

行政院國家科學委員會補助專題研究計畫成果報告

立體影像之關聯性建立方法

計畫類別： 個別型計畫 整合型計畫
計畫編號：NSC 89 - 2212 - E - 034 - 010 -
執行期間：89年 8月 1日至 90年 7月 31日

計畫主持人：張珩
共同主持人：

本成果報告包括以下應繳交之附件：
赴國外出差或研習心得報告一份
赴大陸地區出差或研習心得報告一份
出席國際學術會議心得報告及發表之論文各一份
國際合作研究計畫國外研究報告書一份

執行單位：中國文化大學機械工程系

中 華 民 國 90年 9月 27日

行政院國家科學委員會專題研究計畫成果報告

立體影像之關聯性建立方法

Stereo Image Pattern Recognition Algorithm

計畫編號：NSC-89-2212-E-034-010

執行期間：89年8月1日至90年7月31日

主持人：張珩 中國文化大學機械工程系

一、中文摘要

本計劃關於立體影像(Stereo Image)之視覺辨識與立體影像合成系統之建立，以分析三度空間的對應座標位置進而建立快速計算即時 3D 系統。本計畫案中使用的視覺系統，可應用於自動導引的機械人或是運輸機械上或半導體晶圓之檢測，主要利用一對或一對以上之視覺裝置如 CCD 照相機或攝影機，在已知的兩分隔焦點距離架設此視覺系統，以便使物體成像於兩照相機或攝影機的成像平面上，藉由此兩平面所得到的影像，在已知視角與焦距條件下的兩維空間投影，求得 3D 物體在三維空間中的位置關係，建立一接近瞬時(Real Time)的視覺辨識系統。

本計劃的主要執行方法在使用多階段形式(Hierarchical)立體影像分析方法，以區域性集合(Regional Grouping)影像區域關聯特徵(Field Correlation)，例如像素(Pixels)、角度、連續或不連續的交集等，以關聯性比對(Correlation)，配合照相機的幾何位置與成像次序限制(Ordering Constraints)等，以金字塔(Pyramid)之次分割影像，進行階段式比對亮度(Luminance)特徵，計算出左、右影像中的像素(Pixel)特徵對應差值座標圖(Disparity Map)，在兩維投影的位置座標上，計算差值(Disparity)及像素亮度(Pixel luminance)，以建立三維空間的深度(Depth)相對座標關係。

本計劃中採用低頻濾波之拉普拉斯影像處理(Laplacian Image)，以消除取像時之影像雜訊問題，並以該拉普拉斯金字塔影

像解析，配合統計信賴度之分析，求得左、右影像之最小亮度差值之信賴度，對應三維座標的數學運算問題，以期能達到在瞬時(Real Time)運算下的可接受方案。

關鍵詞：立體影像、影像分析

Abstract

The study is to estimate and construct the stereo image in a binocular vision system with a real-time solution solver. For the vision systems used in the automatic guided vehicles (AGVs) or robotic transportation systems, it is commonly employed two cameras (or video cameras) to form a binocular view. Each picture taken by the camera represents a projection of the object(s) in 2D space. In order to obtain the depth information (or stereograph), it is required to analyze these 2D pictures taken by the two different viewing position cameras.

The method of the study is to implement a field correlation among the left and right images acquired from two positioned cameras, such that each pixel on the image plane are analyzed with intersection angles, and continuous feature sets, from each projection picture in a multi-stage stereo correspondence procedure. By confining windows from coarse to fine, we apply a masked correlation step using the Laplacian image pyramids correspondence schemes on both images to find the corresponding luminance intensities on both images. Thus,

the 3D-depth information referenced in the global coordinate system can be calculated according to the camera parameters. Those data can then be projected onto the ground plane and a disparity map is obtained. It is the aim of this study to implement the stereo correspondence algorithm for the real-time AGV avoidance system.

