

使用二維人耳影像的一般化霍夫轉換實作個人身分辨認

盧傳傑^a、曾建誠^{*a}、李素玲^b
國立高雄第一科技大學電通系^a
長榮大學資工系^b

摘要 — 現今，生物特徵在個人身分辨認上是廣泛應用的，這些生物特徵包括人臉、指紋、掌紋、虹膜和脫氧核糖核酸(DNA)等，為了提高身分驗證、識別及辨認的安全性，本論文將使用人耳做為生物特徵進行身分辨認，因為人耳在生物特徵上是易見、易取得、穩定度高且不易受到時間性的影響改變。本論文將研究如何使用二維(2-D)人耳影像的一般化霍夫轉換(Generalized Hough Transform, GHT)作為開發身分辨認的方法。首先，將對二維人耳影像前處理的方法，進行說明，其次，使用可以檢測任意形狀物件的一般化霍夫轉換的方法作為開發人耳識別的方法。最後，使用自己建立的人耳資料庫進行測試，實驗結果顯示，所提出的人耳識別方法比傳統的豪斯多夫距離的方法具有較高的正確率。

一、簡介

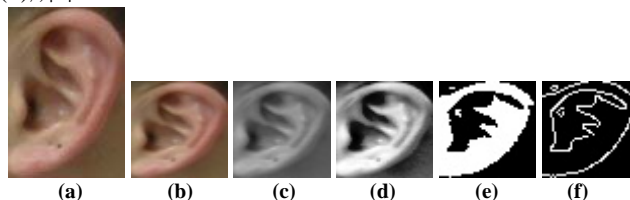
近幾年，人耳辨認、辨識系統在文獻上有逐年增加的跡象，這也顯示著在生物辨認、辨識上，相關研究人員開始發現利用人耳來進行個人身分辨認、驗證，是一件既安全、又方便的事情，人耳在個人身分辨認、驗證上可以區分為二維的人耳影像及三維的人耳影像，而二維的人耳影像所取得的方式較為容易，只要使用一台攝影機拍攝即可，再將拍攝後的人耳影像做後續處理，而人耳定位、及人耳的特徵提取在人耳辨認中是重要的一環，在目前的文獻方法針對特徵提取的方法上在[1]使用SIFT Flow(Scale-invariant feature transform)對人耳做特徵提取來辨認人耳，[2]使用影像像素來特徵提取，[3][4]的主成分分析(Principal Component Analysis, PCA)，[5]使用獨立成分分析(Independent Component Analysis, ICA)，[6]使用邊緣交互點偵測(Edge Interaction Point Detection, EIPD)來提取人耳及特徵提取，[7]使用方向梯度直方圖(Histogram of Oriented Gradients, HOG)來有效解決人耳檢測和識別的問題。

本篇論文將探討二維人耳影像的辨認問題，使用的方法為一般化霍夫轉換，目前一般化霍夫轉換已經應用在許多問題上，如[8]使用在偵測任意形狀上，[9]使用在視覺樣本比對之物件檢索方法中，[10]應用在超音波醫學影像上，[11]將二維的一般化霍夫轉換延伸至三維空間進行人行偵測，因為一般化霍夫轉換可以偵測任意的形狀之物件，故已廣泛的運用在不同的影像處理、辨認、辨識上面，本篇論文也使利用其特殊性，應用在基於人耳影像的身分辨認上。

二、人耳影像前處理方法

2.1 影像正規化

本研究的人耳辨認中，人耳的擷取是由相機擷取，在對取影像作分析，由於人耳影像在進行影像拍攝擷取時，影像會因為拍攝距離、拍攝角度或是每個人的人耳大小而造成擷取的人耳影像，其大小皆不一樣，所以將拍攝後的影像資料庫進行人耳定位後，並加以進行影像正規化，是必須的，因為這樣有利於往後辨認演算法之發展，通常在辨認系統中，辨認時間的長短往往是很重要的，而通常會以事先制定的長寬尺寸進行正規化，並將正規化後的影像放入資料庫中，如此一來可以節省系統消耗資源，加速辨認的處理，在本論文使用人機交談方式來從拍攝之影像裁切出人耳影像，並將裁切後的輸入人耳影像大小設定為 48×48 ，正規化的結果如圖一(b)所示。



