算而得的深度量測值比只用兩張影像計算出的深度量測值來的好，並且實驗的結果是非常精確的。

8 結論與未來展望

本論文中我們經由對散焦測距模組的誤差分析，並以電腦模擬驗證最佳化散焦測距系統的準確度，也確實得到了非常好的測距結果。在這個研究中我們達成了以下幾點具體成果：

1. 我們以散焦影像間的相對模糊關係找尋最佳成像點的過程中，以控制 D_i 來決定模糊半徑大小的方法，判定那張影像較為清晰，接著以較清晰的影像對相對模糊參數為 θ 的 $\text{PSF}(\theta^2)$ 作卷積運算，使較清晰的影像能變成較模糊的影像。依據我們的方法可避免由較模糊的影像去找尋較清晰影像，需經過高通濾 (high pass filter) 處理的狀況。
2. 我們以費氏搜尋法 (Fibonacci search) 大略找出最佳成像點，而後在最佳成像點附近，利用一個二次多項式來近似，再利用二次多項式的內插，求出最佳成像點的位置。這樣的方法不但減少了電腦的運算時間 (比起以往有些散焦測距模組的相關研究者，以線形搜尋法求算最佳解，如 [5] 來得更為有效率。)，而且比只用費式搜尋的方法更提高了估算最佳解的精準度，更進一步的避免掉使用 inverse filtering 演算法，當分母為零時造成估算誤差的情形。
3. 我們藉由 CRLB 的方法推導了散焦測距系統的最佳相機參數值，接著依

據設定最佳相機參數值，我們發現當 s 愈小 J 愈大 ((40) 式)，也就是說如果我們決定使用的散焦影像能夠愈清晰的話，則估算出來的深度值會更精確。在現實生活中我們欲重建的三維深度資訊大多都有著各種不同的深度變化，很少會是一個單一深度值的平行平面，所以我們必須有多張在不同深度位置清晰影像的資訊才能估算更精準的深度量測值。然而當 s 過小時，則影響影像模糊的光源散佈函式會因為 Aliasing 的關係而導致影像失真，所以我們也依據取樣定理的限制，訂立了最佳相機參數值的下界。

4. 在整合多張影像的過程中，爲了避免資料量太過龐大，造成電腦系統運算及記憶體負擔，我們採用 asymptotic Bayesian 遞迴式估測法，以整合多張影像的資訊，這不但可以大幅降低計算所需的記憶體空間，而且也可以在實際系統中一邊取像同時進行運算，節省了時間。
5. 最後實驗的結果中，我們以電腦模擬來驗證散焦測距模組之精確度，發現即使在較複雜的深度變化情形之下，我們的最佳化散焦系統依然估算出很精準的深度量測值。

然而電腦視覺的研究最終目的，就是要能夠將我們的研究技術應用在日常生活中。由於我們電腦模擬的結果是如此的精準，更讓我們想要加快完成最佳化散焦測距系統實際的實驗。

參考書目

- [1] Amos, D.E., “A Subroutine Package for Bessel Function of a Complex Argument and Nonnegative Order,” Sandia National Laboratory Report, SAND 85-1018, 1985.
- [2] Bolles, R.C., Baker, H.H., Hannah, M.J., The JISCT Stereo Evaluation. Proc.DARPA ImageUnderstanding Workshop. pp.263-274,1993.
- [3] Born, M. and Wolf, E., “Principles of Optics”, 1980.
- [4] Darrell, T., and Wohn, K., “Pyramid Based Depth from Focus,” the IEEE CVPR,1988.
- [5] Ens, J., Lawrence, P., “An Investigation of Methods for Determining Depth from Focus,” IEEE Transactions on pattern analysis and machine intelligence, Vol 15, no.2, 1993.
- [6] Grossmann. P., “Depth from focus”, Pattern Recognition Letters, Vol. 5, pp.63-69 ,January 1987.
- [7] Hecht, E., Zajac, A., “Optics,” 1974.
- [8] Horn, B.K.P., Focusing, AI Lab, Memo 160. MIT, Combridge, Mass, 1968.

- [9] Hung, Y.-P., Cooper, D.B., and Cernuschi-Frias, B., “Asymptotic Bayesian Surface Estimation Using an Image Sequence,” *International Journal of Computer Vision*, Vol. 6, No. 2, pp. 105–132, 1991.
- [10] Jain, A. K., “Advances in mathematical models for image processing,” *Proc. IEEE*, 69(5):502-528, 1981.
- [11] Kenneth R. Spring and Michael W. Davidson., “Modulation Transfer Function”
<http://www.microscopyu.com/articles/optics/mtfintro.html>
- [12] Klarquist, W.N., Geisler, W.S., and Bovik, A.C., “Maximum-Likelihood Depth-from-Defocus for Active Vision,” *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, pp. 374–379, 1995.
- [13] Krotkov, E., “Focusing,” *International Journal of Computer Vision*, Vol. 1, pp. 223–237, 1987.
- [14] Nayar, S.K., Masahiro Watanabe, and Minori Noguchi. “Real-Time Focus Range Sensor,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 12 pp. 1186–1198, 1996.
- [15] Oliensis, J., Genc, Y., “Fast and accurate algorithms for projective multi-image structure from motion,” *Pattern Analysis and Machine Intelligence IEEE Trans-*

actions on Vol.23 pp.546–559, 2001

- [16] Oppenheim. Alan V., Willsky. Alan. S., Signals and Systems, second edition.
- [17] Ottmann, T., Widmayer, P. “Algorithmen and Datenstrukturen”, Mannheim, Leipzig, Wien, Zurich, BI-Wiss.-Verl., 1993.
- [18] Pentland, A.P., “A New Sense for Depth of Field,” IEEE Trans.PAMI 9, pp.523–531, 1987.
- [19] Rajagopalan, A.N., Chaudhuri, S., “A block shift-variant blur model for recovering depth from defocused images,” Image Processing, Proceedings., International Conference on Vol:3, pp.636-639,1995.
- [20] Rajagopalan, A.N., Chaudhuri, S., “Optimal selection of camera parameters for recovery of depth from defocused images,” Computer Vision and Pattern Recognition, Proceedings., IEEE Computer Society Conference on pp. 219–224, 1997.
- [21] Scharf, L. L. and McWhorter, L., “Geometry of the Cramér-Rao Bound,” Signal Processing 31, 301–311, 1993.
- [22] Schechner, Y.Y., Kiryati, N., “The optimal axial interval in estimating depth from defocus” Computer Vision, The Proceedings of the Seventh IEEE International Conference on, Vol: 2, pp:843–848,1999.

- [23] Subbarao, M., “Parallel Depth Recovery By Changing Camera Parameters” *Computer Vision*, Second International Conference on pp.149–155
- [24] Surya, G., Subbarao, M., “Depth from defocus by changing camera aperture: a spatial domain approach” *Computer Vision and Pattern Recognition, Proceedings CVPR*, IEEE Computer Society Conference on pp.61–67, 1993.
- [25] Ten-Lee Hwang, James J. Clark, Alan L. Yuille, “A Depth Recovery Algorithm Using Defocus Information,” in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 476–482, 1989.
- [26] Vikramaditya, B. and Nelson, B.J., “Visually Guided Microassembly Using Optical Microscopes and Active Vision Techniques,” *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 3172–3177, 1997.
- [27] Watanabe, M. and Nayar, S.K., “Telecentric optics for focus analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 12 pp. 1360–1365, 1997.
- [28] Xiong, Y. and Shafer, S., “Depth from Focusing and Defocusing,” *Proceedings of the DARPA Image Understanding Workshop*, pp. 967–976, 1993.