Keywords: Stereo Image, Stereo Matching Algorithm

二、緣由與目的

本計劃所探討之問題，廣泛的在 80 年代被討論，有關的早期資料源於電腦的數位影像分析技術，在 1979 年由 Marr 與 Poggio[1]所發表的論文中提到基本的人類三維空間視覺的計算基礎，其中最重要的兩項規則為：(1)在成像的兩維圖片中，每一點幾何特徵僅包含一個深度值，(2)兩相鄰點幾何特徵的深度值幾乎相等。因此建立了若要在兩維的幾何圖形中建立原始三維物體的深度立體資料，則最少需要兩只照相機，以不同的觀察角度或位置，成像於兩個相對的兩維平面圖形上。

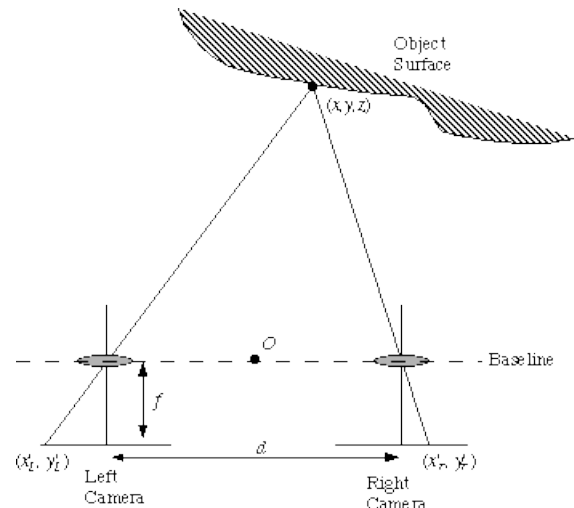
於 1982 年，Barnard 與 Fischler[3]對於立體組合計算中所提出的步驟指出，在人工視覺系統中應包括：(1)影像擷取，(2)照相機幾何模式設定，(3)圖像幾何特徵擷取，(4)影像配對(立體特徵配對)，(5)特徵深度值計算，(6)幾何特徵內插法運算。影像擷取模式包括；(A)兩只照相機，其光學軸線相互平行且相距 d ，(B)兩只照相機的鏡頭中心連線垂直於其光學軸線，稱之為基線，(C)令視界的 X 軸座標平行於基線，(D)且令原點 O 位於兩照相機的鏡頭中心的中點距離處。如圖一所示。

考慮真實物體座標位置 (x, y, z) 在左、右照相機成像面投影之座標值分別為 (x'_l, y'_l) 與 (x'_r, y'_r) ，則依據三角幾何比例可知，當焦距 f 已知時，可由平面投影位置 (x', y')

以下式求得深度 z 值：
$$z = \frac{d \cdot f}{x'_l - x'_r}$$
，其中

$x'_l - x'_r$ 稱之為插值(disparity)，是決定真實

物體座標位置 (x, y, z) 的主要關鍵。



圖一、兩只相機擷取立體影像之示意圖

兩立體影像間之影像物件比對乃是藉由影像上的每一像素對像素間，建立之差值向量(Disparity vectors)所形成，該差值向量連結左、右影像上之二維立體投影物件，且關聯物件實體與相機間之三維座標之深度計算。

Williams 等於 1986 年，提出對於動態影像的控制與策略，其中主要步驟乃是將影像進行階段式(Hierarchical)分割，以減少影像的解析度形成較粗之次影像，且首先於粗影像上進行影像差值(Disparity)之估計，然後再將此估計值衍生至較細緻之次影像，做更精確之估算，該方法中使用單一金字塔影像架構，將影像分割為不同解析度之次影像，藉以建立動態影像之運動向量。

本案中所使用之方法，首先將兩配對之平面影像進行金字塔影像分割；該原始左、右影像經過一低頻濾波器，產生一次影像，而該原始影像即與該次影像相減，以產生一高頻影像，該高頻影像即稱為拉普拉斯影像(Laplacian image)；該次影像接續將鄰近之每隔一行列之像素消除，建立一解析度較低之次影像，並以相同方式低頻濾波該次影像，並相減後得一次拉普拉斯影像，如此反覆進行該步驟，直至所需之頂端金字塔影像之解析度為止。