圖一：(a)原始影像(b)影像正規化(c)灰階影像(d)影像增強(e)閾值擷取(f)邊緣偵測

2.2 灰階影像

在計算機領域中，灰階(Gray scale)影像通常顯示為從最暗黑色到最亮的白色的灰階，通常每個採樣像素用8位元(bits)的來儲存，這樣可以有256種灰階。這種精度剛剛能夠避免可見的條帶失真，並且非常易於編程。

假定有一張數位影像是以灰階的方式呈現，則每一個像素可以有256種變化，最亮是為白色，以數值255表示之。最暗的顏色是為黑色，以數值0表示之。介於黑與白之間的顏色稱為灰色，而灰色依明暗的不同可以有254種變化，以數值1到254來表示之。本論文之彩色人耳影像經由灰階化的結果如圖一(c)所示。

2.3 影像增強

直方圖等化是影像處理領域中利用影像直方圖對對比度進行調整的方法。這種方法通常用來增加許多影像的全局對比度，尤其是當影像的有用數據的對比度相當接近的時候。通過這種方法，亮度可以更好地在直方圖上分佈。這樣就可以用於增強局部的對比度而不影響整體的對比度，直方圖等化通過有效地擴展常用的亮度來實現這種功能。

本論文是以人耳的邊緣影像來進行演算法的計算處理，所以當邊緣越清晰、越明顯結果也會更好。為了讓邊緣

更清楚，讓接下來的辨認更精準，本論文使用直方圖等化來增強影像對比，使邊緣能夠更完整，由本論文之實驗結果發現，在影像增強之後的驗證辨認結果會更加準確，使用直方圖等化使影像增強的結果如圖一(d)所示。

2.4 閾值擷取

影像閾值擷取(Thresholding)，是將二值化影像做分類，將像素分成二個群組，一個群組為像素灰度值大於閾值，另一群組為像素灰度值小於或等於閾值，即可以將像素轉為黑色或白色，而閾值擷取在影像中，擷取某個物體是相當重要的，因為它可以分割影像中，不必要或是不重要的資訊，用來簡化影像，一般來說，只對灰階影像作閾值擷取便可完成。假設灰階影像 $f(x, y)$ ， T 為影像二值化用來切割的門檻值， $g(x, y)$ 是影像閾值擷取後之影像，將灰階人耳影像閾值擷取如方程式(1)所示：

$$g(x, y) = \begin{cases} 255 & \text{if } f(x, y) > T \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

經由上述方法，利用門檻值 T 可以將灰階影像切割成一張二元(黑白)影像，大於門檻值 T 便會被保留下來，小於等於則捨去，所以門檻值 T 的不同可使不必要的資訊得以濾除，其影像閾值擷取的結果如圖一(e)所示。

2.5 邊緣偵測

尋找影像邊緣時，可以使用 Canny 邊緣偵測法或是 Sobel 邊緣偵測法等，本論文所使用的邊緣偵測方法為 Canny 邊緣偵測，Canny 在 1986 年提出邊緣偵測濾波器的評估準則[12]，其設計在於滿足邊緣偵測的三項要求：

- 1.降低偵測錯誤率：能偵測到所有邊緣，並且只有偵測到邊緣。
- 2.邊緣定位：影像中實際的邊緣與演算法所計算出的邊緣之間的距離應縮減至最小。
- 3.單一反應：指的就是只有單一邊緣存在，演算法得出的結果不應為多個邊緣像素。

由於影像中的各個像素點之像素值可能不同而造成明暗變化，當像素值差距大時便會在影像中形成邊緣輪廓，使用 Canny 演算法來對影像進行邊緣偵測的結果如圖一(f)所示。

三、一般化霍夫轉換

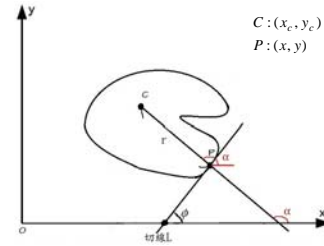
3.1 物件形狀描述

在一般化霍夫轉換中，物件的形狀將使用 R-table 來加以表示和描述。由於不規則無法用數學公式加以表示，因此，我們使用 R-table 來描述物件的重心 C (centroid)和物件邊界 P (boundary)的幾何關係，如圖二所示，若重心 C 的座標為 (x_c, y_c) ，邊界點 P 的座標為 (x, y) 則由 $\vec{OC} = \vec{OP} + \vec{PC}$ 可得方程式(2)：

$$\begin{bmatrix} x_c \\ y_c \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} r \cos \alpha \\ r \sin \alpha \end{bmatrix} \quad (2)$$

