

高等數位訊號處理

期末書面報告

Bundle adjustment 中稀疏矩陣快速求解

系級：台科大電機所碩一

學號：ntust_M10907314

姓名：張祐銓

授課教授：丁建均 教授

中華民國 110 年 6 月

摘要

光束平差法(Bundle Adjustment)是指從視覺影像中提煉出最佳的 3D 模型和相機參數(內參數和外參數)，考慮從任意特徵點發射出來的幾束光線(bundles of light rays)，它們會在幾個相機成像平面上變成像素或是檢測到的特徵點。如果我們調整(adjustment)各相機位姿和特徵點的空間位置，使得這些光線最後收束到相機的光心，就稱為光束平差法(BA)。

光束平差法中非線性優化求解的矩陣有稀疏的性質，所以透過 Schur 消去和 Cholesky 分解方式有效的求解矩陣，在大規模矩陣時效率越明顯。

關鍵字：光束平差法、稀疏矩陣、Schur、Cholesky

目錄

摘要.....	ii
目錄.....	iii
圖目錄.....	iii
第一章 緒論.....	1
1.1 光束平差法(Bundle Adjustment)簡介.....	1
第二章 Bundle Adjustment 最佳化求解.....	2
2.1 參數結構.....	2
2.2 非線性最佳化-高斯牛頓法	2
第三章 稀疏矩陣快速運算.....	3
3.1 矩陣稀疏性.....	3
3.2 快速求解.....	5
3.2.1 Schur 消去	5
3.2.2 Cholesky 分解	6
第四章 實驗結果分析.....	7
第五章 結論.....	8
第六章 參考文獻.....	8

圖目錄

圖 1 多相機位姿拍攝空間點示意圖	1
圖 2 雅克比矩陣稀疏性	4
圖 3 \mathbf{H} 矩陣稀疏性	4
圖 4 \mathbf{H} 矩陣的區域劃分	5
圖 5 \mathbf{S} 矩陣稀疏狀態	6
圖 6 相機 \mathbf{C} 數量固定 200，空間點 \mathbf{P} 數量提升所花費時間	7
圖 7 空間點 \mathbf{P} 數量固定 1000，相機 \mathbf{C} 數量提升所花費時間	8

第一章 緒論

1.1 光束平差法(Bundle Adjustment)簡介

光束指的是三維空間中的座標點投影到像平面上的光束，而重投影誤差正是利用這些光束來構建的，測量儀器精準度不夠和人為因素和外在因素影響下會造成量測誤差，而測量平差的目的是為了求得最可靠的量測成果，測量平差的原理採用的就是最小二乘法，光束平差法會從視覺中提煉出最優的 3D 模型和相機參數[2]，從每個特徵點反射出來的光束，再把相機姿態和特徵點的位置做出最優的調整之後最後收束在光心，在現代有許多 3D 視覺應用都會使用到光束平差法如 Visaul SLAM(Simultaneous localization and mapping) 和 SFM(structure from motion)。