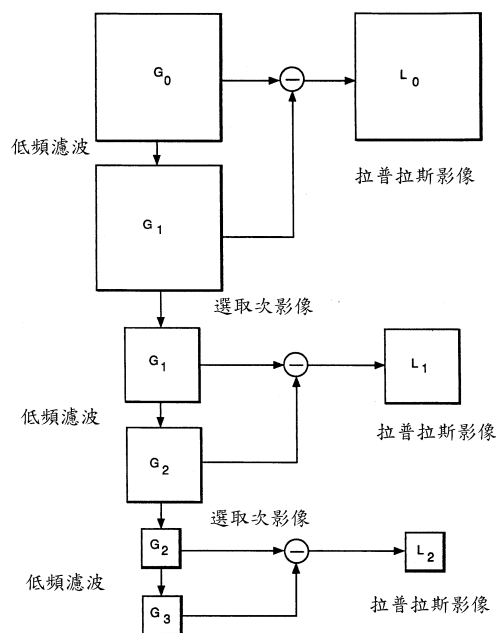
於第二步驟中，首先針對金字塔頂端之

最粗解析度之拉普拉斯次影像進行影像差值估算，該估算值再經過次一級較高解析度之拉普拉斯次影像之精算預估，該步驟反覆進行直至原始影像形成之拉普拉斯影像為止；每一步驟之中該差值向量 (Disparity vector) 之定義為左影像之像素座標為零點而指向相對應之右影像像素座標，亦即以右影像之座標值減去左影像之座標值。

此階段式比對之優點在於每一像素之差值向量計算可時間可相對精簡，並同時建立影像之配對像素，且由反覆計算結果證明，對於影像中背景雜訊之干擾可降至最低，結果可以得到更精確之像素配對。

三、結果與討論

本研究計劃中左右影像之拉普拉斯金字塔影像之建立方法如圖二所示：



圖二、拉普拉斯影像計算之方塊圖

輸入之影像以 G_0 表示，該 G_0 影像經過一低頻濾波處理後，產生另一影像 G_1 ，該濾波影像 G_1 即與 G_0 相減後，留下影像

L_0 ，影像 L_0 保存有高頻物件而減少背景及雜訊之發生，此影像即稱為拉普拉斯影像。

下一步驟該 G_1 影像將隔行與隔列之像素相消除，以產生一次影像 G_1 為原來影像解析度之四分之一，該次影像 G_1 同時再經過一低頻濾波處理，以產生低頻影像 G_2 ，該影像 G_2 再與次影像 G_1 相減，而產生次拉普拉斯影像 L_1 ；該影像 G_2 可再重複以上步驟，經減低解析度後產生次影像，並經濾波處理且與次影像 G_2 相減後，以產生第三層之次拉普拉斯影像 L_2 。此金字塔影像產生之處理步驟可不斷重複，直至所需之金字塔頂端解析度為止。

若以處理 480×512 之原始影像為例， L_3 之影像解析度將為 60×64 ，即可適用於立體影像之像素差值比對計算之用。

影像分析方法首先針對最粗解析度之拉普拉斯次影像計算出整體之像素差值，再進一步以該差值連結次一階段之較高解析度之拉普拉斯次影像進行像素差值估

算，於一階段 $\frac{1}{2^n}$ 解析度次影像經由相鄰之

粗 $\frac{1}{2^{n+1}}$ 解析度次影像轉換估算像素差值

時，僅需將該差值向量乘上 2 以補償金字塔分割時解析度減少之因素；該差值向量可延像素位置的四個方向出現於 $\frac{1}{2^n}$ 解析度次影像，故增加準確度之考量下，該像素差值向量之預測值應有四個，亦即於 $\frac{1}{2^n}$ 解析度次影像之每一像素藉由粗 $\frac{1}{2^{n+1}}$ 解析度次影像相鄰之四個像素差值向量計算而

得其差值向量估計值，故於 $\frac{1}{2^n}$ 解析度次影像之每一像素僅需做一小幅度修正搜尋其正確之差值向量。

該差值向量可以下式表示：

$$\vec{d} = P + u$$

其中， \vec{d} 為差值向量、 P 為差值預測值、 u 為修正值，其搜尋方向可為水平或垂直或兩者並行，對理想之立體影像而言，該差值向量均延水平方向改變，故其