其中， r 是 \vec{CP} 線段的長度， α 是 \vec{PC} 和 \vec{OX} 向量的夾角。再者，若 P 點的切線 L 與 \vec{OX} 的夾角為 ϕ ，則切線的斜率 (slope) 為 $\tan \phi$ ，即切線的方向向量為

$[\cos \phi, \sin \phi]^T$ ，由於在 P 點之梯度向量 (gradient vector) 和切線方向向量垂直，故 P 點的梯度向量方向為 $[\sin \phi, -\cos \phi]^T$ ，故可由計算 P 點的梯度向量後，再由其方向向量來計算角度 ϕ 。



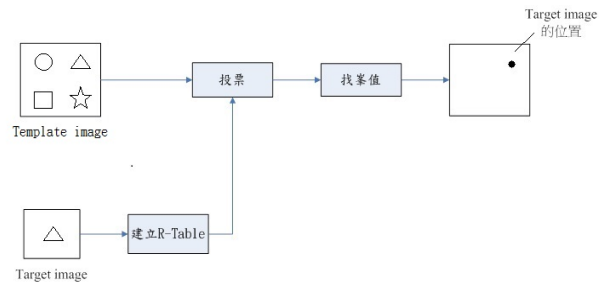
圖二：R-Table 建構的參數說明圖

當 P 點在物體邊界中，繞一圈時，每個邊界點 P 均提供一組參數 ϕ, α, r ，若將 ϕ 量化成 M 個代表值 $\phi_1, \phi_2, \dots, \phi_M$ ，則可以利用 ϕ_i ($i = 1, 2, \dots, M$) 當索引 (index) 將 (r, α) 參數整理成下列 R-table 來描述物件的邊界形狀：

- $\phi_1 : (r_1^1, \alpha_1^1), (r_1^2, \alpha_1^2), (r_1^3, \alpha_1^3) \dots (r_1^{n_1}, \alpha_1^{n_1})$
- $\phi_2 : (r_2^1, \alpha_2^1), (r_2^2, \alpha_2^2), (r_2^3, \alpha_2^3) \dots (r_2^{n_2}, \alpha_2^{n_2})$
- $\phi_3 : (r_3^1, \alpha_3^1), (r_3^2, \alpha_3^2), (r_3^3, \alpha_3^3) \dots (r_3^{n_3}, \alpha_3^{n_3})$
- $\phi_4 : (r_4^1, \alpha_4^1), (r_4^2, \alpha_4^2), (r_4^3, \alpha_4^3) \dots (r_4^{n_4}, \alpha_4^{n_4})$
- \vdots
- $\phi_M : (r_M^1, \alpha_M^1), (r_M^2, \alpha_M^2), (r_M^3, \alpha_M^3) \dots (r_M^{n_M}, \alpha_M^{n_M})$

3.2 物件偵測

在 1962 年，Hough 所提出的霍夫轉換是用來偵測影像中直線，後來被許多研究學者括展成可偵測影像中的圖形、橢圓形和拋物線等，直到 1981 年，Ballard 提出一般化霍夫轉換 (Generalized Hough Transform, GHT) 來偵測影像中一般化形狀的物件，由於一般化形狀，無法用參數化的數學式加以表示，故 Ballard 使用前一節所描述的 R-Table 來表示一般化形狀的物件，然後在影像中使用投票法 (voting method) 來偵測物件，此方法的細節描述如圖三所示：



圖三：一般化霍夫轉換的物件偵測法

由上述範例的說明，歸納一般化霍夫轉換的物件偵測方法之步驟如下：

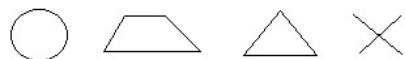
給定 target image 和 template image，且 template image 的大小為 $N_1 \times N_2$ ，偵測步驟如下：