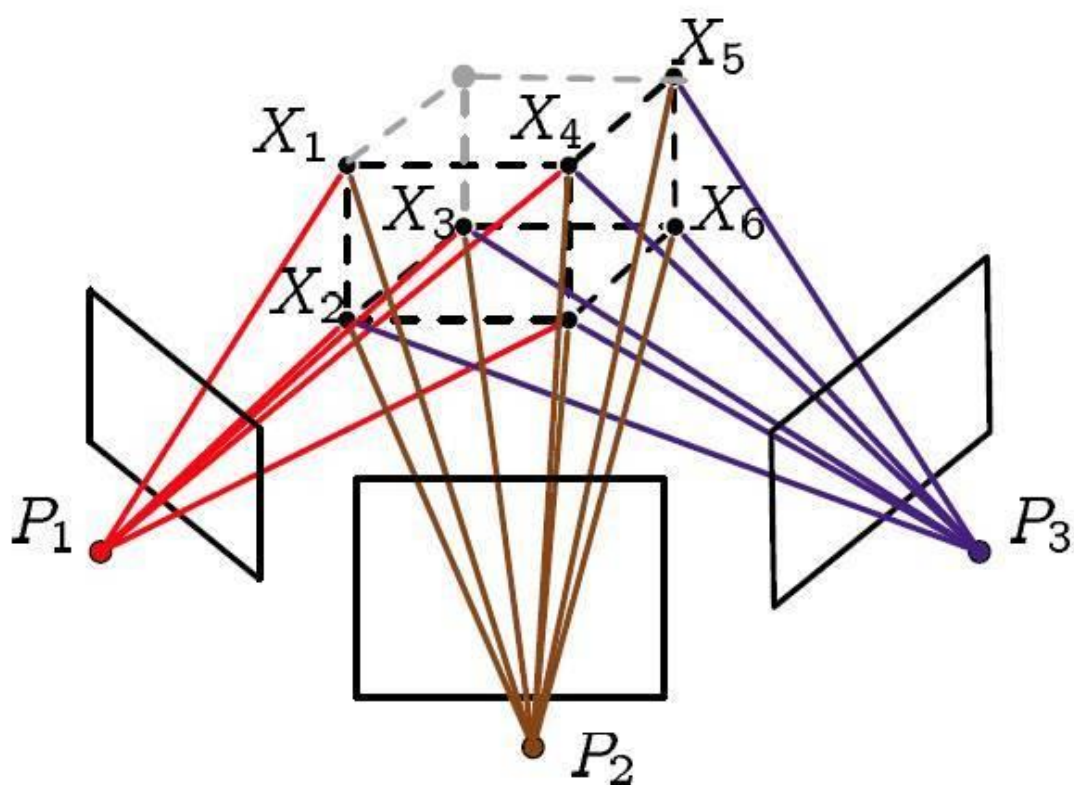


圖 1 多相機位姿拍攝空間點示意圖

第二章 Bundle Adjustment 最佳化求解

2.1 參數結構

這裡我們詳列出各參數意義， \mathbf{T} 為世界座標系至相機座標系的轉換矩陣即是位姿公式為

$$\mathbf{T} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad (1)$$

其中 \mathbf{R} 為旋轉矩陣、 \mathbf{t} 為平移矩陣， \mathbf{T} 對應的李代數為 ξ ， \mathbf{p} 為三維空間座標點，而相機所觀測到的空間座標對應的像素座標是 \mathbf{z} ， $\mathbf{h}(\mathbf{T}, \mathbf{p})$ 為相機這個位姿下 \mathbf{T} 應觀測 \mathbf{p} 對應的像素座標，以最小平方的角度來考慮，那麼可以列寫關於此次觀測的誤差為

$$\mathbf{e} = \mathbf{z} - \mathbf{h}(\mathbf{T}, \mathbf{p}), \quad (2)$$

然後，把其他時刻的量測都考慮進來，可以替誤差增加一個索引，設 \mathbf{z}_{ij} 為在位姿 \mathbf{T}_i 處觀測到 \mathbf{p}_j 產生的資料，那麼整體的代價函數為

$$\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \|\mathbf{e}_{ij}\|^2 = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \|\mathbf{z}_{ij} - \mathbf{h}(\mathbf{T}_i, \mathbf{p}_j)\|^2, \quad (3)$$

對這個最小平方進行求解，相當於對位姿和座標同時做了調整，也就是所謂的 Bundle Adjustment。

2.2 非線性最佳化-高斯牛頓法

前一節中的代價函數不是線性的函數，所以需要使用非線性的方法來優化，根據非線性優化的思想，我們會從某個初始值開始不斷尋找下降方向 $\Delta \mathbf{x}$ ，直到獲得該目標函數的最佳解，我們要最佳化的目標參數為 \mathbf{T} 和 \mathbf{p} ，所以我們將最佳化引數定義成

$$\mathbf{x} = [\mathbf{T}_1, \dots, \mathbf{T}_m, \mathbf{p}_1, \dots, \mathbf{p}_n]^T, \quad (4)$$

增量方程式中的 $\Delta \mathbf{x}$ 則是對整體引數的增量。當我們給引述一個增量時，目標函數變為

$$\frac{1}{2} \|f(\mathbf{x} + \Delta\mathbf{x})\|^2 = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \|\mathbf{e}_{ij} + \mathbf{F}_{ij}\Delta\boldsymbol{\xi}_i + \mathbf{E}_{ij}\Delta\mathbf{p}_j\|^2, \quad (5)$$

其中 \mathbf{F}_{ij} 表示整個代價函數在目前狀態下對相機位姿的偏導數，而 \mathbf{E}_{ij} 表示整個代價函數在目前狀態下對路標三維座標點的偏導數，我們將相機位姿變數和空間點各別放一起

$$\mathbf{x}_c = [\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_m]^T, \quad (6)$$

$$\mathbf{x}_p = [\mathbf{p}_1, \dots, \mathbf{p}_n]^T, \quad (7)$$

那麼式(5)就能簡化表達成

$$\frac{1}{2} \|f(\mathbf{x} + \Delta\mathbf{x})\|^2 = \frac{1}{2} \|\mathbf{e} + \mathbf{F}\Delta\mathbf{x}_c + \mathbf{E}\Delta\mathbf{x}_p\|^2, \quad (8)$$