搜尋均以水平方向為主，在訂定已知之差值向量最小值 d_{min} 與差值向量最大值 d_{max} 之前提下，該左影像之素像將進行區域之像素移動並與右影像產生像素差值，該位移像素整體差值之平方最小值，將被選為適當之估計值。

對像素差值向量之選擇，在實際影像中牽涉統計數學之應用，由於影像與背景及雜訊間的相互干擾，使得像素差值向量需藉由統計方法之最大可能性或貝氏定理計算所修訂；Ohta and Kanade 於 1985 年提出使用 Dynamic Programming 搜尋像素配對之方法；該統計數學之分析將於明年度之計畫案中提出申請。

當以上步驟反覆進行，直至金字塔之最底部，亦即該原始影像之拉普拉斯影像時，影像中每一像素差值向量即可計算得知，此時即可以此差值向量圖(Disparity map)建立三維影像或計算影像中每一物件與相機之距離，用以測距或障礙迴避之用途。

四、計劃成果自評

本案建立一三維立體影像關聯性技術之分析方法，該方法包括使用兩只相機建立影像系統，兩只相機以理想之等高線設置，且相距一適當距離並具有相同焦距，三維物件設定於兩相機之可清晰之感測距離內，同時呈像於兩相機之呈像平面。

該左、右影像像素為 480×512 之解析度以下，首先對左、右影像進行低頻濾波以產生拉普拉斯之影像轉換，再將濾波之低頻影像，於每間隔行列消除像素，取得該影像之次影像解析度，並再對該次影像進行低頻濾波與產生拉普拉斯之次影像，如此反覆進行次影像之產生，而組成一金字塔之影像結構。

第二步驟進行金字塔頂端粗解析度之次影像分析各像素之差值向量，該方法使用一預設之差值向量最大及最小值 d_{min} 與 d_{max} ，將該左次影像水平位移並與右次影像相減，該整體差值平方紀錄之，並重複此步驟，紀錄每一位移之整體像素差值平方總值，再依循選取像素差值之最小平方

和，為該影像之像素差值向量者。

第三步驟進行下一階段之次影像差值向量估算，該次影像之差值向量由前一階段之差值向量估算後，並配合以一小視窗搜尋其像素之補償值，以求得更精確之結果，並同時對整體差值進行最小平方和處理，並求得每一像素之差值向量。

重複以上步驟，直至金字塔底部之原始拉普拉斯影像，並產生差值向量圖(Disparity map)，以建立三維影像之應用。

本計畫案中使用之三維視覺分析方法，可應用於自動導引的機械人或是運輸機械上或半導體晶圓之檢測，使用階段式金字塔分割影像之處理方法，可有效節省左右影像之像素配對時間，且經由拉普拉斯影像之應用，可有效消除雜訊與背景影響，並可有效針對物件輪廓、外型表現顯著性；電腦模擬計算下可有效的分析室內影像檔案，若能結合工業用之影像處理軟體必能測試本案所提出之研究方法在瞬時系統(Real time)下之可行性；故提出本報告的各節結論。

五、參考文獻

1. D. Marr, and T. Poggio, "A Computational Theory of a Human Stereo Vision," Proc. Royal Society of London, B(204), pp. 301-328, 1979.
2. D. Marr, , Vision, W.H. and Freeman Co., 1982.
3. S.T. Barnard, and M.A. Fischler, "Computational Stereo," Comp. Surveys, vol. 14, no. 4, pp. 553-572, Dec. 1982.
4. Williams et al., "A Coarse-to-Find Control Strategy for Stereo and Motion on a Mesh-Connected Commuter," Proceedings of the Conference on Computer Vision and Pattern Recognition, pp 219-226, IEEE, June 1986.
5. T. Kanade and M. Okutomi, "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, No. 9, pp. 920, Sept. 1994.
6. Ohta, Y. and Kanade T., "Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming," IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-7(2):139-154, March 1985.
7. Nishihara, H.K., "Practical Real-Time Imaging Stereo Matcher," Optical Engineering, vol. 23, No.5, Oct. 1984.