- Step1. 找出 target image 的 R-Table。
- Step2. 建立投票矩陣 V ，其大小為 $N_1 \times N_2$ ，且初始值均設定成零。

Step3. 對 template image 中的每一個 edge 像素，執行下列投票步驟：

- (1)計算 edge 像素切線的方向角度，令為 θ° 。
- (2)找出 R-Table 中 ϕ_θ 的向量序列 S 。
- (3)若 edge 像素的位置為 (x, y) ，則在 $(x, y) + S$ 的位置中，各增加一票，而得到新的計票矩陣 V 。

Step4. 找出投票矩陣 V 的最大峯值，令其位置為 (x_t, y_t) ，當最大峯值的票數夠高，則 target image 在 template image 中存在，且位在 (x_t, y_t) 的位置上。

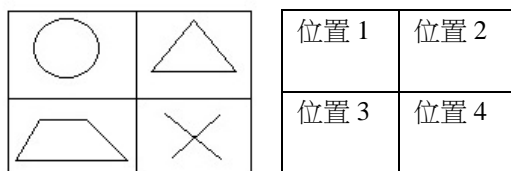


圖四：欲辨識的四種形狀

3.3 物件辨認

最後說明如何將一般化霍夫轉換應用在圖形識別 (pattern recognition) 上，以形狀辨識說明如下：當輸入一張待辨識的 image T，如何決定其形狀為圖四中的那一個形狀呢？使用一般化霍夫轉換解此辨認問題的步驟如下：

Step1. 將圖四的四種形狀建成 template image 如圖五(a)，其對應的位置為圖五(b)：



圖五：(a)template image(b)對應位置

則可建成對應表格如下：

- 位置 1 ↔ 圓形
- 位置 2 ↔ 三角形
- 位置 3 ↔ 梯形
- 位置 4 ↔ 叉型

Step2. 利用一般化霍夫轉換找出待辨識 imageT 在 template image 中的位置 (x_t, y_t) 。

Step3. 由位置 (x_t, y_t) 和對應表格，便可知道 imageT 為何形狀。

本論文的目的，便是使用上述辨識方法來辨識二維人耳影像，進而決定人的身分識別。

四、實驗結果與討論

本論文實驗的人耳影像資料庫採用自行拍攝建構的人耳影像資料庫，每人有兩種距離(30 公分和 50 公分)、五種角度(頭部水平直視前方、頭部向上傾斜約 20°、頭部向下傾斜約 20°、頭部向左傾斜約 20°、向右傾斜約 20°) 每種距離與角度拍攝三張共 30 張，拍攝 40 人，耳朵影像資料庫共有 1200 張影像。再者，當距離為 30 公分時的五種角度之第三張影像來建立樣板，供驗證系統使用，其餘影像皆為待測影像，因此待測影像有 1000 張，樣板影像有 200 張。在 GHT 影像樣板中，所有影像經正規化後影像解析度為 48×48，故集結後之 GHT 樣板影像的大小為 960×480，本論文所用之

人耳辨認系統之 GHT 樣板和對應人的編號如圖六(a)(b)所示。使用 GHT 之人耳辨認系統之架構流程如圖七所示。

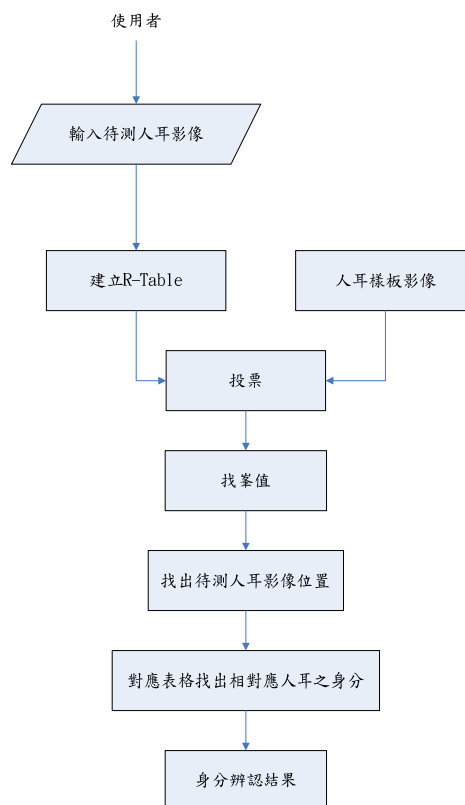


(a)