其中雅克比矩陣可以分段為

$$\mathbf{J} = [\mathbf{F} \quad \mathbf{E}], \quad (9)$$

然後使用高斯牛頓法，其中增量方程式為

$$\mathbf{H}\Delta\mathbf{x} = \mathbf{g}, \quad (10)$$

高斯牛頓法中 \mathbf{H} 矩陣可以由 \mathbf{J} 矩陣近似

$$\mathbf{H} = \mathbf{J}^T \mathbf{J} = \begin{bmatrix} \mathbf{F}^T \mathbf{F} & \mathbf{F}^T \mathbf{E} \\ \mathbf{E}^T \mathbf{F} & \mathbf{E}^T \mathbf{E} \end{bmatrix}, \quad (11)$$

求解增量方程式(10)就可以找到下降方向增量 $\Delta\mathbf{x}$ 。

第三章 稀疏矩陣快速運算

3.1 矩陣稀疏性

當要求解增量方程式時當 \mathbf{H} 太大時求逆會消耗非常多的運算資源，所幸 \mathbf{H} 矩陣有一些稀疏的特性能夠去快速計算，這個稀疏性是由雅克比矩陣 \mathbf{J} 所造成的，考慮這些代價函數中的其中一個 \mathbf{e}_{ij} 這個誤差項只描述了在 \mathbf{T}_i 看到 \mathbf{p}_j 這件事上，只關於這第 i 個相機位姿和第 j 個空間點，其餘部分的變數導數都為0，所以該誤差項對應的雅克比矩陣有下列形式

$$\mathbf{J}_{ij}(\mathbf{x}) = \left(\mathbf{0}_{2 \times 6}, \dots, \mathbf{0}_{2 \times 6}, \frac{\partial \mathbf{e}_{ij}}{\partial \mathbf{T}_i}, \mathbf{0}_{2 \times 6}, \dots, \mathbf{0}_{2 \times 3}, \frac{\partial \mathbf{e}_{ij}}{\partial \mathbf{p}_j}, \mathbf{0}_{2 \times 3}, \dots, \mathbf{0}_{2 \times 3} \right), \quad (12)$$

整體雅克比矩陣以及對應的**H**矩陣的稀疏情況就如圖 2 和圖 3 所示

$$J = \begin{bmatrix} J_{11} \\ J_{12} \\ J_{13} \\ J_{14} \\ J_{23} \\ J_{24} \\ J_{25} \\ J_{26} \end{bmatrix} = \begin{bmatrix} \begin{matrix} C_1 & C_2 & P_1 & P_2 & P_3 & P_4 & P_5 & P_6 \end{matrix} \\ \begin{matrix} \text{blue block} & & \text{blue block} & & & & & \\ \text{blue block} & & & \text{blue block} & & & & \\ \text{blue block} & & & & \text{blue block} & & & \\ \text{blue block} & & & & & \text{blue block} & & \\ & \text{blue block} & & & \text{blue block} & & & \\ & \text{blue block} & & & & \text{blue block} & & \\ & \text{blue block} & & & & & \text{blue block} & \\ & \text{blue block} & & & & & & \text{blue block} \end{matrix} \end{bmatrix}$$

圖 2 雅克比矩陣稀疏性

$$H = J^T J = \begin{bmatrix} \text{blue block} & & \text{blue block} & \text{blue block} & \text{blue block} & & & \\ & \text{blue block} & & & \text{blue block} & \text{blue block} & \text{blue block} & \text{blue block} \\ \text{blue block} & & \text{blue block} & & & & & \\ \text{blue block} & & & \text{blue block} & & & & \\ \text{blue block} & \text{blue block} & & & \text{blue block} & & & \\ \text{blue block} & \text{blue block} & & & & \text{blue block} & & \\ & \text{blue block} & & & & & \text{blue block} & \\ & \text{blue block} & & & & & & \text{blue block} \end{bmatrix}$$

圖 3 **H**矩陣稀疏性

仔細觀察**H**矩陣會發現說左上角矩陣以及右下角矩陣皆為對角塊矩陣，我們可以把**H**矩陣分為四個區塊 **B,E,E^T,C**

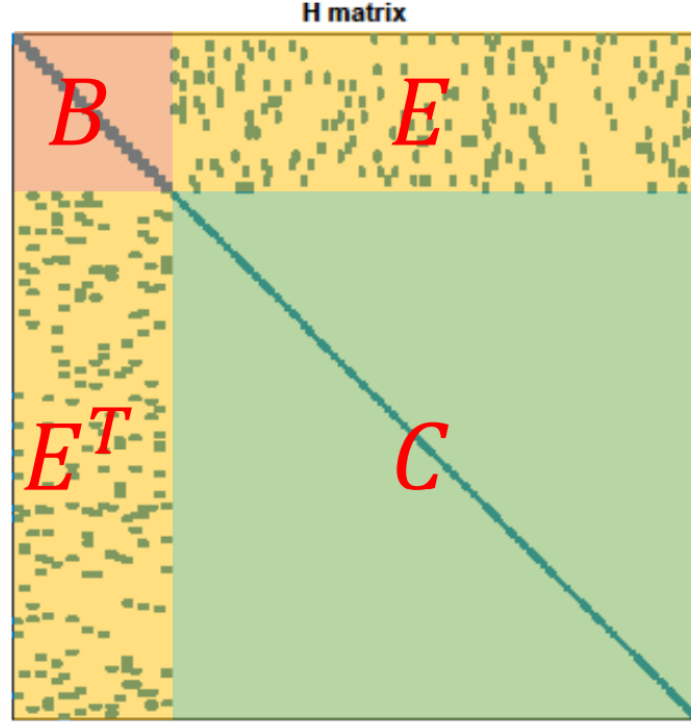


圖 4 \mathbf{H} 矩陣的區域劃分

對應的線性方程式組也可以由式(10)變化成以下形式

$$\begin{bmatrix} \mathbf{B} & \mathbf{E} \\ \mathbf{E}^T & \mathbf{C} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x}_c \\ \Delta \mathbf{x}_p \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} \quad (13)$$

3.2 快速求解

通常空間點 p 數量會遠大於相機 C 數量， \mathbf{B} 和 \mathbf{C} 皆為對角塊矩陣求逆的難度遠小於對於一般矩陣的求逆難度，因此我們只需要對那些對角線區塊分別求逆並使用 Schur 消去與 Cholesky 分解[3]加速運算。

3.2.1 Schur 消去

考慮到對線性方程式的特性我們進行高斯消除，目標是削去右上方的非對角線部分 \mathbf{E} ，得

$$\begin{bmatrix} \mathbf{I} & -\mathbf{EC}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{B} & \mathbf{E} \\ \mathbf{E}^T & \mathbf{C} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x}_c \\ \Delta \mathbf{x}_p \end{bmatrix} = \begin{bmatrix} \mathbf{I} & -\mathbf{EC}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix}, \quad (14)$$

整理可得到

$$\begin{bmatrix} \mathbf{B} - \mathbf{EC}^{-1}\mathbf{E}^T & \mathbf{0} \\ \mathbf{E}^T & \mathbf{C} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x}_c \\ \Delta \mathbf{x}_p \end{bmatrix} = \begin{bmatrix} \mathbf{v} - \mathbf{EC}^{-1}\mathbf{w} \\ \mathbf{w} \end{bmatrix}, \quad (15)$$

消去之後，方程式組第一行變成和 $\Delta \mathbf{x}_p$ 無關的項，單獨把它拿出來，獲得關於位姿部分的增量方程式

$$[\mathbf{B} - \mathbf{E}\mathbf{C}^{-1}\mathbf{E}^T]\Delta \mathbf{x}_c = \mathbf{v} - \mathbf{E}\mathbf{C}^{-1}\mathbf{w}, \quad (16)$$

這個線性方程式的維度與 \mathbf{B} 是一樣的，我們先求解出 $\Delta \mathbf{x}_c$ ，然後再將 $\Delta \mathbf{x}_c$ 代入原方程式求解 $\Delta \mathbf{x}_p$ 這個過程就稱為 Schur 消去，這個做法的優勢在於易於利用到 \mathbf{C}^{-1} 容易求解的特性來加速運算。

3.2.2 Cholesky 分解

在上節中在求解式(16)時，我們記 \mathbf{S} 矩陣為

$$\mathbf{S} = \mathbf{B} - \mathbf{E}\mathbf{C}^{-1}\mathbf{E}^T, \quad (17)$$

其稀疏狀態不規則



圖 5 \mathbf{S} 矩陣稀疏狀態

但 \mathbf{S} 矩陣是一個半正定矩陣，所以可以透過 Cholesky 方式加速求逆。

第四章 實驗結果分析

實驗中我們將比較三組不同的方式分別為直接逆矩陣運算、Schur 消去運算、Schur 消去加上 Cholesky 分解，圖 6 中固定相機 C 數量為 200，隨著空間點 p 數量越多直接逆矩陣運算離另兩種方法差距會越來越大