1	1	1	1	1	2	2	2	2	2	3	3	3	3	3	4	4	4	4	4
5	5	5	5	5	6	6	6	6	6	7	7	7	7	7	8	8	8	8	8
9	9	9	9	9	10	10	10	10	10	11	11	11	11	11	12	12	12	12	12
13	13	13	13	13	14	14	14	14	14	15	15	15	15	15	16	16	16	16	16
17	17	17	17	17	18	18	18	18	18	19	19	19	19	19	20	20	20	20	20
21	21	21	21	21	22	22	22	22	22	23	23	23	23	23	24	24	24	24	24
25	25	25	25	25	26	26	26	26	26	27	27	27	27	27	28	28	28	28	28
29	29	29	29	29	30	30	30	30	30	31	31	31	31	31	32	32	32	32	32
33	33	33	33	33	34	34	34	34	34	35	35	35	35	35	36	36	36	36	36
37	37	37	37	37	38	38	38	38	38	39	39	39	39	39	40	40	40	40	40

(b)

圖六：(a)人耳辨認系統之 GHT 樣板(b) 對應人之編號



圖七：人耳辨認系統之 GHT 架構流程圖

人耳辨認結果如表 I 所示，Accuracy Rate 為準確率，定義為人耳辨認系統中，(使用者進行身分辨認成功辨認之人數)/(使用者進行身分辨認之樣本總數) $\times 100\%$ 。；False Acceptance Rate(FAR)為接受誤差率，是指非使用者 n 辨認成使用者 n 證的比率，定義為人耳辨認系統中，(非使用者 n 辨認成使用者 n 之人數/非使用者 n 進行身分驗證之樣本數) $\times 100\%$ 。通常我們希望 Accuracy rate 愈高愈好，FAR 愈低愈好，實驗結果顯示所提 GHT 方法之平均 Accuracy rate 為 90.9%，平均 FAR 為 0.231%。

表 I
人耳辨認系統之使用 GHT 實驗結果

	Accuracy Rate	FAR
Person 1	88%(22/25)	0.205%(2/975)
Person 2	96%(24/25)	0.103%(1/975)
Person 3	92%(23/25)	0%(0/975)
Person 4	100%(25/25)	0.513%(5/975)
Person 5	72%(18/25)	0.205%(2/975)
Person 6	92%(23/25)	0.205%(2/975)
Person 7	84%(21/25)	0.103%(1/975)
Person 8	88%(22/25)	0%(0/975)
Person 9	76%(19/25)	0%(0/975)
Person 10	100%(25/25)	0%(0/975)
Person 11	100%(25/25)	0.410%(4/975)
Person 12	88%(22/25)	0%(0/975)
Person 13	80%(20/25)	0.205%(2/975)
Person 14	92%(23/25)	0.308%(3/975)
Person 15	88%(22/25)	0.821%(8/975)
Person 16	88%(22/25)	0.410%(4/975)
Person 17	92%(23/25)	0.513%(5/975)
Person 18	100%(25/25)	0.410%(4/975)
Person 19	96%(24/25)	0.205%(2/975)
Person 20	64%(16/25)	0.103%(1/975)
Person 21	96%(24/25)	0.205%(2/975)
Person 22	100%(25/25)	0.410%(4/975)
Person 23	96%(24/25)	0.205%(2/975)
Person 24	96%(24/25)	0%(0/975)
Person 25	92%(23/25)	0%(0/975)
Person 26	88%(22/25)	0.103%(1/975)
Person 27	92%(23/25)	0.103%(1/975)
Person 28	72%(18/25)	0.513%(5/975)
Person 29	100%(25/25)	0%(0/975)
Person 30	84%(21/25)	0.513%(5/975)
Person 31	100%(25/25)	0%(0/975)
Person 32	100%(25/25)	0.308%(3/975)
Person 33	88%(22/25)	0%(0/975)
Person 34	84%(21/25)	0.205%(2/975)
Person 35	80%(20/25)	0.205%(2/975)
Person 36	100%(25/25)	0.718%(7/975)
Person 37	100%(25/25)	0%(0/975)
Person 38	100%(25/25)	0.308%(3/975)
Person 39	100%(25/25)	0.410%(4/975)
Person 40	92%(23/25)	0.308%(3/975)
Average	90.9%(909/1000)	0.231%(90/39000)