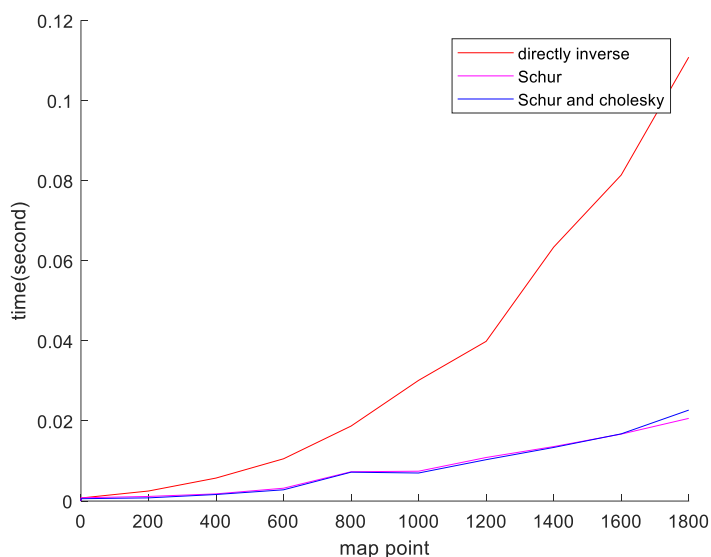


圖 6 相機 C 數量固定 200，空間點 p 數量提升所花費時間

圖 7 當中固定空間點 p 數量為 1000，相機 C 數量越多的情況下有使用 Cholesky 分解在速度上會比較快，因為當相機數量越多時 S 矩陣會越來越大 Cholesky 較一般求逆矩陣速度快但矩陣必須有正定的性質。

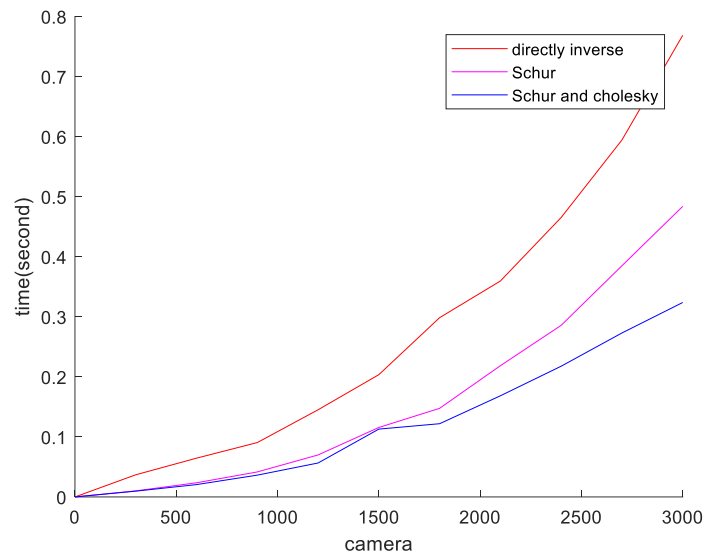


圖 7 空間點 p 數量固定 1000，相機 C 數量提升所花費時間

第五章 結論

在 3D 視覺的應用上光束平差法會因為相機移動而產生出更多的相機 C 數量及空間點 p 數量，會造成要計算的矩陣越來越龐大，藉由將 H 矩陣區域劃分並利用矩陣的稀疏性質使用 Schur 消去和 Cholesky 分解進而達到快速計算，從實驗結果得知使用本篇方式比起計算直接逆矩陣的效率要高出很多。

第六章 參考文獻

- [1] 高翔、張濤、劉毅、顏沁睿(2019 年 12 月)。SLAM 視覺十四講:雙倍內容強化版。臺北市:深智數位。
- [2] Cheng Wei's Blog , Bundle Adjustment 光束法平差。檢自 https://scm_mos.gitlab.io/vision/bundle-adjustment (June.23, 202)
- [3] Sameer Agarwal, Noah Snavely, Steven M. Seitz, and Richard Szeliski, "Bundle Adjustment in the Large"
- [4] Kurt Konolige, "Sparse Sparse Bundle Adjustment," *The British Machine Vision Conference*, doi:10.5244/C.24.102, 2010.