表 II
與過去方法[13]比較

方法	Accuracy Rate
Generalized Hough Transform	90.9%(909/1000)
Hausdorff Distance	86%(860/1000)

表 II 所示，為本論文所提的一般化霍夫轉換方法和[13]所提的豪斯多夫距離(Hausdorff Distance)方法之比較，由於 GHT 方法的 Accuracy rate 較高，故可以說明

本篇論文所提的一般化霍夫轉換法，在人耳身分辨認上是較可行的。

五、結論與未來研究

本論文利用人耳影像的一般化霍夫轉換來辨認人耳，實驗結果有著不錯的效果，在人耳影像前處理的部分是非常重要的，因為不夠細緻化的邊緣影像或是耳朵輪廓內部紋理都可能影響後續的辨認的正確率，論文中實驗結果錯誤率較高的部分都可以藉由增加影像樣版來降低，或是藉由側臉傾斜角度的統一、與定位前頭髮的梳理(減少毛髮散亂干擾)來獲得改善。

除了使用一般化霍夫轉換來辨認人耳外，未來亦可利用尺度不變特徵變換(Scale-invariant feature transform, SIFT)演算法來辨識人耳，測試看是否可以再提升 Accuracy rate。

本論文是使用二維影像，也可以改採用三維人耳影像，增加了人耳影像景深資訊後，提供人耳辨識系統更多特徵，提升系統驗證或辨認能力，以上都是接下來可以努力改進的方向。

參考文獻

- [1] Nejadi H., Li Zhang, Sim, T., Martinez-Marroquin E., Guo Dong, "Wonder Ears: Identification of Identical Twins from Ear Images," International Conference on Pattern Recognition, pp. 1201 - 1204, 2012.
- [2] Kumar, P.R., Dhenakaran, S.S. "Pixel based feature extraction for ear biometrics," International Conference on Machine Vision and Image Processing, pp. 40 - 43, 2012.
- [3] Tharwat, A., Ibrahim, A., Ali, H. A., "Personal Identification Using Ear Images Based on Fast and Accurate Principal Component Analysis," 8th International Conference on Informatics and Systems, pp. MM-56 - MM-59, 2012.
- [4] K. Chang, B. Victor B., K.W. Bowyer, S. Sarkar, "Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 8, pp. 1160-1165, 2003.
- [5] Kurmiawan, F., Shafry, M., Rahim, M., "A Review on 2D Ear Recognition," IEEE 8th International Colloquium on Signal Processing and its Applications, pp. 204 - 209, 2012.
- [6] Gnanasivam, P., Muttan, S., "Ear and fingerprint biometrics for personal Identification," International Conference on Signal Processing, Communication, Computing and Networking Technologies, pp. 347-352, 2011.
- [7] Damer, N., Fuhrer, B., "Ear Recognition Using Multi-Scale Histogram of Oriented Gradients," Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 21-24, 2012.
- [8] D. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," Pattern Recognition, Vol. 13, No. 2, pp. 111-122, 1981.
- [9] Cheng, Shyi-Chyi, Kuo, Chen-Tsung, Chen, Hong-Jay, "Visual object retrieval via block-based visual-pattern matching," Pattern Recognition, Vol. 40, Issue 6, pp. 1695-1710, 2007.
- [10] 吳宇翔(民國100年)。演化式特徵建構技術應用於B模式超音波醫學影像之自動肝器官疾病診斷。國立臺灣海洋大學資訊工程學系碩士學位論文。
- [11] 羅永淇(民國100年)。應用立體樣式比對及三維行為樣版之人行為偵測方法設計。國立臺灣海洋大學資訊工程學系碩士學位論文。
- [12] Canny J., "A Computational Approach to Edge Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, Issue 6, pp. 679-698, 1986.
- [13] 賴宣帆(民國100年)。使用人耳二維影像實作個人身分驗證與辨認。國立高雄第一科技大學電腦與通訊工程學系碩士學位論文。